FACULTY OF SOCIAL SCIENCES
UNIVERSITY OF COPENHAGEN

PhD Thesis

Tseday Jemaneh Mekasha

Macroeconomic Shocks and Inflation, Social Protection and Aid Effectiveness

Copenhagen University,
Department of Economics

Advisors: Finn Tarp
Henrik Hansen

Submitted: November 2015

Preface

There are many people who have contributed to the completion of my PhD and to whom I am greatly indebted to.

I would first like to thank my academic advisors Finn Tarp and Henrik Hansen, who have been very helpful and accessible. Thank you for being so generous with your time and I am very grateful for your unswerving support and advice throughout this work. I have benefited a lot from your comments and feedback on my work. Thank you!

I have also benefited a lot from my one year research visit at the University of California, Berkeley Department of Agriculture and Resource Economics. I am particularly grateful to Elizabeth Sadoulet for her stimulating and insightful comments, particularly in the early stages of my research work.

I would also like to thank John Hoddinott, who is also a co-author in one of the chapters in this thesis, for the very helpful discussions throughout the work. I am also grateful to Matthijs Lof, my co-author on one of the chapters; it has been a pleasure working with you.

My parents and siblings deserve very special thanks for the love, care and support they showed me throughout my study and always. The confidence you have in me, your endless encouragement and the values you have instilled in me help me to continue to strive to be the best version of myself. For this and much more, I will forever remain grateful!

Last certainly not least, my deepest thanks goes to my husband, my coauthor and my best friend, Abbishaye (Kiflu). Thanks for always being there for me and for your love, companionship and all rounded support. You have been a constant source of encouragement and inspiration. This long journey abroad would have been even longer and less enjoyable, had it not been for your sense of humor, love and encouragement. Thank you!

Tseday Jemaneh Mekasha
November 2015
Copenhagen, Denmark

Contents

Summary

This thesis comprises five self-contained chapters which can be read independently. The first two chapters are empirical studies on social protection and household demographic dynamics. The third chapter investigates exchange rate induced inflation in Ethiopia and combines both theory and empirics. The last two chapters present empirical evidence on aid effectiveness using two different empirical methods.

Chapter 1, "Social Protection Programs and Household Demographic Dynamics: Capturing Unintended Outcomes" (joint with John Hoddinott) studies how participation in social protection programs affects household size and composition. This is partly motivated by the longstanding concern in the literature that participation in cash transfer programs may give (unintended) incentives to beneficiary households to increase their size, where the main worry relates to the potential impact of such programs in increasing fertility/number of children. However, little is known about whether participation in cash transfer programs with a focus on poverty reduction leads to changes in household demographic dynamics, particularly through an increase in fertility. In this chapter we investigate this issue using survey data from the Ethiopian productive Safety Net Program, where households receive cash transfers conditional on participation in labor intensive public works. The chapter also looks into potential mechanisms through which participation in the program can lead to a change in household size and composition.

Chapter 2, "Linking Cash Transfers with Labor Supply: Assessing Impact on Household Structure" builds upon Chapter 1 and further examines how the level of financial gain from program participation and the associated labor supply requirement affect household structure. On the one hand, by relaxing beneficiary households' financial constraints, cash transfers from the public works can boost households' capacity to support more members. On the other hand, the labor supply requirement attached to the program is likely to introduce a time/labor constraint on beneficiary/participant households potentially inducing them to change their household structure. Arguably, both the income effect and labor supply requirement aspects of the program can induce changes in household structure, either through attraction of new members or retention of existing ones. The aim of this chapter is thus to disentangle the income effect of the program from the effect that comes through the labor supply conditionality and investigate how these channels work in terms of changing household structure.

Chapter 3, "A Closer Look at Exchange Rate Induced Inflation in Ethiopia" (Joint with Kiflu G. Molla) empirically examines how macroeconomic shocks affect consumer

prices and its disaggregated components in Ethiopia with a particular focus on shocks to the nominal exchange rate. Maintaining export competitiveness in the face of inflationary environments is a major policy challenge for many sub Saharan African countries including Ethiopia. In such instances, even if devaluation measures seem unavoidable in the face of alarming inflation rates and deteriorating trade balance, it is not obvious whether such measures add more to the inflationary pressure or lead to improved competitiveness. This chapter investigates this issue for the Ethiopian economy using a Structural VAR approach where identification of the structural shocks is achieved using a combination of short and long run restrictions. The derivation of the long run identifying restrictions is guided by a simple open economy macro model, which is modified to take into account shocks to the trade balance and external financial flows that are important in the determination of domestic prices and exchange rates in Ethiopia.

Chapter 4, "Aid and Growth: What Meta Analysis Reveals" (Joint with Finn Tarp) looks into the aid effectiveness evidence using a meta-analysis tool. In particular, this paper aims to address two important questions: i) whether the accumulated evidence on the impact of aid on growth is different from zero when one combines results from existing empirical studies on aid and growth? ii) if the observed effect is genuine or an artefact of publication bias? Using a sample of 68 aid-growth studies, our analysis revisits previous work in the area that casts doubt on the role of aid in spurring economic growth. We have employed a meta-analysis tool that takes into account the inherent heterogeneity in effect estimates across studies. The underlying model choice does matter for the conclusions drawn about aid effectiveness using meta-analysis.

Chapter 5, "Aid and Growth: Another Time Series Perspective" (Joint with Matthijs Lof and Finn Tarp) also revisits the aid effectiveness evidence using a Panel VAR approach. This is partly motivated by previous work on the area, particularly, the paper by Nowak-Lehmann, Dreher, Herzer, Klasen, and Martnez-Zarzoso (2012) where we uncovered problems in data handling and model specification that leads to an aid ineffectiveness conclusion. We revisit this evidence and in the process we have shown how data handling and choice of an appropriate model matters for the inference one makes regarding the effect of aid on growth.

Resumé (på dansk)

Denne afhandling består af fem selvstændige indeholdt kapitler, der kan læses uafhængigt af hinanden. De to første kapitler er empiriske studier om social velfærd og husholdningers demografiske dynamik. Det tredje kapitel undersøger valutakurs induceret inflation i Etiopien og kombinerer både teori og empiri. De to sidste kapitler præsentere empiriske resultater om bistandseffektivitet ved hjælp af to forskellige empiriske metoder.

Kapitel 1, *"Sociale Velfærdsprogrammer og Husholdsningers Demografiske Dynamik: Identifikation and Utilsigtede resultater"* (fælles med John Hoddinott) undersøger hvordan deltagelse i programmer for social beskyttelse påvirker husstandens størrelse og sammensætning. Dette er delvist motiveret af den mangeårige bekymring i litteraturen, at deltagelse i transfer-programmer kan give (utilsigtede) incitamenter til husholdningerne til at øge deres størrelse, hvor den vigtigste bekymring vedrører de potentielle virkninger af sådanne programmer i stigende fertilitet/antal børn. Men lidt er kendt, om deltagelsen i transfer-programmer med fokus på fattigdomsbekæmpelse faktisk fører til ændringer i husholdningernes demografiske dynamik, navnlig gennem en stigning i fertiliteten. I dette kapitel undersøger vi dette problem ved hjælp af data fra det etiopiske Safety Net Program, hvor husholdningerne modtager kontante overførsler betinget af deltagelse i arbejdsintensive offentlige arbejder. Kapitlet studerr også de potentielle mekanismer, hvorigennem deltagelse i programmet kan føre til en ændring i husstandens størrelse og sammensætning.

Kapitel 2, *"Sammenkædning af Kontant Overførsler med Arbejdsudbuddet: Vurdering af påvirknignen på husholdningernes struktur"* bygger på kapitel 1, og undersøger, hvordan niveauet for økonomisk gevinst fra program deltagelse og tilhørende krav om arbejdsudbud påvirker husholdningernes struktur. På den ene side gælder at ved at slække på modtagerlandenes husholdningernes finansielle begrænsninger kan kontante overførsler fra offentlige arbejder øge husholdningernes evne til at understøtte flere medlemmer. På den anden side gælder at krav til arbejdsudbud knyttet til programmet kan forventes at resultere i en tids arbejdskraft begræænsning på modtageren/deltagende husstande som potentielt kan inducere dem til at ændre deres husstand struktur. I realiteten kan både indkomst effekt og udbud af arbejdskraft krav aspekterne af programmet resultere i ændringer i husholdningernes struktur, enten gennem tiltrækning af nye medlemmer eller fastholdelse af eksisterende. Formålet med dette kapitel er således at udrede indkomsts effekten af programmet fra den virkning, der kommer gennem arbejdsudbuddet og undersøge, hvordan disse kanaler påvirker i form af nye samlivsformer.

Kapitel 3, *"En Undersøgelse af Valutakursinduceret Inflation i Etiopien"* (med Kiflu G. Molla) er ne empirisk undersøgelse af hvordan makroøkonomiske chok påvirker forbrugerpriserne og dets disaggregerede komponenter i Etiopien med særlig fokus på stød til den nominelle valutakurs. Fastholdelse af eksport konkurrenceevnen i inflationære miljøer er en stor politisk udfordring for mange afrikanske lande, herunder Etiopien. I sådanne tilfælde gælder, at selv om devaluering synes uundgåelige i lyset af alarmerende inflationsrater og den forværrede

handelsbalance, er det ikke indlysende, om sådanne foranstaltninger fører til mere inflationspres eller en forbedret konkurrenceevne. Dette kapitel undersøger dette spørgsmål for den etiopiske økonomi ved hjælp af en Structural VAR tilgang, hvor identifikation af de strukturelle stød opnås ved hjælp af en kombination af kort og lang sigts restriktioner. I analysen af det lange løb identificeres begrænsninger der er styret af en simpel åben økonomi makro model, som er modificeret til at tage hensyn til chok i handelsbalancen og de eksterne finansielle strømme, der er vigtige i forbindelse med fastsættelsen af indenlandske priser og valutakurser i Etiopien.

Kapitel 4, *"Bistand og Vækst: Hvad Meta-analyse Viser "* (med Finn Tarp) studerer bistandens effektivitet af et meta-analyse værktøj. Dette papir har især til formål se på to vigtige spørgsmål: i) om den akkumulerede dokumentation for bisyanden påvirkning på væksten er forskellig fra nul, når man kombinerer resultaterne fra eksisterende empiriske undersøgelser om bistand og vækst? ii) Om den observerede effekt er reel eller en publicerings-skævhed? Ved hjælp af en stikprøve på 68 studier, gennemanalyserer vores studie tidligere arbejde på dette område, der har sete tvivl om den rolle bistand spilelr for væksten. Vi anvender en meta-analyse, der tager hensyn til den iboende heterogenitet i effekt estimaterne på tværs af studier. Det underliggende model valg er bestemmende for konklusionerne om bistandseffektivitet ved hjælp af meta-analyse.

Kapitel 5, *"Bistand og Vækst: En Andet Tidsserieperspektiv "* (med Matthijs Lof og Finn Tarp) genbesøger bistandseffektivitets literaturen ved hjælp af en panel VAR tilgang. Dette er delvist motiveret af tidligere arbejde på området, især et papir af Nowak-Lehmann, Dreher, Herzer, Klasen, og Martínez-Zarzoso (2012), hvor vi afdækker problemer i deres data håndtering og model specifikation, der fører til en konklusion om at bistanden er ineffektiv. Vi gennemgår dette og i processen påviser vi hvordan håndtering af data og valg af en passende model påvirker konklusionerne vedrørende bistandens effekt på væksten.

Chapter 1

Social Protection Programs and Household Demographic Dynamics: *Capturing Unintended Outcomes*

Social Protection Programs and Household Demographic Dynamics:
Capturing Unintended Outcomes

John Hoddinott *   Tseday J. Mekasha†

Abstract

In this paper we investigate the impact of participation in social protection programs on household demographic dynamics. This is done based on data from the four rounds of the Ethiopian Productive Safety Net Program (PSNP) where households receive cash transfers conditional on supplying labor for public works. Identification is achieved using differences in outcomes between participants and non-participants as well as variation in participation status caused by late comers to the program (switchers). The results show that program participation is positively associated with household size and this appears to be mainly driven by a relative increase in the number of adolescent female household members. A more disaggregated analysis, based on the migration status of household members, suggest that the observed increase in the number of female members is due to the program's impact in deterring out-migration of household members. On the other hand, we do not find enough evidence to support the claim that participation in cash transfer programs induce higher fertility. The finding in this paper rather points to a negative and statistically significant association between program participation and households' fertility decision.

Keywords: Conditional Cash Transfers, Public Works, Household Size, Migration
JEL Classification: J10, J12, J13, J22, J23, I38, O15

*Cornell University, Charles H. Dyson School of Applied Economics and Management; jfh246@cornell.edu;

†Copenhagen University, Department of Economics; tseday.mekasha@econ.ku.dk;

1 Introduction

Understanding how social protection programs with targeted income transfers influence household size and composition is an important, yet largely overlooked subject in the program design and impact evaluation literature. Through boosting households' capacity to support more members, participation in cash transfer programs can potentially induce changes in households' decision regarding fertility, marriage and migration; all of which have a role in shaping household size and composition. This will, in turn, have important implications for intra-household resource sharing and eventually for household welfare.[1] In general, since an increase in household size, ceteris paribus, is likely to erode the value of transfer payments it can possibly hamper the effectiveness of poverty focused cash transfer programs.

Thus, the main objective of this study is to empirically examine the household demographic impact of safety nets and poverty focused cash transfer programs in low income countries. This is done using data from the Ethiopian Productive Safety Net Program (PSNP), which is one of the largest social protection programs in Africa next to South Africa. This program involves government cash transfers to poor and chronically food insecure rural households contingent on participation in labor intensive public work activities meant to build community assets. While the cash transfer under the Ethiopian PSNP does not directly target household demographic outcomes like fertility, our objective here is to assess the unintended consequences in terms of changing household size and composition. In order to better understand the potential impact of the program on household demographics, we further investigate the major determinants of household size and composition, namely fertility, marriage and/or migration of members in and out of the household. Accordingly, this paper speaks to the literature on the incentive effects of income gains on fertility, migration and marriage decisions of households.[2]

One strand of the literature focuses on developed countries and assess the fertility response of income gains following government welfare programs.[3] In the context of developing countries, similar concerns on the potential impacts of cash transfer programs on fertility decisions have also been raised. The literature in this regard is, however, mainly limited to cash transfer programs in Latin American countries. In relation to fertility responses of participation in such programs, the literature presents mixed evidence.[4] In addition to the aforementioned concerns in the literature regarding

[1] The literature on the impact of household size on household welfare (as proxied by food consumption) particularly focuses on assessing the existence of economies of scale in consumption in relation to household public goods where household size is used as a scale variable. For earlier debates on this, see for eg. Lanjouw and Ravallion (1995) and Deaton and Paxson (1998) and for a more recent discussion on the issue see Fafchamps and Quisumbing (2007).

[2] See Kearney (2004) , Rosenzweig and Wolpin (1982), Stecklov et al. (2007) and Todd et al. (2012) among others.

[3] See for example, Brewer et al. (2012a), Milligan (2005), Baughman and Dickert-Conlin (2009a), Baughman and Dickert-Conlin (2009b), Brewer et al. (2012b) and Kearney (2004) among others.

[4] For instance, while Stecklov et al. (2007) find a positive impact on fertility following conditional cash transfer program in Honduras, they do not find similar evidence for Mexico and Nicaragua. On

the potential effects of cash transfer programs on fertility, there is also some evidence on how such programs induce changes in households' decision in relation to other determinants of household size. These mainly include households' decision regarding marriage (Bobonis (2011)) and migration (see Hosegood et al. (2009), Angelucci (2012),Stecklov et al. (2005),Hagen-Zanker and Himmelstine (2013) and papers cited therein).

Given the emerging prominence of social protection programs as a tool for poverty reduction (Nio-Zaraza et al. (2012)) and promote micro level growth (Barrientos (2012)) in low income countries, understanding how poverty focused cash transfer programs affect household demographics is important. Despite this, we reiterate that little is known in this regard and we believe that this study will have important contribution.

Our findings show that participation in public works program under PSNP is associated with an increase in household size. This is found to be mainly due a relative increase in the number of adolescent female household members. Digging dipper reveals that the observed increase in female adolescent members appear to be due to increased retention of current members, particularly through early marriage delaying effect of the program. On the other hand, we do not find evidence to suggest that program participation is associated with an increase in the number children or fertility.

The rest of the paper is structured as follows: section 2 presents the setting, including program background, data and sample information and section 3 outlines the identification strategy and the relevant validity checks. While section 4 discusses the analytical framework, section 5 presents the results and related discussion. Finally, concluding remarks are given in section 6.

2 The Setting

2.1 Program Background

Despite the encouraging aggregate economic growth performance in recent years, food security continues as a pressing development challenge in Ethiopia. For decades the country has relied on emergency food aid. While such assistance has been helpful in meeting transitory (acute) food needs and in saving lives, it has not been able to bring about a sustainable solution to chronic poverty and food insecurity problems. Large numbers of rural people continue to live in abject poverty.

After the 2002/03 major drought episode that left 13.2 million people in need of food assistance, the government of Ethiopia, through multi-donor financing, launched the Ethiopian Food Security Program (EFSP) in 2005. The objective of the program is defined to address short term food gaps through predictable transfers and hence help

the other hand, Todd et al. (2012) investigate the fertility impact of a conditional cash transfer program in Nicaragua and document that the program is associated with an increase in birth spacing (see also Rubalcava and Teruel (2006) and Winters et al. (2006)).

households to avoid having to rely on damaging coping mechanisms in the face of negative income shocks. On top of preventing household asset depletion, the program aims to ensure long term food security through building household and community assets that are believed to promote the livelihood of the rural poor. Andersson et al. (2011)

One of the main components of the EFSP is the Productive Safety Net Program (PSNP). Under this program, the government provides two types of transfers: a conditional cash transfer, contingent on participation in labor intensive public works, and a direct support (unconditional cash transfer) provided for labor constrained households only. The majority of PSNP eligible households receive cash transfers based on public works and only around 15 percent receive direct support Hoddinott et al. (2012). Moreover, complementary to the PSNP, EFSP includes extension packages and credit services, collectively referred to as Other Food Security Program (OFSP). These services are mainly targeted to PSNP beneficiaries with the aim of achieving the livelihood promotion objective of the program, facilitating graduation of beneficiaries from the program.

Given that the program targets poor and chronically food insecure households, assignment to the program (eligibility) is not random. The selection process uses both geographic and community level targeting.[5] The basic criteria for selecting beneficiary households into the PSNP include: being member of the community, past history of chronic food insecurity, status of household assets and income from agricultural and non-agricultural activities.[6] In terms of coverage, EFSP is implemented in four major regions of Ethiopia (Amhara, Oromia, SNNPR and Tigray) focusing on chronically food insecure districts and households within these regions. Under normal circumstances, program beneficiaries are expected to graduate in 3-5 years.

2.2 Data and Sample Size

The data used in this study relies on the four rounds of the Ethiopian Productive Safety Net surveys conducted in 2006, 2008, 2010 and 2012. These surveys cover the four PSNP beneficiary regions of Amhara, Ormiya, SNNPR and Tigray. The details about the sampling method is thoroughly discussed in Berhane et al. (2011a) and Hoddinott et al. (2012). The following discussion about the sampling strategy draws on these two studies.

During the first survey in 2006, two stage clustered sampling was used within each region where 68 *woredas* (districts) were randomly selected from a total of 153 chronically food insecure *woredas*. In the second stage, within each selected *woreda*, a random sample of *kebeles* (counties) was selected from a list of *kebeles* where PSNP

[5]This includes the actual identification of target households by the community food security task force and verifying the client list in a public meeting.

[6]Productive Safety Net Program Implementation Manual (2010).

Table 1: Unique Number of Households by Survey Round and Region

	Round 1	Round 2	Round 3	Round 4
Tigray	899	870	846	991
Amhara	900	864	849	985
Amhara HVFB	0	1,160	1,148	1,103
Oromiya	939	870	885	965
SNNP	950	934	917	1,048
Total	3,688	4,698	4,645	5,092

is operational.[7] From each of the selected *kebeles*, a sample of 15 PSNP beneficiary and 10 non-beneficiary (comparison) households were drawn giving a total sample of 25 households per *kebele*. The comparison households were selected from the PSNP beneficiary *kebeles* and are from the same area as PSNP beneficiaries. In this way, a total of 3,668 households were sampled in the 2006 round comprising 899 households from Tigray, 900 from Amhara, 939 from Oromia and 950 from SNNP. Out of these households, 3,095 households appeared in our data set in all of the four survey rounds in 2006, 2008, 2010 and 2012. Table 1 summarizes the total sample information by survey round and region.[8]

The sample size used in this paper is lower than the one indicated in Table 1. This is so for two main reasons. First, beneficiary households are included in our sample only if they are late comers to the program. Due to this, we are not using households that became beneficiaries from the inception of the program as they will not have a before period to apply the Differences-in-Differences (DiD) approach that we rely on here. Second, we also exclude the years 2007, 2009 and 2011 as we do not have household size and composition information for these years.

The questionnaire in the different rounds remains the same ensuring comparability across the different rounds of surveys. (See also Berhane et al. (2011a). For the purpose of this study, the rich set of demographic questions including information on household size, current and former household members are of particular importance on top of information on the status and length of participation and amount of public works payment. Moreover, we also use the information on different time varying household characteristics to account for factors that are likely to affect both participation and household demographic outcomes.

Before moving on to our model and identification strategy, the following sub sections are devoted to elaborating on how beneficiary (treatment) and non- beneficiary (control) households are defined.

[7]'Woreda' and'Kebele' are administrative units where the latter is the smallest administrative unit in Ethiopia.

[8]Amhara-HVFB refer to PSNP beneficiary woredas in Amhara region that receive High Value Food Basket (HVFB) through the support of USAID. Amhara-HVFB sample is included in the survey starting from 2008 onwards for the sake of comparing HVFB woredas with other beneficiary Woredas in Amhara that receive standard PSNP tarnsfer. (Berhane et al. (2011b))

2.3 Beneficiary (Treatment) vs Non-Beneficiary (Control) Households

Beneficiary (Treatment Households)

For our purpose, receiving payment for public works participation is considered as treatment. In particular, for each year, beneficiary ('treatment') households are defined based on whether or not they have received public works payments.

After having defined the pool of beneficiary (treatment) households, we categorized them into three groups based on the year in which they started to take part in the program. As we have already alluded to above, this classification is important in order to use the variation in the treatment status caused by late comers to the program. Constructing the treatment group (restructuring the data) as depicted below has the merit of creating a "before-after period", which is necessary for application of a Differences-in-Differences (DiD) estimation strategy. We refer to these groups as treatment batches and the respective groups are defined as follows:

- Group 1: received payment both in 2007 and 2008 or only in 2008 ;
- Group 2: received payment in 2007-2010 or 2008-2010 or 2009-2010 or only in 2010 and
- Group 3: received payment in 2007-2012 or 2008-2012 or 2009-2012 or 2010-2012 or 2011-2012 or only in 2012

Restructuring the data in the above manner gives us a total of 437 unique beneficiary households (1414 beneficiary observations).

Non-beneficiary (Control) Households

For each treatment batch defined above, a pool of control group household is constructed. This pool includes:

- Households that have never received payment, i.e., during the entire period;
- Households that did not participate in early years (and hence received no payment) and become beneficiaries only in later years and;
- Households that were initially entitled to participate but that did not actually receive any payments;

In all cases, direct support beneficiaries are excluded from the control pool. Having defined the pool of control households in the above manner, an "appropriate control group" is selected mainly based on the following considerations:

1. **Geography**: To get a comparable number of control group households from the same Kebele administration area as the treatment households.

Table 2: Unique Number of Households used in this paper: By Survey Round and Region

	Round 1	Round 2	Round 3	Round 4
Tigray	271	265	269	264
Amhara	213	198	202	201
Amhara HVFB	0	104	119	112
Oromiya	106	97	99	99
SNNP	126	122	124	120
Total	716	786	813	796

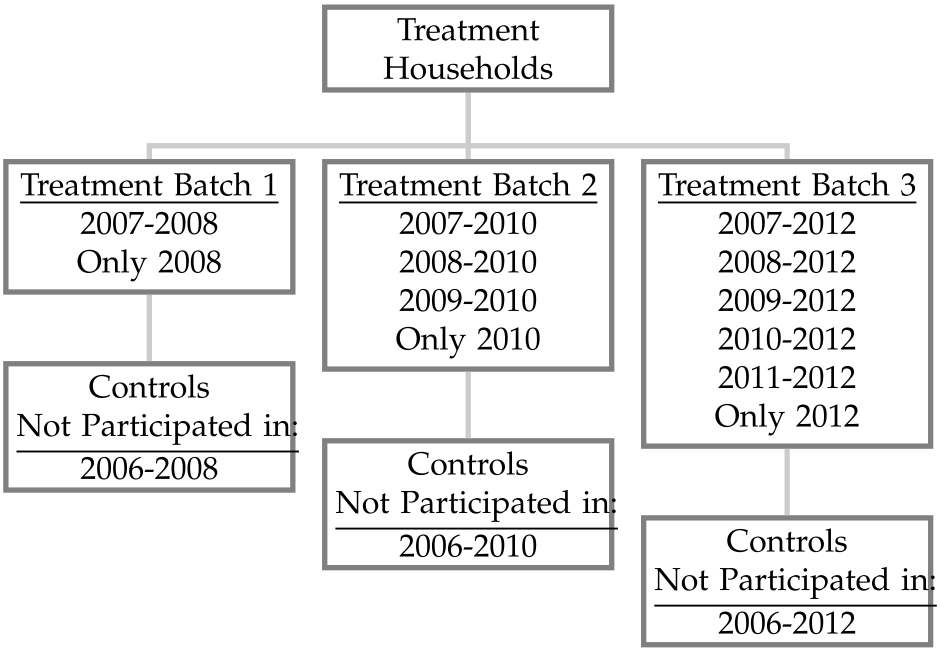
2. **Participation in later years and initial Entitlement**: In choosing controls based on geography, priority is given to households that were not beneficiaries in early years and started to take part in the program in later years as well as for those households that were entitled but did not receive payments. This is important in getting a more comparable control group households vis-a-vis the selected group of treatment households;

If enough numbers of "comparable controls" cannot be found based on the considerations implied in 1 and 2 above; additional controls are added from the same Kebele as the treatment households, using households that never participated in the program i.e., without conditioning neither on participation in later years nor on initial entitlement. When the above steps do not result in a sufficient number of control households, controls are taken from the same Woreda administration area. Following the above method and excluding the years 2007, 2009 and 2011, we have a total number of 466 unique comparison households (1,457 observations.)

In sum, combining beneficiary and non-beneficiary households, we get a total number of 2,871 observations (1,414 beneficiary and 1,457 non-beneficiary households) which are used in this paper.[9] We have summarized the treatment and control households information in Figure 1. Moreover, the sample size by survey round and region is summarized in Table 2.

[9] In the regressions where we have included time varying variables, the total number of observations can be smaller than this as there are many missing values on some of the covariates.

Figure 1: Beneficiary (Treatment) Households and their respective Comparison (Control) Groups



3 Identification Strategy and Validity Check

To identify the effect of interest, we exploit the variation in treatment between beneficiary and non-beneficiary households as well as changes in treatment status of beneficiary households caused by late comers to the program (switchers). The latter is possible due to continuous re-targeting efforts undertaken with the aim of rectifying inclusion and exclusion errors that occurred in the course of selecting participant households. This induces movements of households in and out of the program and is taken as an opportunity for identification.

In particular, some households were not part of the program for some time (for example, until 2008) and become beneficiaries in later years only (say 2010). For such households there is a clear before and after program period and this, together with non-beneficiary households, enables us to apply the Differences-in-Differences (DiD) approach in estimating the impact of the program. Those households who were beneficiaries from the start of the program (2006) are dropped from the sample as these households do not have a 'before program period'.

To practically implement this, noting the fact that the 'before' and 'after' periods are different for different households, treatment households are placed in different cohorts, c, depending on which period(s) they received treatment (i.e. became PSNP beneficiaries). For each of these cohorts of treatment group households, observed over the period p_c, appropriate control group households are assigned following the steps discussed in Section 2.3. The appropriate control group households for each cohort c will be those who have not received any payment during period p_c (See also Figure 1).

Accordingly, the regression equation to be estimated can be given as:

$$H_{it} = \alpha + \beta_1 D_i + \beta_2 T_{it} + \beta_3 D_i \times T_{it} + \beta_4 X_{it} + \gamma_t + \lambda_{tr} + \eta_r + \varepsilon_{it} \tag{3.0.1}$$

where H_{it}, the outcome of interest (size and composition of household i at time t), D_i is a dummy which indicates treatment status of household i, T_{it} is a period dummy which equals one if t is an after treatment period for household i and zero otherwise. It is worth noting that T_{it} varies across households as households switch to the treatment group at different points in time. Thus, the coefficient estimate of the interaction term (β_3) captures the impact of the program on household demographic dynamics. Besides, the model includes: region fixed effects denoted by η_r to control for time invariant factors that are specific to each region; year fixed effects (γ_t) to capture factors that vary over time and are common to all households (and regions); and time fixed effects interacted with region dummies (λ_{tr}) to control for time varying factors that may have differential impact across regions. Finally, X_{it} refers to vectors of time varying household characteristics that may affect both participation and household demographic dynamics. ε_{it} is the usual error term.

The model is augmented with duration of participation and a treatment Batch dummy where the latter captures potential differences across the different batches. The above model is also estimated including household fixed effects u_i to control for time invariant household specific factors.

$$H_{it} = \alpha + \beta_1 D_i + \beta_2 T_{it} + \beta_3 D_i \times T_{it} + \beta_4 X_{it} + \gamma_t + \lambda_{tr} + \eta_r + u_i + \mu_{it} \tag{3.0.2}$$

Testing the Validity of the Identification Strategy

The above identification strategy works under the key assumption that, absent treatment, the underlying trend in the outcome of interest is the same for treatment and control group households. It is therefore important to validate this parallel trend assumption based on the data at hand.

Since the current data set up provides two data points in the 'before' period, we can check whether this assumption holds. For instance, using households that become beneficiaries of the program after 2008 and their respective control households, $\Delta Y^{08-06}_{treat.} - \Delta Y^{08-06}_{cont.}$ should not be significantly different from zero for the parallel trend assumption to hold. Similar argument also applies for households that become beneficiaries in 2010 and 2012.

In principle, those who have never been part of the program are, on average, expected to be relatively better off since the program is designed to help the poor and chronically food insecure households. In reality, however, the difference between the two groups may not be substantial and when there is a difference, it can only be in levels. For instance, there has been considerable exclusion error in targeting which makes the poor and chronically food insecure households to be left out of the program. Besides, since our control group construction gives priority to participants in later years and to households that were entitled but not benefited, this is likely to enhance the comparability of treatment and control group households. Finally, since comparison households were selected from the same communities as the PSNP beneficiaries they may not be starkly different from the ones in the treatment group.

To check whether the parallel trend assumption holds or not, we use a simple Placebo type regression by applying differences-in-differences estimation. For instance, for those who joined the program after 2008 but before 2010, the period between 2006 and 2008 will be their pretreatment period. We can therefore implement a Placebo type test considering 2008 as if it is an after treatment period and 2006 as a pre-treatment period. Similarly, for those who become beneficiaries after 2010, the period 2006-2010 will be their pre-treatment period. For these households, 2006 is used as their before period and 2008 and 2010 are used as their fake treatment periods for the Placebo estimation.

Therefore, for the parallel trend assumption to hold, $\Delta y^{08-06}_{treat.} - \Delta y^{08-06}_{cont.}$, which is given by the estimate for β_3, should be equal to zero.[10] If there is no differential trend in the outcome variable between beneficiaries and non-beneficiaries, β_3 in Equation 3.0.1 should be statistically indistinguishable from zero as this is the case where we do not have *actual* beneficiary households. The result from estimating Equation 3.0.1 and Equation 3.0.2, based on pre-treatment period observations, is reported in Table 3. As can be seen from the results reported in Table 3, the coefficient of the interaction term is statistically insignificant in all cases, corroborating the parallel trend assumption.

Issues of Endogeneity

Since PSNP eligible households can become public works beneficiaries only if they have at least one able bodied member to supply labor for public works, one may raise a question regarding the direction of causation. This may seem concerning as household size/number of members may seem to determine participation, instead of the other way around.

We argue that this is not an issue for the question at hand as the likelihood of selection (participation) for public works does not depend on the number of household members. In particular the fact that selection to the program is a two-step process makes the issue less of a concern. That is, in the first stage household eligibility to PSNP is decided by past history of unmet food needs, wealth and asset assessment and community membership; among other things. Once a household is picked as eligible to PSNP, in the second stage, availability of labor within the household is assessed to decide whether the household should receive conditional transfer under public works or unconditional transfer under direct support. The household will be classified as public works beneficiary if there is at least one able bodied member in the household. To be precise, the question is *not how many able bodied members* the household has, instead whether or not the household has *at least one* able bodied member.

[10] This is done using 2006 as a 'before' period and 2008 as a fake 'after' period. In the empirical estimation we have also added data points where 2006 is used as a 'before' period while considering 2008 and 2010 as fake'after' periods.

Table 3: Pre-Program Household Size: Checking Common Trend (Placebo)

	OLS			Using HH FE		
	HH Size	Females	Males	HH Size	Females	Males
Treatment× Post	0.424	0.140	0.284	-0.247	0.141	-0.388
	(0.346)	(0.243)	(0.258)	(0.317)	(0.147)	(0.239)
Post Period	0.740	0.407	0.333	0.860***	0.481	0.379
	(0.700)	(0.878)	(0.472)	(0.314)	(0.308)	(0.289)
Treatment	-1.026***	-0.348	-0.679**			
	(0.362)	(0.293)	(0.283)			
Constant	4.690***	2.399***	2.291***	4.354***	2.142***	2.213***
	(0.691)	(0.559)	(0.506)	(0.598)	(0.425)	(0.468)
Number of Obs.	467	467	467	467	467	467
$adjR^2$	.17	.066	.14	.069	.04	.054
Prob > F	0.00	0.00	0.00	0.00	0.04	0.00
Household FE	No	No	No	Yes	Yes	Yes
Time FE	No	No	No	No	No	No
Time FE × Region	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	No	No	No

Standard errors clustered at Woreda Level in Parenthesis

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In view of the above, sample selection would have been a problem if beneficiary households are more likely to have one able bodied member compared to non-beneficiaries (control households). However, this is unlikely to be the case as the exclusion of controls from the program is not based on number of able bodied member in the household, it is rather based on their economic and food security status. In view of this, the aforesaid endogeneity issue will not be a problem.

It is however true that public work beneficiary households are more likely to have one able bodied member compared to direct support beneficiaries. Nonetheless, since we do not include direct support beneficiaries in our control pool, the systematic difference in the number of able bodied members between public work and direct support beneficiaries will not pose a threat to our identification.

Another potential threat to the validity of our identification strategy is the issue that comparison households are, on average, wealthier than participants. This is likely the case as the program mainly targets to include poorer households within the community. This, however, will not be a problem as it will only introduce a level difference and does not introduce a systematic/slope difference. In other words, since the stock of household size (whether a household is large or small) is fixed (at least slow moving), it can be taken care of by the group indicator in our difference-in-differences estimation and will not bias our effect estimate.

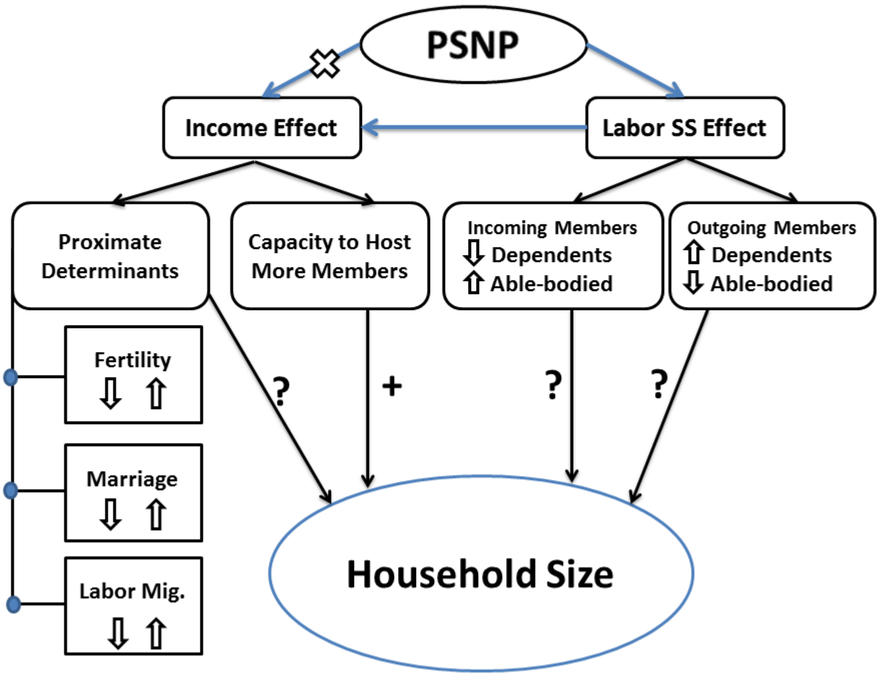
4 Analytical Framework

Changes in household size and composition depend on combinations of economic, social and cultural factors that can either reinforce or offset each other. Summarizing household formation in a single model may therefore conceal more than it can reveal about the complicated demographic dynamics within the household. In this section we present a simple analytical tool in the spirit of a theory of change to map the pathways through which participation in conditional cash transfer programs, such as PSNP, can affect the household size and composition. Since our objective is to lay a foundation for the empirics, instead of deriving a testable prediction from a formal theory model, the discussion below relies on the simple analytical tool depicted in Figure 2 below.

As indicated in section 2, the Ethiopian PSNP has two main components: a direct support component where households get an unconditional income transfers and a public works component which involves income transfers contingent on supplying labor for public works. Our interest in this paper is on the conditional income transfer component and the discussion below focuses only on this aspect of the program.[11] In order to understand how participation in PSNP public works affects household size, it is more instructive to think in terms of the kind of incentive/disincentive program participation creates on beneficiary households. The incentive effects can work either through the income effect of participation or the labor supply requirement which arguably puts a labor pressure within the household.

[11]This is the reason why the direct link from PSNP to income effect is crossed out in Figure 2.

Figure 2: Mapping the Pathways from Public Work Participation to Changes in Household Size



One way through which participation in PSNP affects household size is through increasing the availability of financial resources that may increase the capacity of beneficiary households to host more members. This can incentivize households to attract relatives and/or retain existing members in the household. As depicted in Figure 2 the increased capacity to host more members may lead to increases in household size. Thus, the potential role of a household as a risk mitigating/sharing unit can be one reason why financial gains (economic incentives) from participation in conditional cash transfer programs may affect household size (see Rosenzweig (1988)). As emphasized in Fafchamps (1999): "Given the importance of households in the explicit sharing of risk, it is not surprising to discover that the size and composition of households partly reflects the risk environment surrounding them." (p.27) See also Fafchamps and Quisumbing (2007). [12]

Another (income) incentive effect of cash transfer programs commonly mentioned in the literature is the increased demand for children/fertility following such interventions. The underlying argument is that the income gains from these programs can potentially change the preference of parents regarding quality vs quantity of children (see Stecklov et al. (2007), Becker and Lewis (1973) and Schultz (1997)). Accordingly, the public work cash transfer, by easing households' financial constraints of raising a child, may increase the demand for children. On the other hand, beneficiary households may prefer to use the additional income to improve the human capital of existing children

[12]However, once household size reaches a certain level, the insurance role of a household is likely to cease suggesting there is a finite upper limit to household size (Binswanger and McIntire (1987) and Fafchamps and Quisumbing (2007)).

(in terms of nutrition, health and education) instead of having an additional child. In this case we may not see an increase in fertility following income gains. Apart from changing child bearing decisions within existing marriages, postponing the timing of entry into marriage (delaying the age at first marriage) is another mechanism through which conditional cash transfer programs may change fertility patterns (Stecklov et al. (2007)). Thus, as shown in Figure 2 the impact of an income effect of participation on fertility/the demand for children can go in either direction and remains an empirical matter.[13]

Yet another mechanism through which income gains from participation in public works can influence household size and composition is by changing marriage patterns; that is, by either delaying or facilitating marriage of beneficiary household members. For instance, the availability of financial resources following public works participation can create an incentive to finance wedding ceremonies and hence inducing out-migration of members through marriage. On the other hand, if beneficiary households choose to use part of the cash transfer to invest in the human capital of their young members, they may postpone marriage decisions. Such investment in the human capital of young adults is likely to increase the gains from marriage as predicted by economic theory (see Becker (1973) and Becker (1974)). Taken together, the income effect of public works participation can have either a positive or a negative impact on the timing of marriage of young adults depending on which of the above effects dominate.

Finally, income gains from participation in cash transfer programs can also affect household size and composition by changing the pattern of labor migration (Hagen-Zanker and Himmelstine (2013), Stecklov et al. (2005) and Angelucci (2012)). In the context of PSNP, one can argue that beneficiary households can use part of the transfer to cover migration costs of members who wish to move out in search for better opportunities. This can partly be with the aim of diversifying income sources at the household level through remittances Stark and Bloom (1985). At the same time, because of the program design, the income effect of participation in the program may lower the tendency of labor out-migration from the household as beneficiaries may lose part of the benefit if some of the members out migrate. Similarly labor out migration can also decrease if the income transfer is big enough for beneficiaries to support their members without letting them to leave their place of origin. Thus, the income effect of participation on labor migration is ambiguous. The amount of the transfer as well as the conditions attached to the program are likely to influence how labor migration responds.

So far the discussion revolves on the income effect of participation in public works cash transfer program under PSNP. However, as pointed out earlier the labor supply requirement of participation in PSNP will also have its own role in determining household size and composition. This is depicted in Figure 2 as a labor supply effect.

For instance, beneficiary households may prefer to defer the marriage of their young adults as the labor supply requirement of the program is likely to create an upward

[13]Since the program does not have any component that is directly related with provision of health and family planning services, we do not expect it to have a direct effect on the supply of children.

pressure on the demand for working age labor. If this is the case, participation in public works program may induce beneficiary households to retain their working age adults leading to low out migration due to marriage. The same can be argued for labor out-migration.[14] In general, it can also be argued that participation in the program may lead to a decrease in out-migration of working age members from the household.

Furthermore, the effect of the labor supply requirement following program participation may go beyond labor retention and arguably induce in-migration of labor to the household, particularly of able bodied working age adults. This effect is likely to be pronounced for labor constrained households as the tradeoff between public works labor participation and working on own agriculture/other livelihood activities is relatively high for such households. Thus, households with relatively limited labor supply are likely to welcome an additional member to the family in order to get a helping hand either for activities related to own agriculture or household chores.

The discussion above mostly focuses on how the labor supply effect of participation can influence the movement of able bodied adults in and out of the household. The predictions are likely to be different if we consider how the labor supply requirement affects the size and composition of dependents. In particular, to the extent that the time constraint induced by participation in public works is binding, the labor supply effect of participation can induce out-migration of dependents, like small children. For instance, the labor supply effect of program participation may force beneficiary households to give their small children to relatives. Moreover, since all able bodied men and women should supply labor for public work activities, the program can create a disincentive to have more children as participation in public work activities is likely to compete with the time females use for raising children.[15]

In sum, the above mechanisms can lead to either an increase or a decrease in household size depending on the type of incentive it gives to public works beneficiary households or their members. In some instances the different incentives may also end up having offsetting effects. Therefore, examining the net effect of program participation on household size and composition remain an empirical issue.

5 Results and Discussion

In what follows we present the empirical evidence on the impact of participation in the Ethiopian PSNP on household size and composition. To this end we start by estimating Equation 3.0.1 and Equation 3.0.2 (with household fixed effect) for household size and its disaggregates by age and gender. This is then followed by estimation of the same model on probability of observing, alternatively, a household with at least one in or

[14]The labor supply requirement also implies that direct beneficiaries of the program (members that take part in public work activities) may have to stay in the household in order to get the benefit. This can also reinforce the member retention impact of the labor supply effect of the program.

[15]Even if pregnant female participants under PSNP are expected to shift from public works to direct support after four months of their pregnancy and for ten months after delivery, this incentive may not be high enough to induce them to have children as raising a child requires much more than that.

out migrant member. In all the regressions, the unit of analysis is the household. The DiD estimator of the effect of program participation is given by the coefficient of the interaction term between the group indicator variable (Treatment) and after treatment period indicator variable (Post).

As discussed in section 3, all specifications include region and time fixed effects and the interactions between the two.[16] Moreover, head characteristics and a host of time varying covariates including measures of household wealth, asset, shocks to crops and livestock, distress asset sales and poverty perceptions are included in all specifications. In addition, duration of participation in the program as well as a dummy variable for each treatment batch is included in all estimations. As an extension to this specification, we have also reported results from a specification that includes household fixed effects; in which case all time invariant household specific characteristics are absorbed by the household fixed effect. Moreover, we have used Fixed Effects Poisson (Quasi-ML) estimation technique in cases where our dependent variable is a non-negative count data. Standard errors clustered at Woreda level are used in all cases.[17]

5.1 Public Work Participation and Household Size

In order to estimate the impact of public works participation on household demographic dynamics, we start by examining how participation affects aggregate household size controlling for a host of time varying characteristics. The results are presented in Columns 1-3 of Table 4, which are respectively estimated with OLS, household fixed effects and fixed effects Poisson estimation techniques. In all cases, we have only presented the variables of interest and coefficients of other covariates are excluded for ease of presentation.

Focusing on Columns 1 and 2 of Table 4, PSNP participation is found to be associated with an increase in household size. In both cases, the increase in household size associated with participation appears to be around 0.3 household members. That is, participation in the program is associated with about one additional member in every three households. To put it differently, given the average household size of 5, [18] , this result amounts to a roughly 6 percent increase in household size following participation in the program. This is also consistent with the estimate from the fixed effects Poisson estimation where the coefficient estimates have percentage interpretation. According to the result reported in Column 3 of Table 4, program participation is found to be associated with 5.1 percent increase in household size.

[16] In regressions where we have used within estimation, the region dummies will be absorbed by the household fixed effect.

[17] The Poisson estimation is done using the xtpqml command in stata which allows for clustering.

[18] Information on average household size in the data is provided in Table B.

Table 4: Dependent Variable: Household Size, Female and Male Members

	Household Size		
	OLS	With Household FE	Poisson FE
Treatment × Post	0.295**	0.287**	0.053***
	(0.138)	(0.109)	(0.020)
Treatment	0.143		
	(0.489)		
Post Period	-0.336*	-0.329**	-0.055***
	(0.198)	(0.126)	(0.020)
Constant	2.582***	3.536***	
	(0.398)	(0.406)	
Number of Obs.	2074	2074	2074
$adjR^2$	.17	.076	
Prob > F	0.00	0.00	.
Household FE	No	Yes	Yes
Time FE	Yes	Yes	Yes
Time FE × Region	Yes	Yes	Yes
Region FE	Yes	No	No
Mean of the Dep. Var	5.57	5.57	5.57

Standard errors clustered at Woreda Level in Parenthesis

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

5.2 PSNP Participation and the Demand for Children

As we pointed out in the introduction there is a growing concern in the literature that financial incentive following participation in cash transfer programs might induce beneficiary households to increase their size by increasing the number of children/fertility. In the case of PSNP, however, it is not obvious a priori whether participation in the program will lead to an increase or a decrease in the demand from children. As discussed in section 4, on the one hand, the income effect of public works payment can increase the demand for children by easing the liquidity constraint parents might be facing. This may not, however, be the case if parents value quality over quantity of children. One the other hand, it can also be argued that the obligation to supply labor for public works can compete with the parents' time for child care. Thus, in so far as the time constraint due to public works participation is binding, the demand for children may not increase even if the financial constraint is relaxed through the cash transfer. In the face of public work induced labor pressure, households may actually send out their small children to relatives.

As can be seen from the summary statistics in the appendix (see Table B and Table 18), number of members in 0-6 age category and incoming members due to birth are larger

Table 5: Dependent Variable: Household Members in 0-6 Age Category

	Number of Household Members: 0-6 Age Group		
	OLS	With Household FE	Poisson FE
Treatment × Post	0.036	-0.008	-0.019
	(0.094)	(0.088)	(0.057)
Treatment	-0.123		
	(0.178)		
Post Period	-0.236**	-0.252**	-0.159**
	(0.102)	(0.101)	(0.068)
Number of Obs.	1775	1775	1775
$adjR^2$	.11	.044	
Prob > F	0.00	0.00	.
Household FE	No	Yes	Yes
Time FE	Yes	Yes	Yes
Time FE × Region	Yes	Yes	Yes
Region FE	Yes	No	No
Mean of the Dep. Var	1.44	1.44	1.44

Standard errors clustered at Woreda Level in Parenthesis

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

compared to other age groups and reasons. We thus first explore if the program is associated with an increase in the number of members in the 0-6 age group. Results are presented in Table 5 below. As can be seen from the results there is no evidence to suggest that program participation is associated with an increase in the number of members in the 0-6 age group.

Another way of testing whether participation in the program has impact on households fertility decision is to look at its impact on probability of observing a household with an incoming member in the 0-6 age category. We have presented results from this estimation in Table 6. In this regression, the outcome variable is a binary indicator taking a value one if the household has at least one incoming member in the 0-6 age category (regardless of the reason) and 0 otherwise. Since we have a binary dependent variable, we have presented results using Probit and linear probability models with and without household fixed effects estimations.

As can be seen from Table 6, program participation appears to be negatively associated with probability of observing a household with an incoming member in the 0-6 age category. The effect of interest is statistically significant at 5 percent level of significance in the Probit and LPM models. In the fixed effects model, although both the sign and magnitude are similar with the other estimations, the coefficient is not precisely estimated.

Table 6: Dependent Variable: Probability of Observing a Household with an Incoming Member in 0-6 Age Category

	Probability (0-6 Age Group)		
	Probit	(OLS)LPM	With Household FE
Treatment × Post (d)	-0.078**	-0.083**	-0.067
	(0.039)	(0.039)	(0.040)
Treatment (d)	-0.047	-0.037	
	(0.069)	(0.066)	
Post Period (d)	0.026	0.031	-0.024
	(0.047)	(0.043)	(0.054)
Number of Obs.	2185	2185	2185
$adjR^2$		.11	.074
Prob > F		0.00	0.00
Household FE	No	No	Yes
Time FE	Yes	Yes	Yes
Time FE × Region	Yes	Yes	Yes
Region FE	Yes	Yes	No
Mean of the Dep. Var	0.33	0.33	0.33

Standard errors clustered at Woreda Level in Parenthesis

(d) for discrete change of dummy variable from 0 to 1

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Even if the results presented in Table 5 and Table 6 are informative, the outcome variables do not distinguish between a 0-6 member that joins the household due to birth reason from a member in the same age group that becomes a household member due to other reasons. To get further insight on this we have estimated the impact of program participation on the probability of observing a household with an incoming member due to birth reason. The results from this estimation, presented in Table 7, give a more direct evidence in relation to fertility. This estimation is done conditional on a household having a female member in child bearing age. The results in Table 7 appear to concur with what is observed in the previous two tables. In particular, participation in PSNP public works program is found to be associated with a lower probability of observing a household with an incoming member by birth. This can be observed from the negative and statistically significant coefficient of the interaction term in Table 7.

The negative and significant coefficient observed in Table 6 and Table 7 may not necessarily imply that beneficiary households are less likely to have an incoming member in the 0-6 age group following program participation. The result might rather be an indication of the fact that non-beneficiary households are relatively more likely to have an incoming member in 0-6 age group. This is plausible given that non-beneficiaries are in a relatively better position, in terms of having time for child care, as they do not participate in labor intensive public work activities. This might make them to be more willing to welcome a 0-6 member either from other households or through birth.

On the other hand, another potential explanation for beneficiary households having less number of 0-6 members following program participation can be due to the labor supply effect of the program. Participation in the labor intensive public work activities is likely to compete with the time beneficiary households can have for child care. As a result, public work participants might be forced to give their dependents in 0-6 age group to be looked after by relatives. To substantiate this claim empirically, we have checked whether program participation is associated with a probability of observing a household with an outgoing member in 0-6 age category. Results from this exercise are presented in Table 8 and there is some suggestive evidence that participation in public works activities is positively associated with probability of observing a household with an outgoing member in 0-6 age category (regardless of the reasons).[19]

[19]The major reasons for observing a household with an outgoing child include: child death, to be with parents, orphaned/parents unable to help and to live with relatives

Table 7: Probability of Observing a Household with an Incoming Member:By Birth

	Probit	OLS	Fixed Effect
Treatment × Post (d)	-0.093**	-0.098**	-0.076*
	(0.041)	(0.041)	(0.045)
Treatment (d)	-0.087	-0.066	
	(0.067)	(0.064)	
Post Period (d)	0.040	0.046	-0.021
	(0.052)	(0.047)	(0.058)
Number of Obs.	2044	2044	2044
$adjR^2$		.11	.058
Prob > F		0.00	0.00
Household FE	No	No	Yes
Time FE	Yes	Yes	Yes
Time FE × Region	Yes	Yes	Yes
Region FE	Yes	Yes	No
Mean of the Dep. Var	0.33	0.33	0.33

Standard errors clustered at Woreda Level in Parenthesis
(d) for discrete change of dummy variable from 0 to 1
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Taken together we do not find enough evidence to suggest that participation in public works program is associated with a higher demand children. Unlike the concern of increased fertility following cash transfer programs, the results in this paper show that program participation does not incentivize beneficiary households to have new births. Apart from the labor supply effect of the PSNP, one potential explanation is a change in the preference of beneficiary households towards increasing the quality instead of quantity of children (see Becker and Lewis (1973) and Schultz (1997) for arguments along these lines). However, understanding the reason behind the negative impact of program participation in child bearing (the demand for children) following participation in the Ethiopian PSNP requires further investigation.

Table 8: Probability of Observing a Household with an Outgoing Member in 0-6 age category

	Probit	OLS	Fixed Effect
Treatment × Post (d)	0.043*	0.043**	0.035
	(0.023)	(0.021)	(0.023)
Treatment (d)	-0.014	-0.014	
	(0.031)	(0.043)	
Post Period (d)	-0.007	-0.008	0.025
	(0.019)	(0.018)	(0.020)
Number of Obs.	1665	1665	1665
$adjR^2$		.011	.024
Prob > F		0.00	0.00
Household FE	No	No	Yes
Time FE	Yes	Yes	Yes
Time FE × Region	Yes	Yes	Yes
Region FE	Yes	Yes	No
Mean of the Dep. Var	0.04	0.04	0.04

Standard errors clustered at Woreda Level in Parenthesis
(d) for discrete change of dummy variable from 0 to 1
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

5.3 Change in Household Size: Assessing Differential Impact by Gender and Age

The evidence above makes it clear that the change in aggregate household size following participation presented in subsection 5.1, Table 4 cannot be attributed to an increase in the number of children, particularly in 0-6 age group. In order to understand what is actually deriving the observed change, we pursue a disaggregated analysis. This is at the same time useful to establish whether the observed effect varies by the gender of household members or age. In particular, we start by estimating models where our outcome variables are the number of male and female household members, regardless of age. This will then be followed by a similar analysis on each gender group but across different age groups.

In Table 9 we present results on the impact of program participation on the number of males and females. It can be seen that while the program is found to be associated with an increase in the number of female household members (as can be noted from the fixed effects and Poisson fixed effects estimations), the effect on their male counterparts is not precisely estimated in all cases.

This result appears to suggest that the observed change in aggregate household size is entirely driven by an increase in the number of female household members following program participation. However, the observed lack of evidence of a statistically

Table 9: Dependent Variable: Number of Female and Male Members

	OLS		With HH FE		Poisson	
	Female	Male	Female	Male	Female	Male
Treatment × Post	0.154	0.149	0.176**	0.095	0.072***	0.032
	(0.096)	(0.105)	(0.075)	(0.090)	(0.028)	(0.033)
Treatment	-0.177	0.189				
	(0.302)	(0.306)				
Post Period	-0.125	-0.181	-0.155*	-0.166**	-0.053*	-0.050**
	(0.152)	(0.113)	(0.078)	(0.079)	(0.028)	(0.024)
Constant	0.713	-0.682	1.670***	1.888***		
	(0.486)	(0.431)	(0.246)	(0.233)		
Number of Obs.	2179	2151	2179	2151	2179	2151
$adjR^2$	.068	.14	.04	.04		
Prob > F	0.00	0.00	0.00	0.00	.	.
Household FE	No	No	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE × Region	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	No	No	No	No
Mean of the Dep. Var	2.64	2.87	2.64	2.87	2.64	2.87

Standard errors clustered at Woreda Level in Parenthesis

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

significant change in the number of male members may be due to offsetting effects from different age cohorts, (i.e.,) following program participation, the number of males can increase for some age groups while it can decrease for others. As we have alluded to in section 4 and shown in the simple analytical framework in Figure 2, the public works labor supply requirement under PSNP may compete with the demand for labor for agricultural activities. This may induce households to add working age male members, while they may tend to send out their dependent members to relatives as they may be labor constrained due to participation in PSNP. On the other hand, the income effect from the cash transfer may facilitate migration of male members in some age categories. The same can be said about female household members. It is therefore important to further disaggregate the analysis by different age groups in order to get more insight from the results.

Table 10: Program Participation and Number of Female Members: By Age Groups

	OLS			With HH FE			Poisson		
	7-11	12-18	19-29	7-11	12-18	19-29	7-11	12-18	19-29
Treatment × Post	0.046	0.315***	0.029	0.053	0.270***	0.024	0.097	0.396***	0.040
	(0.099)	(0.102)	(0.066)	(0.101)	(0.095)	(0.061)	(0.140)	(0.115)	(0.091)
Treatment	0.096	-0.221	-0.051						
	(0.124)	(0.181)	(0.113)						
Post Period	-0.085	-0.283***	0.069	-0.119*	-0.162**	0.165*	-0.183**	-0.214**	0.259*
	(0.085)	(0.101)	(0.093)	(0.068)	(0.079)	(0.093)	(0.092)	(0.094)	(0.149)
Constant	0.603***	0.445	0.542	-0.001	0.066	0.648**			
	(0.221)	(0.293)	(0.327)	(0.300)	(0.250)	(0.264)			
Number of Obs.	1233	1269	1289	1233	1269	1289	1233	1269	1289
$adjR^2$	.057	.048	.036	.034	.038	.036			
Prob > F	0.00	0.00	0.00	0.00	0.00	0.00	.	.	.
Household FE	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE × Region	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	No	No	No	No	No	No
Mean of the Dep. Var	0.72	0.81	0.68	0.72	0.81	0.68	0.72	0.81	0.68

Standard errors clustered at Woreda Level in Parenthesis
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The results from the age-gender disaggregation are presented in Table 10 and Table 11 for females and males, respectively. The estimations are done for certain age groups only. In particular, since the effect of the program on the number of members in the 0-6 and > 61 age categories is unlikely to vary by gender, the estimation for these two groups is not disaggregated by gender.[20]

[20]Moreover, similar regressions are also estimated for age groups 30-41 and 42-60 but the coefficients of interest are not precisely estimated for both female and male household members. The results are not reported here for economy of space.

Moving to the age-gender disaggregation result, as can be seen from Table 10, we find statistically significant evidence for a positive association between program participation and the number of female members in the 12-18 age group as. In particular, the coefficient of interest is statistically significant with a magnitude of 0.32 and 0.27 as can be seen in the OLS and fixed effects estimations, respectively. In view of the sample average of female members in 12-18 age group, which is around 0.81, the above estimates imply that program participation is, on average, associated with approximately a 33 and 39 percent increase in female household members in 12-18 age category. Similarly, the corresponding estimate from the fixed effects Poisson estimation technique also tells a similar story where participation is found to be associated with a 39 percent increase in the number of female household members in 12-18 age group.

Table 11: Program Participation and Number of Male Members: By Age Groups

	OLS			With HH FE			Poisson		
	7-11	12-18	19-29	7-11	12-18	19-29	7-11	12-18	19-29
Treatment × Post	0.118	0.024	0.023	0.135	0.014	0.020	0.171	0.026	0.026
	(0.085)	(0.086)	(0.091)	(0.087)	(0.084)	(0.098)	(0.117)	(0.088)	(0.138)
Treatment	-0.045	0.150	-0.093						
	(0.158)	(0.189)	(0.171)						
Post Period	-0.121	0.034	0.119	-0.091	0.027	0.130*	-0.110	0.013	0.211*
	(0.085)	(0.091)	(0.076)	(0.071)	(0.079)	(0.075)	(0.084)	(0.080)	(0.112)
Constant	0.575	-0.099	1.210***	0.346	0.843***	1.047***			
	(0.369)	(0.379)	(0.176)	(0.290)	(0.245)	(0.322)			
Number of Obs.	1268	1350	1237	1268	1350	1237	1268	1350	1237
$adjR^2$	.037	.036	.023	.02	.0095	.017			
Prob > F	0.00	0.00	0.00	0.00	0.00	0.00	.	.	.
Household FE	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE × Region	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	No	No	No	No	No	No
Mean of the Dep. Var	0.77	0.98	0.71	0.77	0.98	0.71	0.77	0.98	0.71

Standard errors clustered at Woreda Level in Parenthesis
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

For male household members, as can be seen from the results in Table 11, we do not find any evidence to suggest that program participation is associated with an increase in male household members. This is contrary to a priori expectation that the labor supply requirement of public works may lead to an increase in the number of male members for beneficiary households relative to non-beneficiary households.

The above results also make it evident that the observed effect of the program on aggregate household size is primarily driven by changes in the number of female members, particularly those in 12-18 age group.

In order to better understand the result observed for female members and to see if further disaggregation tells a different story for male members, we re-estimate the result for 12-18 age groups both for male and female members by splitting the 12-18 age interval into sub-groups. As can be seen from Table 12, the positive and statistically significant impact for female members remains unchanged even with further disaggregation of the 12-18 age group. On the other hand, the result for the number of males continues to be statistically indistinguishable from zero. (See results in Appendix Table C)

Table 12: Program Participation and Number of Female Members: Further Disaggregation of the 12-18 Age Group

	With HH FE				Poisson			
	12-14	12-16	14-18	16-18	12-14	12-16	14-18	16-18
Treatment × Post	0.236**	0.277***	0.245**	0.203**	0.449***	0.427***	0.462***	0.492**
	(0.091)	(0.087)	(0.094)	(0.092)	(0.173)	(0.122)	(0.150)	(0.198)
Post Period	-0.114	-0.123	-0.088	-0.119	-0.210	-0.199	-0.150	-0.238
	(0.101)	(0.089)	(0.095)	(0.129)	(0.194)	(0.136)	(0.130)	(0.224)
Constant	-0.480	0.101	0.459	0.747***				
	(0.391)	(0.324)	(0.376)	(0.233)				
Number of Obs.	960	1127	1044	730	960	1127	1044	730
$adjR^2$	.053	.056	.032	.021				
Prob > F	0.00	0.00	0.00	0.00	.	.	.	.
Household FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE × Region	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	No	No	No	No	No	No	No	No
Mean of the Dep. Var	0.55	0.70	0.63	0.50	0.55	0.70	0.63	0.50

Standard errors clustered at Woreda Level in Parenthesis
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In summary, participation in labor intensive public works cash transfer program under the Ethiopian PSNP is found to be associated with an increase in household size which appears to be mainly due to an increase in the number of female household members in the age 12-18 category. This implies that program participation not only increases aggregate household size but also induces changes in the composition of participant household members. In particular, the observed participation induced increase in the number of female household members implies that females are more needed in participant households relative to non-participants.

One explanation can be that the female members are needed for their labor either for public works, agriculture or household chores. However, if the reason relates to the need for labor either for public works and agriculture, one would expect to see a relative increase in the number of male members instead.[21] In view of this, if the

[21] To verify this claim, we use the member level information in our data and checked whether the probability of observing a household member working on public works and agriculture differs by gender. In particular we use a dummy variable indicating whether the member has actually participated in public works activities and worked on agriculture as alternative outcome variable. Our estimation takes into account unobserved household characteristics and time specific effects and the interaction between the

need for labor is the explanation for increasing the number of females, it is likely that participant households need the female labor for household chores. This can partly be attributed to the established gender norms in rural parts of Ethiopia where females are normally expected to take care of household tasks.

Another possible explanation for the observed positive impact of participation in the number of female members is that the income effect from participation gives beneficiary households the capacity to support their current female members in the 12-18 age category for longer period, instead of letting them to get married or migrate to the cities at early age. In view of the evidence in the literature that female marriage and migration to the cities can be a manifestation of the income risk the household is facing (For e.g. Rosenzweig (1988) and Rosenzweig and Stark (1989)), one can argue that the increase in income following public works participation, by mitigating potential income risks, can induce beneficiaries to retain their female members for long.

In either cases, since the observed increase in female members in 12-18 age group is a net increase which can either come from a reduction in the number of outgoing current female members or an increase in incoming female household members, we cannot make a firm conclusion at this stage about the reason behind the observed increase in female household members. In the next section we therefore assess how participation is associated with the probability of observing a household with an incoming or outgoing female member in the 12-18 age category.

5.4 Moving In and Out of the Household: Does Participation in PSNP Matter?

In this section we assess if participation in PSNP public works is associated with the probability of observing a household with a female outgoing or incoming member. As we have argued in Figure 2, a priori it is not obvious whether public works participation will make it more/less likely to observe a beneficiary household with an outgoing or incoming member. This by and large depends on the kind of incentives program participation creates for beneficiary households and their members. This is thus an empirical matter and below we empirically assess how public works participation affects movements of members in and out of the household focusing on female household members in the 12-18 age category. We emphasize on female members as our previous result suggests that the positive impact of program participation on household size is mainly driven by changes in the number of female household members.[22]

Accordingly, following our discussion in Figure 2 we expect female out-migration to be higher following participation if for instance, the income effect of participation

two. The evidence shows that participant female members are less likely to work both on public works and agriculture compared to males. Results not reported but can be made available upon request.

[22] All results are also done for male members but we do not report them here for economy of space and they can be made available upon request. Unless otherwise stated, in the regressions for male household members, the coefficient estimates on the variable of interest are not precisely estimated in all cases.

relaxes the liquidity constraint of beneficiary households. For instance, by giving beneficiaries the capacity to finance activities that they could not do before, such as wedding ceremonies and/or cover migration costs of potential female migrants, cash transfers can lead to changes in the number of female household members. On the other hand, income gains from participation in PSNP can improve beneficiary households' capacity to feed/support more members and hence can also deter out-migration of females in the 12-18 age group. This can be, for example, by delaying marriage or by making migration in search for work unnecessary. Moreover, the labor supply requirement of the program may also reduce out migration of females, if for instance females in 12-18 age group are needed for household chores while adults are participating in public works. Thus the net effect of participation on the probability of observing a household with a female out migrant depends on which of the above effects dominate.

Participation and Female Outgoing Members

We first assess impact of program participation on the probability of observing a household with at least one female outgoing member with a particular focus on the 12-18 age interval. We started off by estimating the impact of participation on female out-migration regardless of the reason for the out migration of the member.

As before, our unit of analysis is the household. Our outcome of interest is a binary indicator taking a value of 1 if the household has at least one female outgoing member in the specified age category and 0 otherwise. All the estimations are done conditional on the household having a female member in the specified age group. Since our dependent variable is binary, we use Probit models on top of the linear probability and fixed effects models. In the case of the Probit estimations we have reported the marginal effects.

As can be seen from results reported in Table 13, public works participation is found to have a statistically significant negative impact on the proportion of households with a female out migrant/outgoing member in the age groups 12-18, 14-18 and 16-18. On the other hand, for female household members in the 19-29 age group, we find a statistically significant positive association between participation and probability of observing a household with an out-migrant member in this age category. In the case of the fixed effects estimation, however, the coefficient estimate on the variable of interest is not precisely estimated. These results suggest that the member retention effect of public works participation is mainly apparent in adolescent female members. This result can be attributed to either the income effect or the labor supply effect of participation as discussed above.

We next turn to examining how public works participation affects out-migration of female members by reason of out-migration. Among the reasons for out migration, we focus on female out migration due to marriage as this is the dominant reason for out-migration of females indicated in the data.[23] Accordingly we estimate the potential

[23] The second common reason for females to leave the household is work related (that is, looking for/to be close to work). However, we do not estimate the impact of participation on female out-migration due to work as we do not have enough variation to do the estimation disaggregated by gender and age. In the appendix we have reported the result for out migration due to work reason without disaggregating it

Table 13: Probability of Observing a Household with a Female Out-Migrant: By Age Groups

	Probit				OLS(LPM)				With HH FE			
	12-18	14-18	16-18	19-29	12-18	14-18	16-18	19-29	12-18	14-18	16-18	19-29
Treatment × Post (d)	-0.046**	-0.061**	-0.082**	0.057*	-0.052*	-0.069*	-0.099*	0.052*	-0.057*	-0.082*	-0.109*	0.033
	(0.023)	(0.028)	(0.033)	(0.031)	(0.031)	(0.040)	(0.053)	(0.027)	(0.034)	(0.045)	(0.060)	(0.027)
Treatment (d)	0.028	0.034	0.005	0.010	0.032	0.027	0.003	0.011				
	(0.057)	(0.070)	(0.066)	(0.039)	(0.059)	(0.072)	(0.067)	(0.046)				
Post Period (d)	0.078**	0.098**	0.107**	-0.042	0.075**	0.093**	0.100**	-0.038	0.075**	0.103**	0.075	-0.030
	(0.034)	(0.040)	(0.043)	(0.029)	(0.031)	(0.036)	(0.039)	(0.032)	(0.035)	(0.042)	(0.053)	(0.041)
Number of Obs.	1286	1024	773	1346	1286	1024	773	1346	1286	1024	773	1346
Household FE	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE × Region	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No

Standard errors clustered at Woreda Level in Parenthesis
(d) for discrete change of dummy variable from 0 to 1
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

impact of public work participation on the probability of observing a household with a female out migrant due to marriage.

As can be seen from results depicted in Table 14, that the coefficient estimate of the variable of interest is negative in all cases, but it is precisely estimated only in the case of the Probit regression. This can therefore be taken as a suggestive evidence that program participation has a marriage delaying effect on female members in their early and late adolescent age.

Overall, the results on female out-migration suggest that program participation is, on average, negatively associated with out-migration of female household members. We also find some suggestive evidence regarding the potential impact of program participation in delaying marriage of female household members, particularly of those in their adolescent ages. This member retention tendency observed on the part of participants can be partly explained by beneficiary households' improved capacity to feed/educate their young female members and/or by the increase in labor demand following participation in labor intensive public works.

Participation and Female Incoming Members

Next we examine how public work participation is linked with the probability of observing a beneficiary household with a female incoming member in the12-18 age group. It can be the case that the income effect of participation induces beneficiaries to attract new members to the household (for eg. hosting relatives as a way of informal risk sharing mechanism). Moreover, to the extent that the labor supply requirement competes with the time beneficiary households can use for other livelihood activities, participation can make beneficiary households more likely to add a member for labor reason.

In Table 15, we report results where the probability of observing a household with at least one incoming female member is used as an outcome. Looking at the results, although the impact of the program on the probability of observing a household with female incoming member in the 12-18 age category is found to be positive, the coefficient estimates of the variables of interest are in all cases statistically indistinguishable from zero.

---

by age and gender. It can be seen that program participation is negatively associated with the probability of observing a household with an out-migrant member due to work in all age groups but 19-29 (see Table 21.

Table 14: Probability of Observing a Household with a Female Out-Migrant Due to Marriage: By Age Groups

	Probit				OLS(LPM)				With HH FE			
	12-18	14-18	16-18	19-29	12-18	14-18	16-18	19-29	12-18	14-18	16-18	19-29
Treatment × Post (d)	-0.035*	-0.047**	-0.033*	0.004	-0.042	-0.063	-0.046	0.007	-0.036	-0.064	-0.048	-0.032
	(0.020)	(0.022)	(0.020)	(0.017)	(0.036)	(0.044)	(0.041)	(0.024)	(0.039)	(0.045)	(0.047)	(0.026)
Treatment (d)	-0.003	-0.019	-0.017	-0.011	0.001	-0.022	-0.015	-0.005	0.000	0.000	0.000	0.000
	(0.045)	(0.052)	(0.051)	(0.021)	(0.063)	(0.075)	(0.082)	(0.035)				
Post Period (d)	0.040*	0.051*	0.034	-0.020	0.041	0.052	0.034	-0.027	0.020	0.023	0.037	-0.002
	(0.024)	(0.028)	(0.030)	(0.019)	(0.029)	(0.035)	(0.041)	(0.028)	(0.038)	(0.045)	(0.057)	(0.036)
Number of Obs.	1022	852	618	1061	1022	852	618	1061	1022	852	618	1061
Household FE	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE × Region	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No

Standard errors clustered at Woreda Level in Parenthesis
(d) for discrete change of dummy variable from 0 to 1
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Marriage is also one of the major reasons for observing a household with a female incoming member. We have therefore tested if this can explain the observed increase in household size following program participation. Specifically, we estimate the probability of observing a household with a marriage related female in-migrant member following program participation. The result from this estimation is reported in Table 22 and it can be seen that there is no strong evidence to suggest that participation is associated with the probability of observing a household with an incoming member due to marriage.

Table 15: Probability of Observing a Household with a Female In-Migrant: By Age Groups

	Probit				OLS(LPM)				With HH FE			
	12-18	14-18	16-18	19-29	12-18	14-18	16-18	19-29	12-18	14-18	16-18	19-29
Treatment × Post (d)	0.007	0.010	-0.000	0.004	0.012	0.017	0.001	-0.005	0.014	0.023	0.009	-0.009
	(0.013)	(0.015)	(0.009)	(0.016)	(0.017)	(0.018)	(0.011)	(0.020)	(0.015)	(0.017)	(0.011)	(0.022)
Treatment (d)	0.014	0.008	0.009	0.009	0.016	0.014	0.011	0.008				
	(0.020)	(0.021)	(0.019)	(0.026)	(0.022)	(0.023)	(0.023)	(0.036)				
Post Period (d)	0.042***	0.034***	0.027**	0.002	0.036***	0.030***	0.023**	0.003	0.036**	0.029*	0.017	-0.017
	(0.014)	(0.012)	(0.012)	(0.018)	(0.011)	(0.011)	(0.010)	(0.019)	(0.017)	(0.015)	(0.014)	(0.020)
Number of Obs.	2098	2021	1812	2050	2098	2021	1812	2050	2098	2021	1812	2050
Household FE	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE × Region	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No

Standard errors clustered at Woreda Level in Parenthesis
(d) for discrete change of dummy variable from 0 to 1
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

6 Conclusion

In this paper we started out asserting that uncovering how participation in safety nets and targeted cash transfer programs influence household demographic dynamics is an important policy relevant issue for low income countries. As a case in point, we focused on one of the largest cash transfer schemes in Africa, the Ethiopian Productive Safety Net Program (PSNP). The program was started in 2005 and the government transfers cash to poor households conditional on participation in labor intensive public work activities that aim to build community assets. Our analysis made use of four rounds of the Ethiopian Productive Safety Net survey data over the period 2006-2012. The survey covers both beneficiaries and non-beneficiaries. Identification is achieved using differences in outcomes between participants and non-participants as well as variation in participation status caused by late comers to the program (switchers). Estimation is done applying Differences-in-Differences approach.

Our results suggest that the Ethiopian PSNP is indeed positively associated with household size; and the more disaggregated analysis shows that the observed increase in household size mainly comes from changes related to adolescent female household members. Furthermore, digging into how the program influences in/out-migration of household members, it emerges that the observed increase in the number of female household members is mainly related to the negative association between program participation and number of out-migrant members. In other words, the program encourages retention of current female household members.

The estimated negative impact of program participation on female out migration is found to mainly work through marriage related out-migration. And this marriage delaying effect of the program appears to be particularly important for female adolescent members. There is also evidence that the program associated with a decline in out-migration due to work related reasons (labor out-migration). Overall, the member retention tendency observed on the part of participants can be explained by beneficiary households' improved financial capacity to support their members and/or by the increased labor demand following participation in the labor intensive public work activities. However, in this paper we cannot disentangle the income effect of participation from the labor supply effect as we are looking at the net effect of program participation.

In contrast, we do not find enough evidence to suggest that program participation has led to more in-migrant members or increased fertility (child bearing). We rather found a negative association between program participation and the proportion of households having a new member due to birth related reason. This can also be explained by the labor supply requirement of the program as participation in labor intensive public work activities possibly competes with the time beneficiaries can have for child care as discussed in section 4. However, this needs further investigation.

One important implication of our findings is the potential role of cash transfer programs in reducing early marriage. Though unintended, this is a desirable outcome as early marriage continues to be a challenge in rural parts of Ethiopia. Even if the marriage delaying effect can be considered as a desirable outcome in its own right, it will be more welfare enhancing if the adolescent females are retained for long for the purpose of human capital investment like education.

Another takeaway message of the paper relates to the potential role of the Ethiopian PSNP in reducing rural-urban migration. This result can be explained by the labor conditionality attached to the program and/or by the program induced increase in income that enables beneficiaries to cope with adverse shocks without leaving their place of origin. Given the high unemployment rates in the urban areas of the country and the resulting socio-economic pressures, this too can be considered as a desirable outcome of the program. However, the result at best reflects the short term effects on labor migration and a fall in current labor out-migration does not necessarily imply a decline in future migration patterns. The latter by and large depends on the success of the program in creating sustainable income and livelihood opportunities that can deter out-migration of recipients even after graduation from the program.

The above findings also have implication for participation in public works cash transfer programs and fertility. For one, the fact that the program is not found to be associated with an increase in child birth is a good news in light of the concern in the literature that financial incentives induced by cash transfer programs can increase fertility/the demand for children. Moreover, the marriage delaying result implied by the program can, unintentionally, play a role in reducing fertility by delaying the timing of entry into marriage.

In sum, policy makers need to assess both the intended and unintended incentives that public work cash transfer programs may generate as these matter for the overall effectiveness of such programs in achieving targeted outcomes.

Appendices

Appendix A: Covariate Balance

Figure 3: Kernel Density Graph of Predicted Propensity Scores

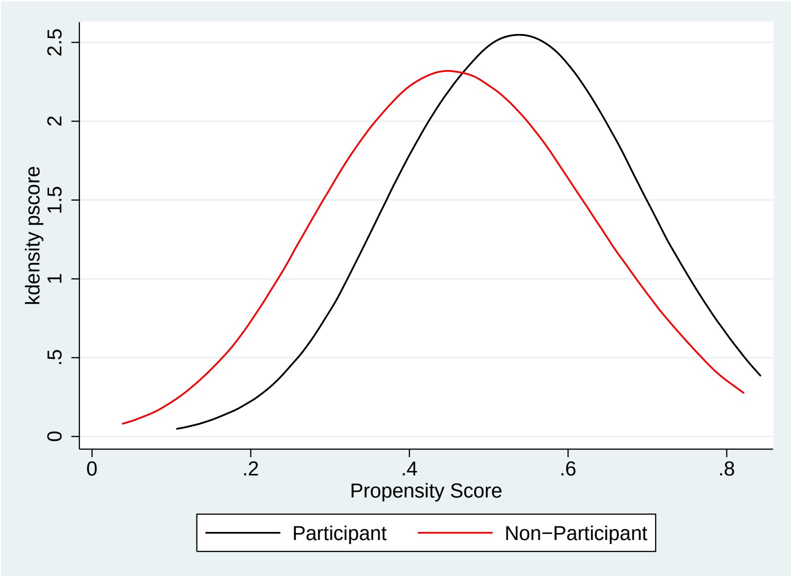


Table 16: Pre-Program Mean Comparison of Covariates across Beneficiary and Comparison Households

	Comparison	Obs	Beneficiary	Obs	Difference	p-val
Male Headed Household	0.87	447	0.81	449	0.06**	0.015
Head Age	46.14	447	44.31	449	1.83*	0.056
No Formal Education	0.77	447	0.76	449	0.00	0.967
Primary Education	0.20	447	0.20	449	-0.00	0.893
Secondary Education	0.03	447	0.03	449	-0.00	0.860
Value of Productive Assets	5.09	447	5.05	449	0.04	0.670
Corrugated Metal Roof	0.17	447	0.11	449	0.06**	0.012
Distress Asset Sell	0.51	447	0.53	449	-0.03	0.424
Shocks to Crop Production	0.66	447	0.65	449	0.01	0.816
Poblems in Crop Production	0.30	447	0.31	449	-0.01	0.644
Shock to Livestocks	0.66	447	0.67	449	-0.00	0.962
Poor Relative to Others	0.46	447	0.50	449	-0.04	0.181
Average Relative to Others	0.41	447	0.41	449	0.00	0.956

Appendix B: Descriptives: Summary Statistics

Table 17: Household Size and Composition

Variable	Mean	Std. Dev.	Min.	Max.	N
Household size	5.46	2.23	1	17	2185
No. of Males	2.83	1.6	0	11	2185
No. of Females	2.63	1.39	0	9	2185
No. of HH Members(0-6)	1.17	1.07	0	6	2185
No. of HH Members(7-11)	0.86	0.86	0	4	2185
No. of Males(12-18)	0.6	0.8	0	5	2185
No. of Females(12-18)	0.48	0.71	0	4	2185
No. of Males(19-29)	0.4	0.63	0	4	2185
No. of Females(19-29)	0.41	0.55	0	4	2185
No. of Males(30-41)	0.31	0.47	0	2	2185
No. of Females(30-41)	0.38	0.49	0	2	2185
No. of Males(42-60)	0.35	0.49	0	2	2185
No. of Females(42-60)	0.29	0.46	0	2	2185
No. of HH Members($>$ 61)	0.22	0.48	0	4	2185

Table 18: Summary Statistics on In-Migration of Members by Age, Gender and Reason

Variable	Mean	Std. Dev.	Min.	Max.	N
Dummy for In-Migrants	0.4	0.49	0	1	2179
Dummy for Male In-Migrants	0.21	0.41	0	1	2179
Dummy for Female In-Migrants	0.25	0.43	0	1	2179
Dummy for In-Migrants(0-6)	0.32	0.47	0	1	2179
Dummy for In-Migrants(7-11)	0.02	0.14	0	1	2179
Dummy for Female In-Migrants(12-18)	0.03	0.17	0	1	2179
Dummy for Male In-Migrants(12-18)	0.02	0.14	0	1	2179
Dummy for Female In-Migrants(19-29)	0.03	0.18	0	1	2179
Dummy for Male In-Migrants(19-29)	0.02	0.15	0	1	2179
Dummy for Female In-Migrants(30-41)	0.01	0.09	0	1	2179
Dummy for Male In-Migrants(30-41)	0.01	0.09	0	1	2179
Dummy for Female In-Migrants(42-60)	0	0.05	0	1	2179
Dummy for Male In-Migrants(42-60)	0	0.05	0	1	2179
Dummy for Birth	0.31	0.46	0	1	2149
Dummy for Marriage	0.05	0.21	0	1	2149
Dummy for To be Close to School	0.01	0.07	0	1	2149
Dummy for To be Close to Work	0.02	0.12	0	1	2149
Dummy for To live with Parent/Relat.	0.03	0.16	0	1	2149
Dummy for Other Reasons	0.03	0.18	0	1	2149
Dummy for Divorced	0.01	0.07	0	1	2149

Table 19: Summary Statistics on Out-Migration of Members by Age, Gender and Reason

Variable	Mean	Std. Dev.	Min.	Max.	N
Dummy for Out-Migrants	0.26	0.44	0	1	2069
Dummy for Male Out-Migrants	0.16	0.36	0	1	2003
Dummy for Female Out-Migrants(12-18)	0.14	0.35	0	1	2041
Dummy for Out Mig (0_6)	0.04	0.19	0	1	1336
Dummy for Out Mig (7_11)	0.02	0.12	0	1	1213
Dummy for Female out-Migrants(12-18)	0.06	0.24	0	1	804
Dummy for Male out-Migrants(12-18)	0.03	0.18	0	1	945
Dummy for Female out-Migrants(19-29)	0.04	0.21	0	1	832
Dummy for Male out-Migrants(19-29)	0.07	0.26	0	1	726
Dummy for Female Out-Migrants(30-41)	0	0.07	0	1	815
Dummy for Male Out-Migrants(30-41)	0.01	0.09	0	1	658
Dummy for Female Out-Migrants(42-60)	0	0.04	0	1	638
Dummy for Male Out-Migrants(42-60)	0	0.05	0	1	762
Dummy for Child Death	0.04	0.18	0	1	2055
Dummy for Marriage	0.06	0.24	0	1	2055
Dummy for To be Close to School	0.02	0.14	0	1	2055
Dummy for Work Related	0.06	0.24	0	1	2055
Dummy for Adult Death	0.04	0.19	0	1	2055
Dummy for Divorced	0.01	0.12	0	1	2055
Dummy for Other Reasons	0.03	0.16	0	1	2055

Appendix C: Additional Regression Results

Table 20: Program Participation and Number of Male Members: Further Disaggregation of the 12-18 Age Group

	With HH FE				Poisson			
	12-14	12-16	14-18	16-18	12-14	12-16	14-18	16-18
Treatment × Post	-0.022	0.043	-0.017	-0.005	-0.020	0.073	-0.001	-0.017
	(0.076)	(0.094)	(0.075)	(0.088)	(0.133)	(0.116)	(0.103)	(0.165)
Post Period	-0.035	-0.072	0.085	0.115	-0.095	-0.114	0.121	0.227
	(0.079)	(0.090)	(0.085)	(0.092)	(0.148)	(0.111)	(0.109)	(0.170)
Constant	0.580**	0.582*	0.664**	0.206				
	(0.283)	(0.304)	(0.266)	(0.335)				
Number of Obs.	1057	1233	1180	900	1057	1233	1180	900
$adjR^2$	.016	.022	.018	.012				
Prob > F	0.00	0.00	0.00	0.08	.	.	.	.
Household FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE × Region	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	No	No	No	No	No	No	No	No
Mean of the Dep. Var	0.57	0.81	0.76	0.57	0.57	0.81	0.76	0.57

Standard errors clustered at Woreda Level in Parenthesis
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 21: Probability of Observing a Household with an Outgoing Member due to Work

	OLS(LPM)				Probit				With HH FE			
	12-18	14-18	16-18	19-29	12-18	14-18	16-18	19-29	12-18	14-18	16-18	19-29
main												
Treatment × Post (d)	-0.064***	-0.064***	-0.065**	-0.008	-0.032***	-0.032***	-0.031**	0.004	-0.066***	-0.066**	-0.058*	-0.001
	(0.019)	(0.023)	(0.026)	(0.020)	(0.009)	(0.010)	(0.013)	(0.017)	(0.022)	(0.026)	(0.032)	(0.022)
Treatment (d)	-0.001	-0.009	-0.003	-0.015	-0.001	-0.008	-0.005	-0.009				
	(0.035)	(0.036)	(0.040)	(0.034)	(.)	(0.027)	(0.025)	(0.026)				
Post Period (d)	0.031**	0.033**	0.031	0.028	0.023**	0.024**	0.021*	0.019	0.034	0.039	0.023	0.029
	(0.015)	(0.016)	(0.019)	(0.024)	(0.011)	(0.012)	(0.012)	(0.017)	(0.025)	(0.026)	(0.029)	(0.025)
Number of Obs.	1318	1202	1004	1533	1318	1202	1004	1533	1318	1202	1004	1533
$adjR^2$	.011	.01	.0096	.022					.021	.021	.02	.039
Prob > F	0.00	0.00	0.01	0.00					0.00	0.00	0.00	0.00
Household FE	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	No

Standard errors clustered at Woreda Level in Parenthesis
(d) for discrete change of dummy variable from 0 to 1
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 22: Probability of Observing a Household with Female Incoming Member due to Marriage: By Age Groups

	Probit			
	12-18	14-18	16-18	19-29
main				
Treatment × Post (d)	0.013	0.025	0.025	0.005
	(0.014)	(0.021)	(0.023)	(0.013)
Treatment (d)	-0.008	-0.024	-0.029	0.008
	(0.014)	(0.017)	(0.020)	(0.014)
Post Period (d)	0.027***	0.020***	0.011*	-0.013
	(0.009)	(0.007)	(0.007)	(0.013)
Number of Obs.	1689	1689	1393	1801
	OLS			
	12-18	14-18	16-18	19-29
Treatment × Post	0.013	0.019	0.014	-0.003
	(0.012)	(0.013)	(0.010)	(0.016)
Treatment	-0.002	-0.009	-0.012	0.011
	(0.012)	(0.010)	(0.009)	(0.021)
Post Period	0.008	0.006	0.001	-0.009
	(0.006)	(0.006)	(0.004)	(0.014)
Constant	-0.000	-0.001	-0.017	0.020
	(0.038)	(0.035)	(0.019)	(0.028)
Number of Obs.	2185	2185	2185	2185
$adjR^2$	.0047	.0041	-.002	.011
Prob > F				
Household FE	No	No	No	No
Time FE	Yes	Yes	Yes	Yes
Time FE × Region	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
	With HH FE			
	12-18	14-18	16-18	19-29
Treatment × Post	0.012	0.019*	0.016	-0.006
	(0.010)	(0.011)	(0.010)	(0.016)
Post Period	0.009	0.003	-0.004	-0.010
	(0.013)	(0.009)	(0.007)	(0.014)
Constant	-0.007	-0.014	0.003	0.015
	(0.024)	(0.024)	(0.021)	(0.043)
Number of Obs.	2185	2185	2185	2185
$adjR^2$	.0032	.0032	-.0026	.014
Prob > F				
Household FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Time FE × Region	Yes	Yes	Yes	Yes
Region FE	No	No	No	No

Standard errors clustered at Woreda Level in Parenthesis
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

References

Andersson, Camilla, Alemu Mekonnen, and Jesper Stage, "Impacts of the Productive Safety Net Program in Ethiopia on livestock and tree holdings of rural households," *Journal of Development Economics*, 2011, *94* (1), 119 – 126.

Angelucci, Manuela, "Conditional Cash Transfer Programs, Credit Constraints, and Migration," *LABOUR*, 2012, *26* (1), 124–136.

Barrientos, Armando, "Social Transfers and Growth: What Do We Know? What Do We Need to Find Out?," *World Development*, 2012, *40* (1), 11 – 20.

Baughman, Reagan and Stacy Dickert-Conlin, "The earned income tax credit and fertility," *Journal of Population Economics*, 2009, *22* (3), 537–563.

_ and _, "The earned income tax credit and fertility," *Journal of Population Economics*, 2009, *22* (3), 537–563.

Becker, Gary S., "A Theory of Marriage: Part I," *Journal of Political Economy*, 1973, *81* (4), pp. 813–846.

_, "A Theory of Marriage: Part II," *Journal of Political Economy*, 1974, *82* (2), pp. S11–S26.

_ and H. Gregg Lewis, "On the Interaction between the Quantity and Quality of Children," *Journal of Political Economy*, 1973, *81* (2), pp. S279–S288.

Berhane, Guush, John Hoddinott, Neha Kumar, Alemayehu Seyoum Taffesse, Michael Tedia Diressie, Yisehac Yohannes, Rachel Sabates-Wheeler, Mulugeta Handino, Jermey Lind, Mulugeta Tefera, and Feyera Sima, "Evaluation of Ethiopia's Food Security Program: Documenting Progress in the Implementation of the Productive Safety Nets Programme and the Household Asset Building Programme," Mimeo, Washington, D.C:International Food Policy Research Institute 2011.

_ , _ , _ , and _, "The impact of Ethiopia's Productive Safety Nets and Household Asset Building Programme: 2006-2010," International Food Policy Research Institute 2011.

Binswanger, Hans P. and John McIntire, "Behavioral and Material Determinants of Production Relations in Land-Abundant Tropical Agriculture," *Economic Development and Cultural Change*, 1987, *36* (1), pp. 73–99.

Bobonis, Gustavo J., "The Impact of Conditional Cash Transfers on Marriage and Divorce," *Economic Development and Cultural Change*, 2011, *59* (2), pp. 281–312.

Brewer, Mike, Anita Ratcliffe, and Sarah dSmith, "Does welfare reform affect fertility? Evidence from the UK," *Journal of Population Economics*, 2012, *25* (1), 245–266.

_ , _ , and _, "Does welfare reform affect fertility? Evidence from the UK," *Journal of Population Economics*, 2012, *25* (1), 245–266.

Deaton, Angus and Christina Paxson, "Economies of Scale, Household Size, and the Demand for Food," *Journal of Political Economy*, 1998, *106* (5), pp. 897–930.

Fafchamps, Marcel, "Rrual Poverty, Risk and Development," Report Submitted to the Food and Agricultural Organization, Center for the Study if African Economies October 1999.

_ **and Agnes R. Quisumbing**, "Chapter 51 Household Formation and Marriage Markets in Rural Areas," in T. Paul Schultz and John A. Strauss, eds., , Vol. 4 of *Handbook of Development Economics*, Elsevier, 2007, pp. 3187 – 3247.

Hagen-Zanker, Jessica and Carmen Leon Himmelstine, "What do we know about the impact of social protection programmes on the decision to migrate?," *Migration and Development*, 2013, *2* (1), 117–131.

Hoddinott, John, Guush Berhane, Daniel O. Gilligan, Neha Kumar, and Alemayehu Seyoum Taffesse, "The Impact of Ethiopia's Productive Safety Net Programme and Related Transfers on Agricultural Productivity," *Journal of African Economies*, 2012, *21* (5).

Hosegood, Victoria, Anne Case, and Cally Ardington, "Labor Supply Responses to Large Social Transfers: Longitudinal Evidence from South Africa," *American Economic Journal: Applied Economics*, 2009, *1* (1), 22–48.

Kearney, Melissa, "Is There an Effect of Incremental Welfare Benefits on Fertility Behavior? A Look at the Family Cap," *Journal of Human Resources*, 2004, *39* (2), 295–325.

Lanjouw, Peter and Martin Ravallion, "Poverty and Household Size," *The Economic Journal*, 1995, *105* (433), pp. 1415–1434.

Milligan, Kevin, "Subsidizing the Stork: New Evidence on Tax Incentives and Fertility," *The Review of Economics and Statistics*, 2005, *87* (3), 267–290.

Nio-Zaraza, Miguel, Armando Barrientos, Samuel Hickey, and David Hulme, "Social Protection in Sub-Saharan Africa: Getting the Politics Right," *World Development*, 2012, *40* (1), 163 – 176.

Rosenzweig, Mark R., "Risk, Implicit Contracts and the Family in Rural Areas of Low-Income Countries," *The Economic Journal*, 1988, *98* (393), 1148–1170.

_ **and Kenneth I. Wolpin**, "Government Interventions and Household Behavour in a Develooping Country Anticipating the Unanticipated Consequences of Social Programs," *Journal of Development Economics*, 1982, *10*, 209–225.

_ **and Oded Stark**, "Consumption Smoothing, Migration, and Marriage: Evidence from Rural India," *Journal of Political Economy*, 1989, *97* (4), 905–926.

Rubalcava, Luis and Graciela Teruel, "Conditional Public Transfers and Living Arrangements in Rural Mexico," California Center for Population Research, On-line Working Paper Series 2006.

Schultz, T. Paul, "Demand for Children in Low Income Countries," *Handbook of population and Family Economics. Edited by M.R. Rosenzweig and O.Stark*, 1997, pp. 349–430.

Stark, Oded and David E. Bloom, "The New Economics of Labor Migration," *The American Economic Review*, 1985, *75* (2), pp. 173–178.

Stecklov, Guy, Paul Winters, Jessica Todd, and Ferdinando Regalia, "Unintended effects of poverty programmes on childbearing in less developed countries: Experimental evidence from Latin America," *Population Studies*, 2007, *61* (2), 125–140.

_ , _ , Marco Stampini, and Benjamin Davis, "Do Conditional Cash Transfers Influence Migration? A Study Using Experimental Data from the Mexican Progresa Program," *Demography*, 2005, *42* (4), pp. 769–790.

Todd, Jessica., Paul Winters, and Guy Stecklov, "Evaluating the impact of conditional cash transfer programs on fertility: the case of the Red de Proteccin Social in Nicaragua," *Journal of Population Economics*, 2012, *25* (1), 267–290.

Winters, Paul, Stecklov Guy, and Todd Jessica, "The Impact of Conditional Cash Transfers on household composition, fertility and migration in Central America," Technical Report 2006.

Chapter 2

Linking Cash Transfers with Labor Supply: *Assessing Impact on Household Structure*

Linking Cash Transfers with Labor Supply: *Assessing Impact on Household Structure*

Tseday J. Mekasha*

Abstract

We use a longitudinal data from the Ethiopian productive Safety net program, which constitutes labor intensive public works as one of its major components, to assess how income transfers and labor supply conditions attached to such programs affect household structure. Our identification strategy relies on changes in the participation status of households that is caused by the continuous re-targeting effort observed in the course of program implementation. This enables us to identify the effect of the program using differences between beneficiary and non-beneficiary households as well as differences in their before and after program participation. We found that both the financial gains from the program and the labor supply requirement associated with it have a positive impact on household size. While the income effect is found to be associated with an increase in the number of adolescent female household members, the labor supply effect works through an increase in the number of male household members. The relative increase in the number of adolescent female household members appears to be due to the (early) marriage delaying effect of the program. Contrary to findings of the existing literature on related programs elsewhere, we do not find evidence to associate the program with an increase in fertility.

Keywords:
JEL Classification: J10, J12, J13, J22, J23, I38, O15

*tseday.mekasha@econ.ku.dk

1 Introduction

In recent years, social protection programs that provide cash transfers to households either as a pure income transfer or conditional on certain requirements being fulfilled are becoming common in Africa. These programs range from age based cash transfer programs that target certain age groups, like children and the elderly as in the South African Child Support Grant and Old Age Pension Scheme, or broader cash transfer schemes that focus on the poor and most vulnerable. The latter is becoming more common in sub Saharan African countries where it is being promoted as a tool to fight poverty and bring sustainable changes in the livelihood of the poor (see Nino-Zarazúa et al. (2012), Barrientos (2012) and Schubert and Slater (2006)).

In the face of binding resource constraints, households' effort to fight poverty and bring meaningful changes in their livelihoods by and large depends on how resources are shared within the household. This in turn depends, among other things, on household structure. According to existing evidence in the literature household structure in low income countries, as captured by size and composition, responds to both economic gains and losses (see Rosenzweig (1988), Winters et al. (2009) and papers cited therein). In view of this, understanding the process of household formation and partition in the context of low income countries is crucial for ensuring effectiveness of interventions that aim to improve the livelihood of the poor.

However, little is known how cash transfers programs and conditions associated with them affect households' behavior in terms of changing their size and composition. In this regard, the bulk of the literature on Africa is on the South African social pension scheme assessing how stable income gains affect labor migration/labor supply response of prime aged adults (Bertrand et al. (2003), Posel et al. (2006), Hosegood et al. (2009)), living arrangements (Edmonds et al. (2005)), household composition (Maitra and Ray (2003)), among others. Although these studies provide useful insight, the generalizability of their results to poverty focused cash transfer programs is limited due to the distinct features of the social pension scheme in South Africa; in terms of eligibility criteria, target group and program objective.[1] It is thus important to fill the knowledge gap on how income gains from poverty focused programs and conditions attached to these programs affect household structure as this has a clear implication for the effectiveness of the programs in achieving their desired goals. In this paper we aim to do this using the Ethiopian Productive Safety Net program (PSNP) as a case in point.

Hoddinott and Mekasha (2015) show that participation in the Ethiopian Productive Safety Net Program (PSNP), where households receive cash transfer contingent on supplying labor to public works, is associated with an increase in household size. They further show that the increase in household size is mainly due to a relative increase in number of female household members. Even if these findings reveal interesting facts

[1] For a detailed discussion on how cash transfer programs based on age vulnerability like the South Africa social pension scheme differs from poverty focused cash transfer programs can be found in Nino-Zarazúa et al. (2012)

on the relationship between program participation and household structure, it is mute as to which component of the program induces the observed change in household structure. That is, does this change reflect households' response to the benefit, the conditions attached to it or both? Thus, in this paper we aim to explore to what extent the economic incentive following participation in the public works component of the Ethiopian Productive Safety Net Program and the labor supply conditionality tied with it have a role in explaining the change in household size and composition. In particular, effort is made to disentangle the income effect of participation from the effect that can originate from the labor supply requirement of the program.

On the one hand, by relaxing the financial constraint rural households are likely to face, public works cash transfers can induce beneficiaries to add new members, retain existing ones for longer than usual or facilitate household partitioning either through marriage or labor migration (Foster (1993)). On the other hand, the labor supply requirement attached to the income benefit is likely to introduce another resource constraint, particularly a time/labor constraint within the household. This may induce participant households to attract new(working age) members and/or retain their able-bodied members.[2] To what extent this additional constraint affects household structure depends on whether the constraint is binding or not, which in turn depends on the total labor pool of the household and how much time the household is left for agriculture and other household chores net of pubic work participation (see Hosegood et al. (2009)).[3] The effect of participation in PSNP public works on household structure will thus depend on the degree to which the income gains relax households' financial constraints and the extent to which the time constraint induced by the labor supply requirement is binding. It is therefore important to empirically quantify how each of these two mechanisms affects household structure and this is the main objective in this paper.

Although there are efforts to evaluate the effectiveness of the Ethiopian PSNP, the principal focus of these evaluations have been on the behavioral response of households in terms of targeted outcomes like agricultural investment and productivity, the level of food security and asset holding/building (Hoddinott et al. (2012), Berhane et al. (2011b) and Andersson et al. (2011)). The effect of the program on these intended outcomes by and large depends on how income gains from the program are shared and the extent to which the program affects labor allocation within the household. The results of this paper will thus add a different perspective to the assessment of the program effectiveness.

Our data relies on the four rounds of the Ethiopian Productive Safety Net surveys that have been conducted during the period 2006-2012. On top of participation information, these surveys have information on the level of actual benefits/payments households receive and the number labor days households worked on public works activities. Moreover, the longitudinal nature of the data coupled with the rich set

[2]One can also argue that, the time/labor constraint can also induce participants to send their small children to relatives as they will not have time to look after them.

[3]Given that all able bodied working age adults participate in public works, conditional on the household becoming eligible for the program, there can be a trade-off between public works participation and other household activities including agriculture and daily household chores.

of question on household demographics and migration status of each member in the household makes this data set suitable to address the question at hand. Our identification takes advantage of the variation created by the continuous re-targeting efforts during program implementation which changes the participation status of some of the households. In particular, the continuous re-targeting gives the data a unique structure which gives us the chance to observe beneficiary households both before and after becoming participants in public works program. We thus employ a differences-in-differences approach to identify the effect of interest.

The analysis reveals interesting findings: both the income effect from public works participation and the labor supply requirement attached to it are found to induce changes in household structure. We also find evidence pointing to differential impacts based on the attributes of household members including their age and gender. In particular, an increase in public works per capita income is associated with an increase in household size which mainly arises due to participants delaying the marriage of their adolescent female members. The income transfer is however associated with an increase in female marriage in the 19-29 age category. On the other hand, the labor supply requirement appears to be associated with an increase in number of male household members, particularly those in the 19-29 age category. Lastly, we do not find any evidence to suggest that public works income transfer increases number of members in the 0-6 age group. We instead find suggestive evidence associating an increase in public works labor months with an increase in the probability of observing a household with an outgoing member in 0-6 age group; which can be explained by the time constraint beneficiary households are likely facing for child care.

The rest of the paper proceeds as follows: Section II presents a brief overview of the Ethiopian productive Safety Net program with a particular emphasis on the public work component of the program, introduces the data and provides some descriptive statistics. While Section III describes the model specification and estimation techniques, the main findings and related discussions are presented in Section V. Finally, a concluding remark is given in Section VI.

2 Program Background, Data and Descriptive Statistics

2.1 The Ethiopian Productive Safety Net Program and its Public Works Component

Ethiopia has been reliant on emergency food assistance for decades. Following increased consensus on the limitations of such assistance in bringing sustainable solution to chronic food insecurity problems in the country, the government of Ethiopia has launched the National Food Security Program (EFSP) in 2003 with the aim of addressing the root causes of food insecurity. Later in 2005, the Ethiopian Productive Safety Net Program, henceforth PSNP, is started as one major component of the EFSP.

The main objective of PSNP is to smooth household consumption and prevent asset depletion through predictable cash transfer to poor and chronically food insecure households while building household and community level assets. The program runs with an annual budget of nearly 500 million US dollar reaching more than 7 million people and is among the largest social protection program in sub-Saharan Africa. Gilligan et al. (2008). The program is being implemented in eight regions across the country focusing on chronically food insecure woredas in these regions. The regions include: Amhara, Oromiya, SNNPR, Tigray, Harrari, Afar, Somiali and Dire Dawa. Targeting under PSNP is done using a combination of methods including geographic and community targeting where the latter includes the actual identification of beneficiary households by the community food security task force.

Public Works under PSNP[4]

PSNP has a public works and direct support components. The public work component is the major part where beneficiary households receive cash transfers conditional on participating in labor intensive public works activities that aim to build community assets. The direct support component is, on the other hand, unconditional transfer for labor constrained households that could not supply labor for public works.

- **Public Works Labor Supply**

Once a household becomes eligible to participate in PSNP, based on assessment of past history of unmet food needs and asset and wealth levels, the next step is to decide whether to classify the household as either a public works beneficiary or direct support beneficiary. This depends on the availability of able bodied members in the household.[5] If the household does not have any able bodied member, it will be classified as a direct support beneficiary. On the other hand, if some of the members are able bodied and some are not, the able bodied members can work on behalf of the dependents to secure payment entitlement on behalf of the latter.[6]

However, an able bodied member can work for dependents only up to a labor cap, which is 15 days per month per economically active member. In view of this, the maximum number of days a public work beneficiary household can work is given by *No. of able-bodied individuals* $\times$ *15*. This shows that there is no uniform labor cap that is applied at the household level. Within a household, since each able bodied member is required to work 5 days per month, the 15 days labor cap implies that an able bodied member can work for himself/herself and for 2 more other dependents.

The labor cap particularly applies on households whose number of dependents exceed the number of able bodied members. In particular, in view of the above infor-

[4]The discussion below heavily draws on the PSNP program Implementation Manual 2010

[5]For a household member to be eligible to participate in public works, the member should be able bodied and above 15 years of age; and particularly for females, they should not be pregnant after 4 months and should not be a lactating woman within the first 10 months after giving birth.

[6]Targeting after 2008 becomes Full Family Targeting (FFT), meaning that once the household becomes eligible to participate in the program, all family members are entitled to get the benefit.

mation, it can easily be seen that a labor cap will be applied if the ratio of dependents to able bodied members is greater than 2. Such households will still be classified as public work beneficiaries but receive a combination of public works and direct support transfer.

- **Public Works Payment**

In a given year and for a given wage rate, beneficiary households are required to participate in public works for 6 months in non-agricultural season and hence are entitled to a certain amount of payment.[7] Over the years, public works wage rates have been adjusted for the rise in the cost of living. Accordingly the wage rates in the respective years were: 6 Birr (Ethiopian currency) in 2006 and 2007, 8 Birr in 2008, 10 Birr in 2009 and 2010, 12 Birr in 2011 and 14 Birr in 2012. After the introduction of FFT in 2008, the entitlement varies with the number of household members. Accordingly, the annual total entitlement for a household under FFT is calculated as:

$$\begin{aligned} \textit{Entitlement} &= \textit{Total Household Size} \times \textit{5 days of Work/ month/able bodied adult member} \\ &\quad \times \textit{Wage Rate} \times \textit{6 months} \end{aligned} \tag{2.1.1}$$

2.2 Data and Descriptive Statistics

The data we use for the empirical analysis relies on the four rounds of the Ethiopian Productive Safety Net (PSNP) surveys that have been conducted over the period 2006-2012 in four PSNP beneficiary regions, namely Amhara, Oromiya, SNNPR and Tigray.[8] Surveys were conducted in June/July of 2006, 2008, 2010 and 2012. Thus, for the survey years, data on the number of public work days and amount of public works payment is observed only for the first five months of the respective survey years. On the other hand, for years in 2007, 2009 and 2011, information on public work days and payment is observed for the whole year. Moreover, while data on household size and composition is available only for the survey years, information on entering and exiting members is available for all years.

One particular aspect of the program that we make use of in this analysis is the continuous re-targeting that is carried out in the course of program implementation. In particular, every year the Community Food Security Task force updates the existing PSNP client list with the aim of identifying households that are ready to graduate, correcting for exclusion error and selecting new entrants. This aspect of the program enables us to form a longitudinal data by focusing on beneficiary households that are late comers to the program.

[7]Note that before the introduction of the Full Family targeting (FFT), payment entitlements do not vary by household size. That is, after FFT, the benefit becomes per household member rather than per household.

[8]Details about the sampling method is given in Berhane et al. (2011a) as well as Hoddinott et al. (2012).

Accordingly, we get three groups/batches of treatment households, where a household is considered as a treatment household if it receives a positive amount of public works payment. For the empirical analysis we have used two different (continuous) measures of levels of treatment. While one captures the amount of public works payments received by the household and the other captures the number of labor days supplied by the household for public work activities.

Regarding classification of beneficiary households into different batches, the first treatment batch includes households that were not part of the program in 2006 but become beneficiaries in 2008 only or both in 2007 and 2008. For these households, 2006 is used as their before period. In doing so effort is made to select comparable number of control/non-beneficiary households from the same geographical area as the beneficiary households. The second treatment batch includes households that participated in the program during 2007-2010, or 2008-2010, or 2009-2010 or in 2010 only. The *before* periods for these households vary depending on when they started to participate in the program and it can respectively be stated as 2006, 2006-2007, 2006-2008, 2006- 2009. Finally, the same thing applies to the last treatment batch households that stayed in the program until 2012. Again, the before period varies based on the year they joined the program. In all cases, direct support beneficiaries are not included.

For each of the above treatment batch, the corresponding comparison groups are households that never received public works payments in the treatment period in question. For the first treatment batch households, for example, households that did not participate in the program in the three years from 2006-2008 are used as comparison households. And the same follows for the other treatment batch households. To increase comparability, we give priority to households that were not beneficiaries in early years but become part of the program in later years in our construction of comparison group households as well as for households that were initially entitled to participate but did not actually get any payment. Moreover, geographical considerations are also used to get a comparable number of households from the same 'Kebele' as in the treatment households.[9] In our construction of comparison group households, direct support beneficiaries are excluded.

The final sample, excluding the years 2007, 2009 and 2011 includes 2871 total observation comprising of 1414 beneficiary and 1457 non-beneficiary households. The data is then collapsed so that for each household we have one before and one after period. That is, for outcome variables that are continuous (for e.g. number of members) we take the respective averages of the before and after periods. Whereas for binary outcome variables (for e.g. indicators for households having an incoming/outgoing members), we generate an indicator variable which equals 1 if the household has at least one incoming/outgoing member and zero otherwise, in the respective periods. The same applies for covariates as well. On the other hand, for our treatment level indicator variables (payment and public work days), we take the average in the post period and entered it in both periods as these variables are used as indicators of levels of treatment. Following this, we end up with a total observation of 1334 households. Descriptive statistics of the variables used in this paper is presented in the appendix

[9] Kebele is the smallest administrative unit in Ethiopia

Table 9-Table 13

3 Empirical Model and Identification

Model

This section presents the empirical model that we use to disentangle the channels through which public works participation affects household structure, as measured by its size and composition. In particular, as alluded to in the introduction, the main objective is to empirically assess how households change their size and composition following financial gains from public works programs as well as to the labor supply requirements attached to it.

In theory, holding, among others, public works labor days constant, an increase in public works income transfer is expected to increase household size (and its disaggregated components) by relaxing households liquidity constraint. Assuming economies of scale in household chores like child care and food preparation, there can potentially be gains from increasing size of the household. Presence of household public goods can also be an additional source of economies of scale that can justify increasing household size (Deaton and Paxson (1998) and Lanjouw and Ravallion (1995)). However, after a certain point the non-rivalry nature of the household public goods will cease and congestion can arise. There is thus a certain optimal household size beyond which the cost of increasing household size outweighs the potential benefits (Fafchamps and Quisumbing (2007)). The above arguments shows that there is a non-linear relationship between household size and income; we thus do not expect households to steadily increase their size following increases in income. In our empirical model given below we have therefore included a quadratic term in income transfers to capture this potential non-linear relationship(see Equation 3.0.1).

Since the net change in aggregate household size (its disaggregated gender and age components) is a combined effect of the flow of members in and out of the household, we also estimate, separately, the probability of observing a household with at least one incoming or outgoing member.

In relation to outgoing members, an increase in the level of public works payments, all else equal, can increase or decrease the probability that the household has an outmigrant member. On the one hand, households can use the additional income to finance activities that would have been impossible without the income transfer, for example, it can be used to finance migration of working age adults or wedding ceremonies. On the other hand, the income effect from transfer payments can increase households' capacity to support more members inducing them to increase retention of their existing members. Due to this, the effect of transfer payments on probability of observing outmigration (and hence on household size) is ambiguous; but will have important implications on the age and gender composition. Similar argument goes to incoming members. In particular, an increase in transfer payments is likely to increase a household's probability of adding new member/s, either through birth or hosting relatives.

In relation to the labor supply requirement, an increase in public works labor days (labor supply for public works) can compete with the labor time beneficiaries have for agriculture and other livelihood activities. Holding, among others, transfer payments constant, this can increase the demand for working age labor within the household and hence induce beneficiaries to invite new adult members or retain their working age adults longer than usual. We would therefore expect an increase in public works labor days to affect household size and composition; as well as the probability of having outgoing or incoming members.

As pointed out before, the impact of supplying labor to public works on household structure will, however, depend on the extent to which the household's time/labor constraint is binding. And the labor pressure (constraint) ensued by supplying labor to public works will in turn hinges on the number of able bodied household members and the size of agricultural land that is available to the household. This is important as it determines the amount of labor the household is left with for agriculture and other household chores, net of public works participation.[10] Thus, controlling for the per capita land available to the household, the number of per capita (average) public work days can capture/measure the level of pressure the household is likely to face from supplying labor to public works.[11] Conditional on per capita land, an increase in per capita public works labor supply would be expected to lead to an increase in household size, but only after a certain point, beyond which the household will face a labor pressure. Or, it should be expected to instigate inflow of members and/or increase retention of current members only after a certain point. To capture this potential non-linear relationship, we have therefore included square of the per capita public works days in the estimated model given in Equation 3.0.1.

In view of the above discussion the empirical model to be estimated in this paper takes the following form:

$$\begin{aligned} Y_{it} \;=\; & \alpha_1 + \beta_1 Post + \beta_2 payment_{t-1} * Post + \beta_3 (payment_{t-1})^2 * Post \\ & + \beta_4 days_{t-1} * Post + \beta_5 (days_{t-1})^2 * Post + \beta_6 X_{it} + \mu_i + u_{it} \end{aligned} \tag{3.0.1}$$

where Y_{it} refers to aggregate household size or its disaggregated components, number of members by age and gender; or alternatively a binary variable which equals one if the household has at least one incoming (outgoing) member and zero otherwise. *Post* is a dummy which equals one for (after) treatment period, $days_{t-1}$ is lag of public work days per economically active members in the household (measured in labor months), $payment_{t-1}$ is lagged public work transfer payments divided by average size

[10]For instance, households that supply the same per capita (average) public works days can face different degree/level of labor pressure depending on the total labor days they have left for agriculture.

[11]Per capita land and per capita labor days are calculated/given by dividing total land and total public works days by the number of economically active members, respectively.

of the household in the sample period, X_{it} captures other time varying factors that can potentially affect both the outcome and the main variables of interest (i.e) public work labor months and payment. Finally, μ_i is household fixed effect capturing time invariant household characteristics. Other control variables include: household head characteristics, measures of household economic and wealth status, poverty perception, shocks to crops and livestock, region dummies, per capita land holding, duration of participation as well as the number of years the household is observed in the data.

In Equation 3.0.1, both the amount of public works transfer payment and labor are expressed in per capita terms to effectively capture the actual benefit the household is getting and the actual labor pressure the household is facing, respectively.[12] The parameters of interest in Equation 3.0.1 are β_2, β_3, β_4 and β_5.

Estimation Issues

Equation 3.0.1 is estimated using Differences-in-Differences approach. In particular, we use the variation in the level of treatment between beneficiaries and non-beneficiaries coupled with the overtime variation in level of treatment (payment and labor supply) as we observe beneficiaries in both the before and after their treatment period. On top of this, since we are using continuous treatment indicators, we are also making use of the variation in the level of treatment across beneficiary households to identify the effect of interest.

Although households can be observed in the data for more than two years, the data that we used for estimation is averaged over the before and after periods so that each household is observed only for two period periods. In particular, for the treatment indicator variables (public work payments and labor months) we take average of the lagged per capita payment received and lagged per capita labor months worked over their respective treatment periods. For the outcome variables and other covariates, we take averages of the variables in the before and after periods.

In all estimations, we have included ”duration” of participation and the ”number of years” each household is observed in the data. The latter is important because the wider the time window we observe a household in our data (regardless of participation status), the more likely it will be to observe an incoming or outgoing member.[13]

One thing that is worth noting in relation to the variation in public works per capita payment is that, once we hold labor supply to public works constant, the across beneficiary variation in the level of payments mainly comes from the variation in public works wage rates as well as transfers from the direct support component of the program. The latter is made possible since labor capped households can receive a mix of benefits (direct support plus public works) as they will not have enough able

[12]The use of per capita payment does not matter in fixed effects estimations as the household fixed effect absorbs whether the household is large or small.

[13]It is worth noting that the variable ”number of years” is different from duration in the sense that the former also takes into account the number of years we observe the control/comparison households in our data, while the variable ”duration” only takes positive values for treatment households.

bodied members to secure payment entitlement for dependents. Moreover, part of the variation can also be due to differences in payments overtime and across geographical locations, where the latter can arise due to differences in program implementation. For instance, there are cases where we observe households receiving different levels of payments for the same amount of public work labor days.

One potential challenge in estimating Equation 3.0.1 is the issue of endogeneity that can arise since shocks to household size and composition are also likely to be correlated with our variables of interest (public work payments and labor months). To minimize this problem, we use lagged per capita payments and lagged per capita public works labor supply as these are less likely to be correlated with shocks at period t. This amounts to assuming that shocks to household size at time t are orthogonal to shocks at $t-1$, $u_{it} \perp u_{it-1}$ ensuring that the errors are uncorrelated with past, present and future values of the regressors.[14]

4 Results and Discussion

As we have indicated in the introduction, Hoddinot and Mekasha (2015) find participation in the Ethiopian Productive Safety Net Program (PSNP) to be associated with an increase in household size. Given the nature of the program, where participants receive cash transfers conditional on supplying labor to public works activities, two mechanisms might be at work behind this change in household size. While one mechanism may work through relaxing households' income constraints, the other mechanism may operate through creating a labor/time constraint on participant households. In view of this, there are two important questions that we address in this paper: 1) If and to what extent these two mechanisms affect household demographic dynamics 2) Do these mechanisms have differential impact based on member attributes like gender and age? Answering these two questions will be the focus of the following sub sections.

In all estimations, standard errors are clustered at Woreda level in order to allow for arbitrary autocorrelation of the errors across households within a Woreda.[15] Our unit of analysis in all regressions is the household. Moreover, in all of our estimations we include a host of covariates that can potentially affect both household size/composition and our variables of interest (public work payments and public work labor supply). For the sake of brevity, we focus on our variables of interest and do not report coefficients of the other covariates. Unless otherwise stated, all regressions include: household head characteristics, measures of household wealth, economic status and poverty perceptions, shocks to crops and livestock, per capita land size, region dummies, duration

[14]We take the lags of these variables before collapsing the data into two periods. Using the lag of per capita payment and public work labor months also helps us to use full year information on payment and public work labor months instead of five month information available during survey years. As indicated in subsection 2.2, for the survey years (2006, 2008, 2010 and 2012), public works payment and public work days worked are observed only for five months (January-June).

[15]Woreda is the third administrative unit in Ethiopia next to Region and Zone.

of participation in the public works program and the number of years the household is observed in the data.

As it is discussed in section 3, we also include the quadratic terms of our variables of interest and their respective interactions with the post treatment period indicator.

In the DiD model, the parameters of interest are given by the coefficient of the interaction term between the post treatment period indicator and public works payment, public works labor months and their square terms. While interaction terms that involve public work payment are meant to capture the income effect of participation in labor intensive public works, the labor pressure effect that goes beyond and above the income effect is captured by interaction terms that involve public works labor supply. In all cases, we have demeaned the treatment intensity indicators using the average level of treatment in the sample so that the coefficients can be interpreted as responses of households receiving the average levels of treatment.

4.1 Financial Incentives and Changes in Household Size

This subsection presents results on how financial gains from participation in labor intensive public works affect household structure, controlling for, among others, labor supply to public works activities. We start with a discussion about the income effect of the program on aggregate household size and its components, numbers of male and female household members. This will then be followed by a discussion about the income effect on the numbers of male and female household members in the different age categories. In each of the cases, we estimate Equation 3.0.1 using OLS, OLS with household fixed effects and fixed effects Poisson estimation techniques. The fixed effects Poisson estimation is applied in view of the count data nature of the dependent variable; and the coefficients from this estimation have a percentage interpretation.

Coming to the results, holding the number of public work months and other covariates in the model constant, an increase in public works income transfer is found to increase household size but at a decreasing rate, although the squared term is not precisely estimated (see Table 1).

Regarding the size of the income effect, on average, an additional one thousand birr increase in public works per capita income is found to increase household size by about 1.2 members as can be seen from Column 1 of Table 1. In view of the mean household size in the sample, which is 5.54, the estimate amounts to a 22 percent increase in household size following an additional 1000 birr per capita public works income payment per year.[16] Following a similar increase in per capita income payments from public works, the corresponding coefficient estimate from the fixed effects Poisson estimation also shows a 19.3 percent increase in household size.[17]

[16] At the current exchange rate, 1 Ethiopian birr is approximately 0.05 USD, thus a 1000 birr increase is about 50 USD.

[17] Given the average per capita transfer in the sample, which is 263 birr, a 1000 birr increase in per capita payment amounts to a 300 percent increase in the level of public work transfer.

Looking at the impact of the income transfer on the numbers of male and female household members (gender disaggregation), the income effect of the public works transfer payments appear to vary by gender. In particular, the income effect of the cash transfers observed at aggregate household level appear to come solely from the effect on the number of female household members. As can be seen from both the fixed effects and Poisson estimations results, the coefficient of the level of payments on the number of female household members is positive and statistically significant. On the other hand, in the case of the number of male household members, the income effect of the public works cash transfers is found to be statistically indistinguishable from zero in all cases.

In relation to the magnitude of the income effect observed for females, an increase in public works income is found to be associated with an increase in the number of female household members by 0.962 members. More specifically, as can be observed from the fixed effects estimation result (see Column 2 of Table 1), a 1000 birr increase in per capita public works income payments is found to be associated with a 36 percent increase in the number of female household members. The corresponding estimate for females in the Poisson model also tells similar story. (see Column 5 of Table Table 1).

The above findings makes it clear that the income effect of an in increase in public works cash payments tend to work only through increasing the number of female household members. The observed increase in the number of female household members can be either due to an increase in the number of incoming female members or increased tendency of retaining them for a longer period following an increase in the intensity of treatment (payment). Since we have data on migration of members in and out of households, we will be able to test which channel is in play here. Before that, however, it is interesting to see which age group, among female household members, is deriving the observed increase in number of female household members. Even among female members, households may tend to increase their female members in certain age groups over others. In order to test if there is such heterogeneity in the income effect, based on the age of female members, we split the sample of female members into 3 different age groups and run separate regressions for each age group.[18]

The result for the number of females disaggregated by age is presented in Table 2. As can be seen from this table the estimated impact of income transfers on aggregate household size and on the number of female members appears to be manifested through an increase in the number of female household members in the 12-18 age group. However, as can be seen from the negative coefficient of the quadratic terms (Sq. Payment×Post) the effect is decreasing with the level of income transfers. In all cases, the coefficient estimates for both the linear and square terms are found to be statistically significant at 1 percent. As indicated in section 3 the observed non-linear relationship between public works income transfers and household size is theoretically expected; we do not expect households to expand their size steadily as they get more income. From a theoretical standpoint, because of the presence of a diminishing returns

[18]We focus on these three age categories as females in these age groups are the ones that are more likely to either move-into households or are the likely candidates to be retained (kept for longer in the household). As can be seen from Figure 2, this is also supported by the data.

in terms of sharing household public goods, a linear relationship between household size and public works transfer payments is less tenable (See also Foster and Rosenzweig (2002)). It is rather reasonable to assume that there is a point of congestion beyond which the cost of adding a new member outweighs its benefits. Contrary to the results observed for females, the age disaggregation in the case of male household members does not reveal much as is also observed at the aggregate level (See Table 3).

Taken together the evidence thus far shows that the income/financial gain from participating in labor intensive public works activities induces an increase in household size which is only explained by an increase in the number of female members in the 12-18 age group. In the next sub section, we will explore whether the observed heterogeneous (across gender and age groups) impact observed for income transfers is also reflected when we look into the labor supply effect of the program on household size and composition.

4.2 Impact of Labor Supply Conditionality

In this sub section we examine if the labor supply conditions attached to the income transfer has impact on household size and composition and if there is differential impact by age and gender of the household members. As argued before, the labor supply condition requires every economically active household member (regardless of gender) to supply labor for public works. In some cases able bodied members of the household will have to work on behalf of the dependents. There is likely to a tradeoff between supplying labor for public works and engaging in own agricultural and other household activities including child care, food preparation and other tasks related to livelihood (eg. herding of livestock and preparation of land for the next agricultural season etc). It is therefore reasonable to assume that the labor supply requirement of the program can increase demand for labor within the household .We expect the public works labor supply requirement to have a more pronounced effect on economically active members of the household. However, it can also be the case that the pressure on labor demand can also lead to an increase in the number of household members in non-working age groups. [19]

Estimation results on the labor supply effect of the program on household size and composition depicted in Table 1 reveal some interesting facts. In particular, as can be seen from this table, holding income from public works and other covariates in the model constant, we do not find any evidence to suggest an increase in labor months is associated with the number of female household members. In all cases the coefficients of ($PW\ Days \times Post$) and ($Sq.\ PW\ Days \times Post$) terms are not statistically significant. For the number of male members, on the other hand, there is some evidence showing that an increase in households labor pressure (per capita public works labor days) leads to an increase in the number of male members. Even if the linear term capturing

[19] In rural Ethiopia it is not uncommon for young boys and girls to take part in some labor activities, for eg. herding, water fetching, food preparation and babysitting. In view of this, it is not unrealistic to expect a labor supply effect on the number of non-working age (7-11 or 12-14) members.

the intensity of treatment is negative, but statistically insignificant, the positive and statistically significant coefficient on the quadratic term indicates that the labor pressure effect kicks in only after a certain level of per capita labor supply. To some degree, this effect is also reflected on aggregate household size (see Table 1).

A closer look at the age disaggregated results reported in Table 3 shows that the labor supply effect of public work participation observed for male members is mainly reflected in the number of males in the 19-29 age group. In particular, conditional on public works income gains, an increase in public works labor months is found to be positively associated with the number of male household members in the 19-29 age category. This result appears to be consistent in both the fixed effects and Poisson estimations.[20]

In a nutshell, both the income gain and the labor supply requirement/condition tied to it are found to explain the increase in aggregate household size. In particular, while the income effect of participation is manifested only through a relative increase in the number of female household members in early and late adolescent age groups, the labor supply effect mainly works by increasing the number of working age (19-29) male members . However, as noted before, since the observed increase in aggregate household size can be either through addition of new members or retention of existing ones, it is early to conclude that beneficiary households are expanding their size following the income and labor supply effects of program participation.

The change in household size and composition observed above is obviously a net/combined effect of changes in the number of members moving in and out of the household. It is therefore important to assess how the income and labor supply effects of participation influences households' behavior in terms of their decision to retain current members and welcome new ones. This is discussed in the section below with a particular focus on the age-gender categories that are found to be important in explaining the increase in household size observed at the aggregate level.

[20]For female members, even if the age disaggregated result in Columns 5 and 8 of Table 2 show that the total effect of an increase in public works labor months is positive, the coefficient of the squared terms is statistically indistinguishable from zero.

Table 1: Dependent Variable: Household Size, Number of Females and Males

	OLS			HH FE			FE Poisson		
	HH Size	Num Fem	Num Male	HH Size	Num Fem	Num Male	HH Size	Num Fem	Num Male
main									
Payment× Post	0.885	0.981**	-0.074	1.224**	0.962***	0.254	0.193**	0.326***	0.071
	(0.585)	(0.394)	(0.360)	(0.515)	(0.348)	(0.341)	(0.093)	(0.119)	(0.130)
Sq. Payment× Post	-0.484	-0.439*	-0.052	-0.410	-0.405	-0.000	-0.041	-0.110	0.005
	(0.342)	(0.243)	(0.201)	(0.347)	(0.266)	(0.197)	(0.103)	(0.149)	(0.118)
PW Days× Post	-0.586	-0.624	-0.008	-1.275	-0.778	-0.478	-0.170	-0.260	-0.040
	(1.054)	(0.682)	(0.742)	(0.967)	(0.588)	(0.776)	(0.195)	(0.236)	(0.329)
Sq. PW Days× Post	3.562	0.721	2.759	6.722*	1.782	4.936***	1.194	0.749	1.451*
	(3.709)	(2.639)	(1.787)	(3.677)	(2.540)	(1.808)	(0.736)	(0.943)	(0.818)
Constant	3.723***	1.815***	1.317***	5.007***	3.019***	2.012***			
	(0.566)	(0.372)	(0.446)	(0.480)	(0.430)	(0.296)			
Number of Obs.	1014	1010	1008	1014	1010	1008	1014	1010	1008
$adjR^2$	.13	.051	.1	.11	.057	.084			
Prob > F	0.00	0.00	0.00	0.00	0.00	0.00	.	.	.
Household FE	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	-	-	-	-	-	-
Mean of the Dep. Var	5.54	2.67	2.89	5.54	2.67	2.89	5.54	2.67	2.89

Standard errors clustered at Woreda Level in Parenthesis
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Dependent Variable: Number of Female Household Members by Age

	OLS			With HH FE			Poisson		
	7-11	12-18	19-29	7-11	12-18	19-29	7-11	12-18	19-29
main									
Payment× Post	0.236	1.742***	0.254	0.227	1.621***	0.282	0.137	1.904***	0.376
	(0.427)	(0.345)	(0.374)	(0.442)	(0.362)	(0.374)	(0.593)	(0.449)	(0.613)
Sq. Payment× Post	-0.409	-0.952***	-0.378	-0.374	-0.851**	-0.340	-0.421	-1.272***	-1.266
	(0.368)	(0.321)	(0.291)	(0.403)	(0.344)	(0.284)	(0.919)	(0.408)	(0.918)
PW Days× Post	0.124	-1.900**	-0.661	0.191	-1.885**	-0.610	0.614	-2.145*	-0.918
	(0.698)	(0.730)	(0.660)	(0.664)	(0.781)	(0.709)	(0.871)	(1.142)	(1.120)
Sq. PW Days× Post	-1.881	0.754	1.738	-1.213	1.569	3.249	-1.451	3.982	7.185
	(1.590)	(1.842)	(2.000)	(1.833)	(1.998)	(2.572)	(2.674)	(3.301)	(4.768)
Post Period	0.289**	0.123	0.074	0.243	0.000	0.035	0.310	-0.072	0.017
	(0.122)	(0.122)	(0.104)	(0.161)	(0.147)	(0.138)	(0.213)	(0.202)	(0.210)
Constant	0.619***	0.555**	1.032***	0.641*	0.567	1.817***			
	(0.173)	(0.272)	(0.178)	(0.376)	(0.440)	(0.239)			
Number of Obs.	570	582	592	570	582	592	570	582	592
$adjR^2$	.0081	.067	.051	.02	.1	.064			
Prob > F	0.00	0.00	0.00	0.03	0.00	0.00	.	.	.
Household FE	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	-	-	-	-	-	-
Mean of the Dep. Var	0.74	0.85	0.69	0.74	0.85	0.69	0.74	0.85	0.69

Standard errors clustered at Woreda Level in Parenthesis
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Dependent Variable: Number of Male Household Members by Age

	OLS			With HH FE			Poisson		
	7-11	12-18	19-29	7-11	12-18	19-29	7-11	12-18	19-29
main									
Payment× Post	-0.394	-0.115	0.025	-0.364	-0.004	0.017	-0.475	-0.040	0.360
	(0.498)	(0.437)	(0.352)	(0.498)	(0.419)	(0.368)	(0.677)	(0.444)	(0.601)
Sq. Payment× Post	0.674	0.134	0.244	0.672	0.107	0.776	0.945	0.057	0.439
	(0.449)	(0.314)	(0.776)	(0.415)	(0.311)	(1.042)	(0.610)	(0.303)	(2.238)
PW Days× Post	1.043	0.673	-1.280*	0.675	0.952	-0.876	0.598	1.277	-2.312
	(0.844)	(0.885)	(0.645)	(0.832)	(0.867)	(0.648)	(0.996)	(0.992)	(1.662)
Sq. PW Days× Post	-0.015	-1.108	5.609**	-0.359	-1.657	7.611***	1.394	-1.948	17.125**
	(2.030)	(2.043)	(2.179)	(1.869)	(1.898)	(2.183)	(2.457)	(2.629)	(7.691)
Post Period	-0.001	0.158	-0.314**	-0.008	0.266*	-0.414**	-0.112	0.279*	-0.845**
	(0.149)	(0.127)	(0.150)	(0.177)	(0.134)	(0.161)	(0.238)	(0.153)	(0.358)
Constant	1.202***	1.093***	0.948***	1.198***	0.630*	1.293***			
	(0.250)	(0.248)	(0.180)	(0.365)	(0.329)	(0.395)			
Number of Obs.	584	644	602	584	644	602	584	644	602
$adjR^2$	.011	.037	.07	.019	.076	.12			
Prob > F	0.00	0.00	0.00	0.00	0.00	0.00	.	.	.
Household FE	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	-	-	-	-	-	-
Mean of the Dep. Var	0.78	0.98	0.71	0.78	0.98	0.71	0.78	0.98	0.71

Standard errors clustered at Woreda Level in Parenthesis
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4.3 Does the Program affect Member Retention/Out-Migration?

Income gains from cash transfer programs and the conditions attached to them can either facilitate or deter out-migration of household members. In particular, if the cash transfer is big enough to relax households' resource constraints, it can induce out-migration of working age members (among others, see Hosegood et al. (2009)). On the other hand, it can also be argued that income gains from cash transfers can give households the capacity to support more members inducing them to retain their members for longer period. This member retention tendency on the part of program beneficiary households will however be optimal as long as the cost (both the direct and opportunity costs) of keeping a member does not exceed the benefit.[21] In a similar manner, conditions attached to cash transfer programs can also deter out-migration, for example if the part or all of the benefit is going to be lost when the member moves out of the household (see for eg. Stecklov et al. (2005)).

Against this background, the observed increase in aggregate household size, particularly the increase in the number of adolescent females and working age male members found in this paper can plausibly be explained by the retention of current members. As indicated above, households' decision whether to retain their members or not can be influenced by the income and/or labor supply effects. Before we explore this further using regression analysis, it is important to discuss why members leave the household. We have illustrated the reasons given for out-migration of male and female household members in Figure 1.

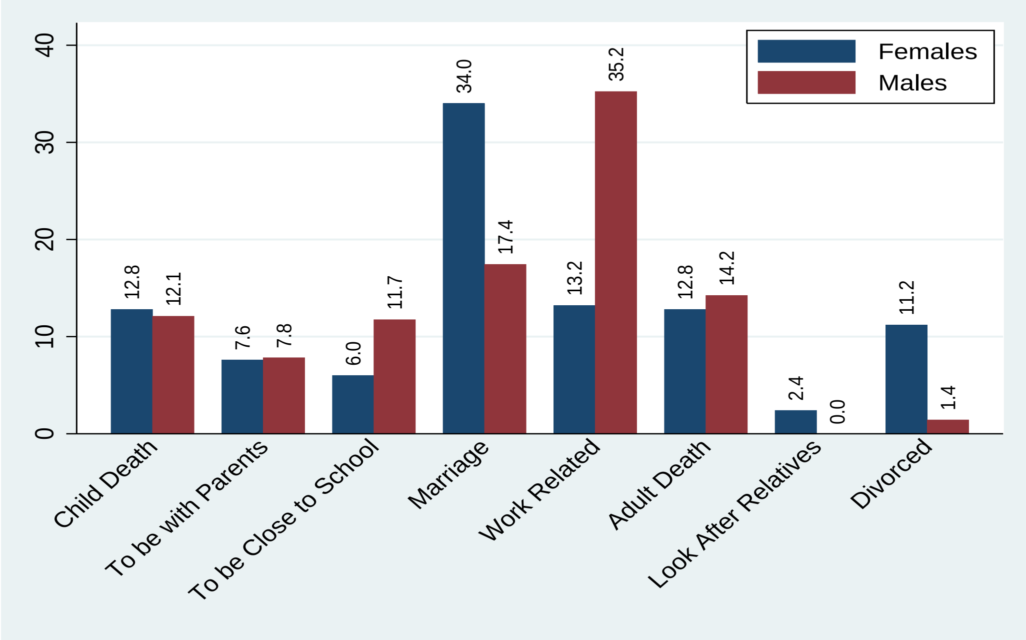
Graphical Analysis

As can be seen from Figure 1, which is done using the member level data, members leave the household for various reasons, and for some of these reasons there are clear gender differences. For example, for females, the major reason for leaving the household is found to be marriage. Of the total number of females that move out of the household, 34 percent are due to marriage. The corresponding figure for males is only 17 percent. In the case of male members, the main reason indicated for moving out of the household is due to work related reasons. This accounts for about 35 percent of the total male move outs and it is more than double when compared to females. Only 13 percent of the total female move out is due to work related reason.[22]

[21] The direct costs include costs for food, clothing and other costs of sharing household public goods, while the opportunity costs include the remittance potential migrants will send to the family (See Stark and Bloom (1985). On the other hand, in addition to the social and psychological benefits of living together, the benefits of member retention can be, for instance, if the stayer contributes to the income or livelihood of the household or if s/he supports the household in non-monetary terms, for example through his/her labor.

[22] Education (to be close to school) and divorce are the other two reason where we observe substantial difference between males and females. While a relatively higher proportion of male members indicate education as a reason for out migration, out migration due to divorce is clearly higher for females (11.2 percent for females vs 1.4 percent for males)

Figure 1: Reasons for Moving Out of the Household: By Gender

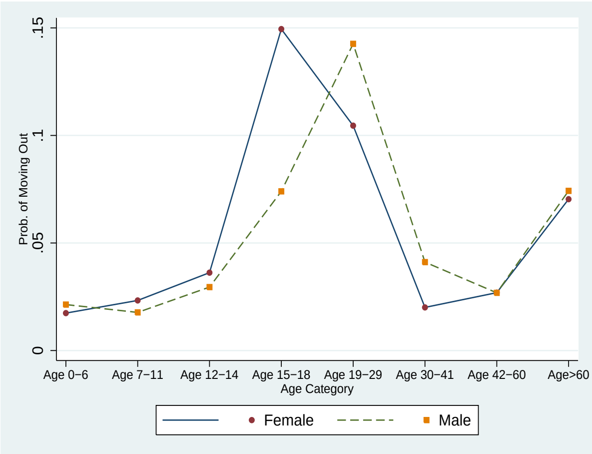


From the above figure it is clear that out migration due to marriage and work are the two most important reasons for members to move out of the household. In view of this, if member retention (or deterring out migration) is the main explanation for the observed increase in the number of female and male members, we should expect the financial gains from participation and the labor supply requirement attached to it to be inversely related with out-migration of current members due to marriage and work, respectively.

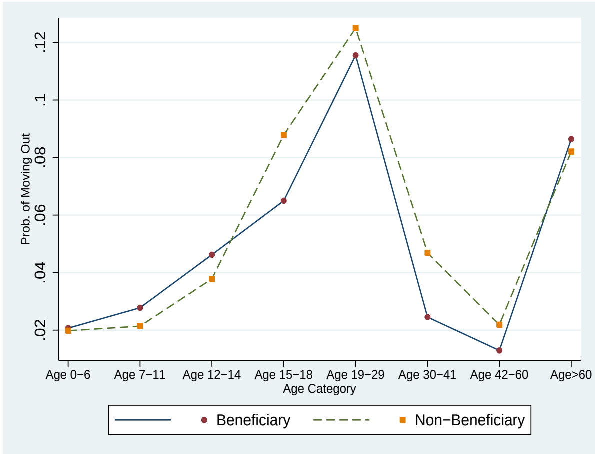
In order to assess if certain age groups are more likely to leave the household compared to others, and see if there is a differential pattern by gender and participation status, we have plotted migration propensities of the different age groups in Figure 2.

As can be seen from Panel (A), the outmigration propensity for females in the 15-18 age group is substantially larger when compared to that of males in the same age category. While the opposite is true in the case of the 19-29 age category, we do not see any substantial difference between the two genders for other age categories. Panel (b) plots the outmigration propensities of beneficiary and non-beneficiary household members across age groups. This is a done regardless of the gender of household members and it does not show any clear difference between beneficiary and non-beneficiary households. In Panel (c) we compare outmigration probabilities of female members in beneficiary and non-beneficiary households. Panel (d) does the same for male household members. As can be seen from Panel (c), there is a clear difference in the outmigration propensities of female members in the age groups (15-18 and 19-29) where marriage and work related out migrations are more likely to occur. For these age groups, females in non-beneficiary households have higher probabilities of out-migration. For the 19-29 age group, while there is a steeper decline in the probability of outmigration for non-beneficiary female household members, the decline

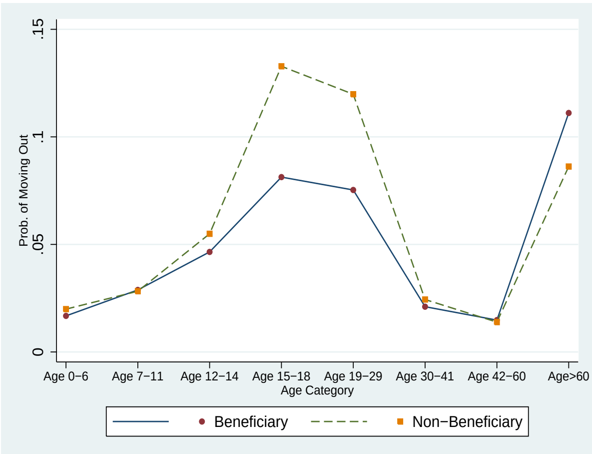
Figure 2: Probability of Moving Out of the Household for Different Age Groups



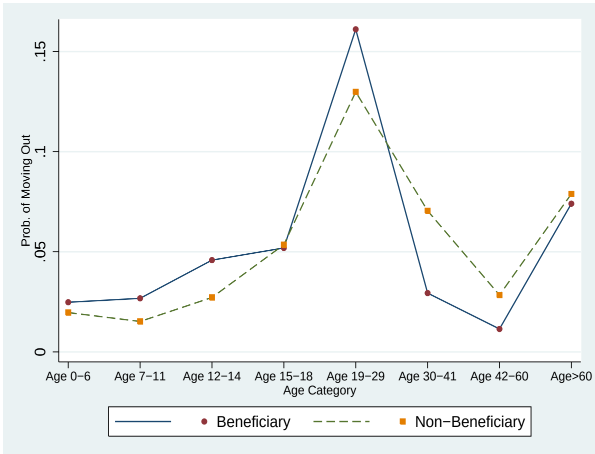
(a) By Gender



(b) By Participation Status



(c) Females: By Participation Status



(d) Males: By Participation Status

for beneficiary household members is relatively flatter. In other age intervals, however, the probability of observing females moving out from the household does not appear to differ that much between beneficiary and non-beneficiary household members. In the case of male household members, on the other hand, we do not see any clear differences (see Panel (d)).

Marriage and labor migration are more likely to be observed, plausibly, in the 15-30 age interval. This, coupled with the observation that marriage and work are the two major reasons for leaving the household, observing an out-migrant in this age interval can be attributed mainly, if not fully, to these two reasons.[23]

Overall, even if the above is based on post treatment period comparison between beneficiary and non-beneficiary households, without taking into account the before and after changes, there is clear indication that, in age intervals where we expect to see

[23]Given that early child marriage is still a problem in rural parts of Ethiopia, observing females getting married in their mid-teens is not uncommon.

relatively more out-going members, females in beneficiary households are less likely to move out from the household.[24] Given that marriage is the main reason for females to move out of the household, the above pattern gives an indication about the (early) marriage delaying effect of the program.

Regression based analysis

We have tested the above predictions empirically by estimating the model specified in section 3, Equation 3.0.1 using a binary dependent variable which equals one if the household has at least one marriage related female out-migrant member, and zero otherwise. This is done separately for the 14-18, 15-18, 16-18 and 19-29 age groups and the result from this exercise is presented in Table 4. We chose to focus on these age groups since, given the context we are looking at, these are the age categories where female marriage (including early marriage) is likely to happen. Since the regressions are done conditional on a household having female members in each specific age group, the sample size used in the estimations is smaller compared to what we used in the previous regressions. Given the binary nature of the dependent variable, we have done the estimations using linear probability model (LPM), with and without household fixed effects, and Probit model.

As can be seen from Table 4, for females in the 14-18, 15-18, 16-18 age groups, an increase in per capita public works income transfer, conditional on public works labor supply and other covariates in the model, is associated with a decrease in the probability of observing a household with a female out-going member due to marriage. This can be seen from the negative and statistically significant coefficient on the quadratic term. Although the coefficient on the linear term that captures the level of treatment (*Payment* $\times$ *Post*) is positive, but statistically insignificant, in the fixed effects model, the stronger coefficients on the quadratic term can be taken as an indication that the income effect is weak for low levels of transfer payments. Overall, this implies that income gains from public works participation has marriage delaying effect for female adolescent household members.

On the other hand, the total income effect on female marriage is positive and statistically significant for females in the 19-29 age group. The positive relationship between income gains and the probability of observing a household with at least one female member in 19-29 age category that moved out due to marriage is an indication that the female marriage delaying decision of beneficiary households is done in a rather rational way. That is, beneficiary households are willing to let their female members get married when they reach an acceptable age for marriage. Given that age of the female at first marriage is one of the proximate determinants of fertility, the marriage delaying effect observed above will also have its own implication for the effect of the program on fertility.

In view of the evidence in the literature that marriage can also be used as a means to

[24]Of course, this is also without controlling for any covariates and without taking the levels of treatment into account.

smooth consumption and mitigate risk in the face of income short falls (see Rosenzweig and Stark (1989)), the marriage delaying effect of the public work income transfers found in this paper can be indicative of the improvement in beneficiary households' financial capacity to support their members.

On the other hand, conditional on the income effect, we do not find evidence to suggest that beneficiary households are delaying the marriage of their adolescent female members due to the labor supply effect of participation in public works activities. In almost all cases, the parameters of interest for adolescent females are not precisely estimated.

In Table 5 we have presented results showing income and labor supply effects on work related out migration of male members; where the dependent variable is a binary variable indicating whether the household has at least one male member that moved out for work related reasons. Again, this is estimated conditional on the household having a male member in the relevant age groups. Looking at the results, there is only weak evidence that the income effect of participation increases out-migration of male members in the 19-29 age group at decreasing rate. However, this evidence is weak and it is not robust across specifications. Moreover, we do not find evidence to suggest that an increase in public work labor months is associated with a decline in out-migration of males, which is what one would expect to observe if the labor supply effect induces retention of male labor.

Table 4: Probability of Observing a Household with an Outgoing Female Member Due to Marriage: By Age

	OLS(LPM)				Probit				With HH FE			
	14-18	15-18	16-18	19-29	14-18	15-18	16-18	19-29	14-18	15-18	16-18	19-29
Payment× Post	0.023	-0.040	-0.024	-0.033	-0.185**	-0.287***	-0.130*	-0.042	0.119	0.030	0.050	-0.053
	(0.119)	(0.132)	(0.171)	(0.116)	(0.072)	(0.084)	(0.072)	(0.050)	(0.137)	(0.152)	(0.183)	(0.118)
Sq. Payment× Post	-0.343**	-0.416***	-0.462***	0.435***	-0.753**	-1.202***	-0.543	0.098*	-0.401**	-0.417***	-0.464***	0.510***
	(0.169)	(0.117)	(0.135)	(0.151)	(0.370)	(0.391)	(0.344)	(0.055)	(0.168)	(0.138)	(0.158)	(0.127)
PW Days× Post	-0.260	-0.222	-0.079	0.326	-0.048	-0.078	-0.009	0.221**	-0.527*	-0.486	-0.335	0.364
	(0.250)	(0.289)	(0.381)	(0.216)	(0.059)	(0.083)	(0.026)	(0.107)	(0.287)	(0.331)	(0.409)	(0.249)
Sq. PW Days× Post	0.261	0.259	-0.141	-0.506	0.964	1.485*	0.618	-0.068	0.227	0.091	-0.381	-0.444
	(0.428)	(0.509)	(0.611)	(0.526)	(0.718)	(0.795)	(0.509)	(0.192)	(0.537)	(0.666)	(0.790)	(0.577)
Post Period (d)	0.019	0.015	0.019	0.016	-0.044	-0.060	-0.038	0.015	-0.020	-0.024	0.006	0.009
	(0.040)	(0.046)	(0.060)	(0.034)	(0.037)	(0.050)	(0.038)	(0.020)	(0.049)	(0.059)	(0.069)	(0.040)
Number of Obs.	568	503	423	688	568	503	423	688	568	503	423	688
$adjR^2$	.031	.038	.035	.077					.083	.1	.095	.091
Prob > F	0.00	0.00	0.00	0.00	.	.	.	.	0.13	0.00	0.00	0.00
Household FE	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	-	-	-	-
Mean of the Dep. Var	0.06	0.06	0.06	0.05	0.06	0.06	0.06	0.05	0.06	0.06	0.06	0.05

Standard errors clustered at Woreda Level in Parenthesis
(d) for discrete change of dummy variable from 0 to 1
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Probability of Observing a Household with an Outgoing Male Member Due to Work: By Age

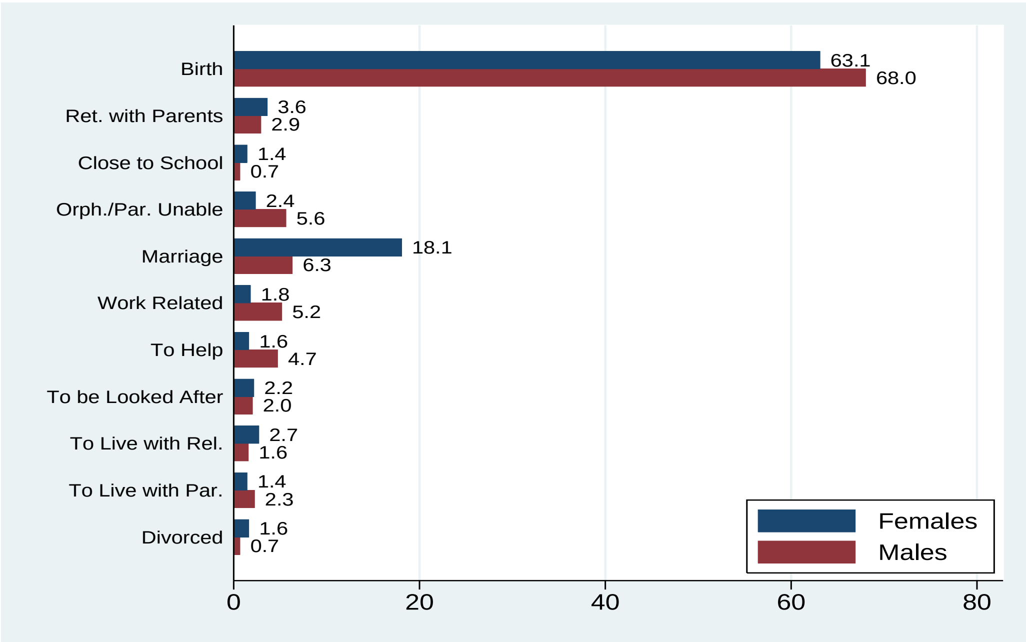
	OLS(LPM)				Probit				With HH FE			
	14-18	15-18	16-18	19-29	14-18	15-18	16-18	19-29	14-18	15-18	16-18	19-29
Payment× Post	0.012	0.044	0.052	0.259**	0.046*	0.048*	0.035*	0.249***	-0.001	0.045	0.032	0.212
	(0.050)	(0.039)	(0.047)	(0.117)	(0.024)	(0.025)	(0.020)	(0.077)	(0.053)	(0.034)	(0.037)	(0.132)
Sq. Payment× Post	0.002	-0.029	-0.035	-0.108	0.240*	0.201	0.166	0.567	0.008	-0.037	-0.029	-0.100
	(0.046)	(0.036)	(0.036)	(0.100)	(0.133)	(0.124)	(0.107)	(0.442)	(0.049)	(0.032)	(0.027)	(0.095)
PW Days× Post	0.004	0.078	0.052	-0.295*	0.002	0.018	0.007	-0.150	0.076	0.132	0.132	-0.304
	(0.160)	(0.186)	(0.233)	(0.176)	(0.048)	(0.037)	(0.029)	(0.146)	(0.176)	(0.200)	(0.251)	(0.225)
Sq. PW Days× Post	0.101	-0.017	0.262	0.367	0.089	0.073	0.096	1.537	-0.096	-0.248	-0.148	-0.412
	(0.419)	(0.448)	(0.668)	(0.451)	(0.145)	(0.096)	(0.087)	(1.509)	(0.472)	(0.502)	(0.721)	(0.451)
Post Period (d)	-0.018	0.002	-0.007	0.042	-0.017	-0.009	-0.010	-0.028	-0.031	-0.013	-0.020	0.049
	(0.039)	(0.040)	(0.050)	(0.036)	(0.016)	(0.012)	(0.012)	(0.051)	(0.050)	(0.051)	(0.061)	(0.042)
Number of Obs.	623	567	492	709	623	567	492	709	623	567	492	709
$adjR^2$	.017	.0088	.017	.026					.009	.019	.023	.059
Prob > F	0.00	0.01	0.00	0.00	.	.	.	.	0.28	0.14	0.18	0.03
Household FE	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	-	-	-	-
Mean of the Dep. Var	0.03	0.03	0.03	0.07	0.03	0.03	0.03	0.07	0.03	0.03	0.03	0.07

Standard errors clustered at Woreda Level in Parenthesis
(d) for discrete change of dummy variable from 0 to 1
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4.4 Effect of the Program on Probability of Adding a New Member

In understanding how the income and labor supply effects of public works participation affect household size, it is important to examine if the probability of attracting new/incoming members increases with the level of treatment. Broadly speaking, the addition of a new member to a household can reflect two things: first, assuming that the flow of benefit is mostly from the host to the new member, it shows the household's willingness and capacity to support more members; second, it can also show a mutual benefit where the incoming member also contributes resources to the household.[25] To better understand the main motive behind adding members it is helpful to look at the reasons for a new member joining the household. In Figure 3 we have summarized the major reasons for observing a member moving-in to the household by gender.

Figure 3: Reasons for Moving-In to the Household: By Gender



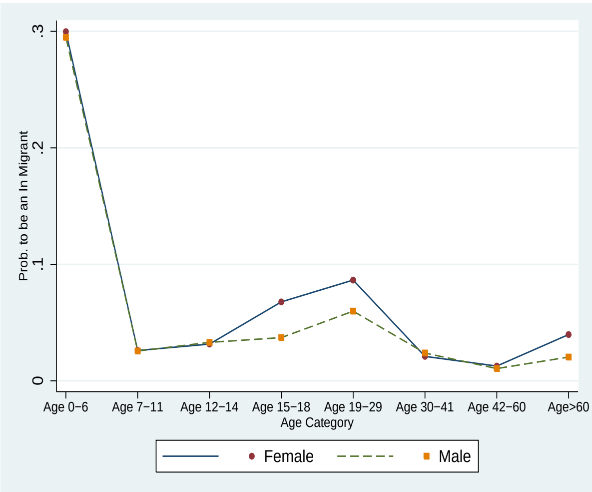
As can be observed from Figure 3, in our data, of the total households who reported to have an incoming member, more than 60 percent indicated birth as the reason for adding a new member to the household. Moreover, marriage is found to be the second major explanation for observing a member joining the household. Here it should also be noticed that there is a clear gender difference. In particular, while 18 percent of households indicated marriage as a reason for adding a new female member, the proportion of households that indicated marriage as a reason for adding male members is only 6 percent.

Given that new births and marriage appear to be the major reasons/channels for having a new member to the household, in what follows we will focus on these two reasons. Before turning to a regression analysis to estimate the effects of public works

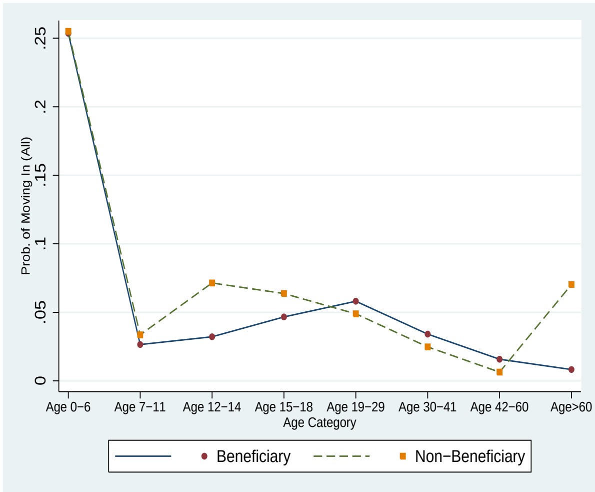
[25] This can be either by directly participating in activities that can generate income to the household or by giving a helping hand in daily household chores.

income and labor supply on the probability of observing new members to the household through birth and marriage, we first do a graphical inspection of the probabilities of being a new member to the household for the different age groups. Accordingly, using the member level data, we have plotted the probabilities of being an incoming/new member for members in the different age categories in Figure 4. This is done: by gender (Panel A), by participation status (Panel B), for females by participation status (Panel C) and for males by participation status (Panel D).

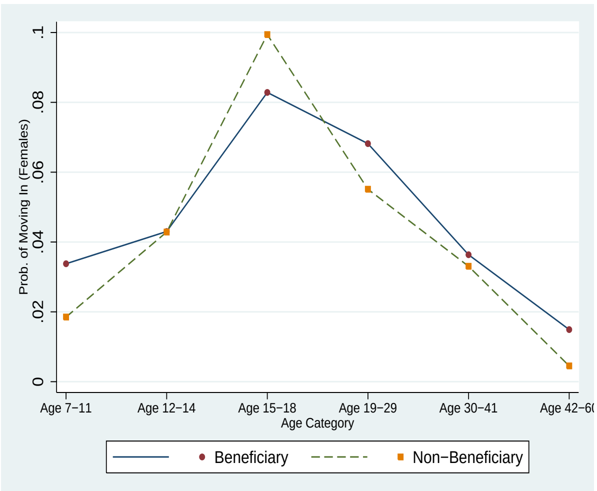
Figure 4: Probability of Moving-In to the Household for Different Age Groups



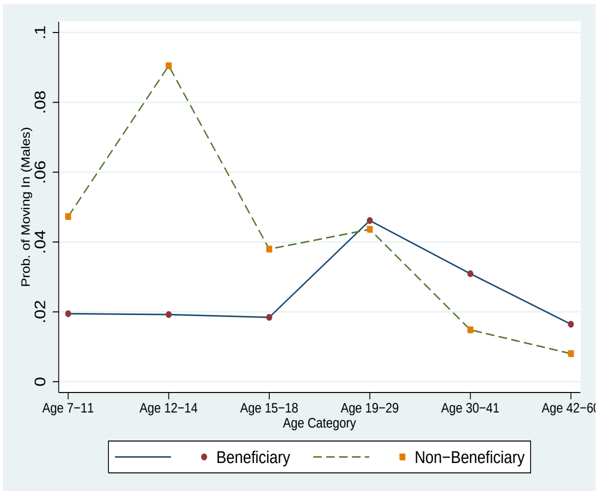
(a) By Gender



(b) By Participation Status



(c) Females: By Participation Status



(d) Males: By Participation Status

As it is evident from Figure 4 the probability of being a new member in the age interval 0-6 is the highest and this probability appears to be the same for males and females (as one would expect). Specifically, 30 percent of both male and female members in the 0-6 age group are new members to the households. Coupled with the observation from Figure 3 that birth is the major explanation for having a new member, the high probability of being a new member observed for those in the 0-6 age category can be explained by birth. Given the long standing concern in the literature that income transfers can lead to an increase in the demand for children and hence fertility, one may expect the probability of having a new member by birth to be higher for beneficiaries compared to non-beneficiaries. However, as the graph in Panel B of Figure 4 makes it clear, we do not see a difference in probability of having a new member in 0-6 age interval (by Birth) between beneficiary and non-beneficiary households.

Looking at the probabilities of being a new member for the other age groups in Panel B of Figure 4, we see non-beneficiaries to have slightly higher probability of having new members in the 15-18 age interval. For members in age intervals 19-29 and onwards, the difference in probability of being a new member is not that different between the two groups. The gender disaggregation in Panel (C) and (D) also tells a similar story. For females, the overall pattern looks similar for beneficiaries and non-beneficiaries except that we observe slight differences in age intervals where (early) marriage is likely to occur (i.e., in the 15-18 age interval). In particular, the probability of having an incoming member in the 15-18 age interval is slightly higher for non-beneficiaries, while the opposite is true for members in the 19-29 age interval. For Males in Panel (D), though the probability of observing an incoming member in age groups 7-11 and 12-14 is way smaller for beneficiaries compared to non-beneficiaries, beneficiary households appear to have the highest probability of having an incoming male member in 19-29 age interval compared to other age categories. [26]

From the above graphical analysis, it is apparent that child birth is the main explanation for having a new member in a household, though we do not observe a difference in the probability of having a new member through birth between beneficiaries and non-beneficiaries. Although this in itself is interesting, given the fear of an increase in fertility following financial gains from cash transfer programs, it needs to be substantiated empirically. The fact that the Ethiopian PSNP has been there for long makes it suitable to evaluate the fertility impact of the program.

The income effect from public works cash transfers and the labor supply requirement attached to it are likely to affect fertility in different ways. On the one hand, the increase in income may increase the demand for children as the public work cash transfer brings additional resources to the household and lessens the burden of raising a child. However, if parents value quality over quantity, they may prefer to use the public works income to give a better life to their existing children, in which case we will not expect the additional income to have an impact on the demand for more children. Thus, the income effect of public works program can depend on the extent to which the additional income relaxes households' financial constraints and their preference for quality vs quantity of children.

When it comes to the effect of the labor supply requirement on fertility/demand for children, the time constraint beneficiary households are facing is what counts most. Specifically, to the extent that working on public works competes with the available time beneficiary households can have for raising children, the labor supply requirement of

[26]The spike in the probability of being a new member in 12-14 age group observed for non-beneficiary households is puzzling and remains unexplained.

the program can reduce the demand for children. However, whether the time constraint is going to be binding or not also depends on the options beneficiary households have in terms of child care. For instance, if beneficiaries can give their newborn to their relatives or if they can have someone in the household to look after the newborn, the labor supply requirement may not necessarily lead to a lower demand for children. Overall, the impact of the public works labor supply requirement on the demand for children/fertility by and large depends on whether the time constraint beneficiary households face is binding or not.

We start by estimating the probability of observing a household with 0-6 incoming members (Panel A of Table 6) and with incoming member due to birth (Panel B of Table 6). As can be seen from Panel A, we do not find any statistically significant evidence to suggest that the income effect of the program is associated with an increase in the probability of observing households with an incoming member in the 0-6 age group. The same is true for the labor supply effect. These results can be taken as suggestive evidence for the lack evidence regarding the income and labor supply effects on households' fertility decisions.

Results reported in Panel B of Table 6 directly estimate for the income and labor supply effects on households' fertility decisions. In this case we observe some weak evidence for a positive (after some level) income effect on fertility. In particular, conditional on public works labor supply, after some point, an increase in income from public works cash transfers is found to be positively associated with the probability of observing a household with an incoming member due to birth. However, this is evidence is rather weak and it is not precisely estimated in the fixed effects model. We do not also find any evidence to suggest an important labor supply effect on households' fertility decisions.

Table 6: Probability of Observing a Household with an Incoming Member in 0-6 Age Group

Dep. Var: In-Migrants in the 0-6 Age Group			
	Using OLS	Using Probit	With HH FE
Payment× Post	-0.218	-0.280	-0.040
	(0.236)	(0.266)	(0.242)
Sq. Payment× Post	0.270	0.418	0.224
	(0.220)	(0.273)	(0.187)
PW Days× Post	-0.042	-0.028	-0.250
	(0.427)	(0.467)	(0.485)
Sq. PW Days× Post	0.765	0.816	0.465
	(1.028)	(1.122)	(0.888)
Post Period (d)	-0.164*	-0.197**	-0.130
	(0.086)	(0.100)	(0.104)
Number of Obs.	1137	1137	1137
$adjR^2$	.18		.08
Prob > F	0.00		0.00
Household FE	No	No	Yes
Region FE	Yes	Yes	-
Mean of the Dep. Var	0.39	0.39	0.39

Dep. Var: Move-In Due to Birth=1			
	Using OLS	Using Probit	With HH FE
Payment× Post	-0.242	-0.249	-0.007
	(0.234)	(0.270)	(0.244)
Sq. Payment× Post	0.371**	0.405*	0.241
	(0.176)	(0.207)	(0.177)
PW Days× Post	0.045	0.024	-0.255
	(0.412)	(0.471)	(0.478)
Sq. PW Days× Post	0.065	0.201	-0.148
	(1.008)	(1.213)	(1.080)
Post Period (d)	-0.119	-0.140	-0.099
	(0.077)	(0.090)	(0.091)
Number of Obs.	1071	1071	1071
$adjR^2$	.18		.07
Prob > F	0.00		0.00
Household FE	No	No	Yes
Region FE	Yes	Yes	-
Mean of the Dep. Var	0.39	0.39	0.39

Standard errors clustered at Woreda Level in Parenthesis
(d) for discrete change of dummy variable from 0 to 1
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Additional results are presented in Table 7, where we have the total number of

Table 7: Dependent Variable: Number of Members in 0-6 Age Group

No. of Household Members: 0-6 Age Group			
	Using OLS	With HH FE	FE Poisson
Payment× Post	0.355	0.354	0.142
	(0.365)	(0.340)	(0.257)
Sq. Payment× Post	0.116	0.123	0.324
	(0.285)	(0.267)	(0.337)
PW Days× Post	-0.940	-0.740	-0.347
	(0.665)	(0.681)	(0.499)
Sq. PW Days× Post	-0.694	-0.069	0.008
	(2.217)	(2.160)	(1.441)
Post Period	-0.187	-0.291*	-0.207*
	(0.151)	(0.160)	(0.117)
Constant	2.118***	2.005***	
	(0.254)	(0.353)	
Number of Obs.	814	814	814
$adjR^2$	.13	.08	
Prob > F	0.00	0.00	.
Household FE	No	Yes	Yes
Region FE	Yes	-	-
Mean of the Dep. Var	1.43	1.43	1.43

Standard errors clustered at Woreda Level in Parenthesis
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

members in 0-6 age interval as an outcome variable. Since this variable measures the net change (the combined effect of incoming and outgoing members) in the number of 0-6 members, the results can be used to infer whether or not these groups have contributed to the increase in aggregate household size. As can be seen from Table 7 there is no evidence to suggest that the income effect of the public works cash transfer and the labor supply requirement attached to it induce changes in the number of members in the 0-6 age category.

One thing that is worth noting in Table 7 is that, in the case of public work labor months, the coefficients of both the linear and the square terms are negative implying that an increase in public work labor months is negatively associated with the number of members in the 0-6 age category. Though not precisely estimated, this results can be taken as a suggestive evidence for the fact that the time constraint beneficiary parents are facing for child care activities is binding, forcing them to give their 0-6 members to their relatives. If this is the case, we should see an increase in outgoing members in the 0-6 age group following an increase in public work labor months.

We have tested this by estimating the labor supply effect on the probability of observing a household with an out migrant member in the 0-6 age group. Accordingly results reported in Table 8 give some suggestive evidence on the above claim. In

Table 8: Probability of Observing a Household with an Outgoing Member in 0-6 Age Group

Dep. Var: Out Mig of 0-6	Using OLS	Using Probit	Fixed Effect
main			
Payment× Post	-0.007	0.003	-0.074
	(0.111)	(0.088)	(0.112)
Sq. Payment× Post	0.013	0.354	0.056
	(0.080)	(0.328)	(0.079)
PW Days× Post	0.515**	0.637***	0.571**
	(0.205)	(0.181)	(0.242)
Sq. PW Days× Post	-0.490	-0.616	-0.487
	(0.545)	(0.405)	(0.738)
Post Period (d)	0.083**	0.095**	0.078
	(0.041)	(0.043)	(0.052)
Number of Obs.	909	909	909
$adjR^2$	.0027		.02
Prob > F	0.00		0.47
Household FE	No	No	Yes
Region FE	Yes	Yes	No
Mean of the Dep. Var	0.06	0.06	0.06

Standard errors clustered at Woreda Level in Parenthesis
(d) for discrete change of dummy variable from 0 to 1
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

particular, conditional on income, an increase in public works labor supply is associated with an increase in probability of observing a household with an outgoing member in the 0-6 age group.

Taken together, based on the above evidence we can say that, the increase in the number of 0-6 members does not explain the observed increase in aggregate household size.

Finally, apart from incoming members in 0-6 age group, beneficiary households might be more likely to attract adult/working age members. In order to check this we have focused on female members joining the household due to marriage since marriage is the dominant reason for females to join a household. For males on the other hand, we do it regardless of reason as we don't have enough observations to focus on a single reason and there is no as such one dominant reason for observing male incoming members. In both cases the estimations are done focusing on relevant age category and the results are presented in Table 17 and Table 18.

As can be seen from Table 17, there is some evidence to suggest that an increase in public work labor supply is associated with an increase in the probability of observing households with an in-migrant male member in the 19-29 age category. From the positive and statistically significant coefficient on the quadratic term on (*Sq PW Days* × *Post*), we can say that the positive effect of labor supply starts only after a certain level of labor supply. Moving to female incoming members, as can be seen from results presented in Table 18, we do not find enough evidence to suggest either an income or a labor supply effect on the probability of observing female incoming members due to marriage.

5 Conclusion

As part of the effort to fight poverty and promote livelihood of the rural poor in a sustainable way, cash transfer programs, conditional or otherwise, are becoming common in low income countries. The Ethiopian public works program where beneficiary households receive cash transfers conditional on participating in public work activities is one example. This paper empirically assess how financial gains from public works cash transfers and the labor supply conditions tied to it impact household structure. This is based on the argument that, while income gains from the program can relax households' liquidity/financial constraint, the labor supply condition may introduce resource constraint, in terms of time/labor constraint. Impact of participation in the public works program on household size and composition will thus depend not only on the degree to which the cash transfer relaxes households' financial constraint (income effect) but also on the extent to which the time constraint induced by the labor supply requirement is binding (labor supply effect).

Accordingly, the findings of this paper show that both the income and labor supply effects are positively associated with household size. We also found evidence for differential effects based on attributes of household members, particularly gender and age. Specifically, the increase in household size induced by the income effect is found to be associated with an increase in the number of female adolescent household members. Moreover, we find this to be mainly due to the effect of public works income in increasing households' tendency of retaining female adolescent members for longer period through delaying (early) marriage. We also found evidence that the income effect increases marriage related outmigration of female members in the 19-29 age category; implying that the marriage delaying effect of income goes only until a certain age.

On the other hand, an increase in public work labor supply is found to increase household size mainly by increasing the number of male household members. Consistent with this result, we have found the probability of observing a household with a male incoming member to increase with an increase in public works labor supply; and this is particularly true in the case of males in the 19-29 age category. Finally, while we do not find enough evidence to support the claim that income transfer increases number of children/fertility, there is evidence that an increase public works labor supply to be positively associated with the probability of observing a household with an outgoing member in the 0-6 age category.

Taken together, the result that households are retaining their female adolescent members for long is suggestive of the fact the cash transfers may have contributed towards relaxing the financial constraints of some households giving them the capacity to support their female members for longer period. On the other hand, the increase in male incoming members and outgoing children might be an indication that the time constraint following public works labor supply might be binding for some households forcing them to attract new members and/or send out their small children.

The above results have important implications for the design and implementation of poverty focused cash transfer programs in low income countries. In particular, the findings accentuate the need for program designers and policy makers to fully understand how household structure responds to not only the benefits but also the conditions attached to the benefits. Failure to do so can restrict the effectivenesss of cash transfer programs and limit their scope as a tool for fighting poverty and vulnerability in low income countries.

Appendices

Appendix .1: Descriptive Statistics

Table 9: Household Size and Composition

Variable	Mean	Std. Dev.	Min.	Max.	N
Household size	5.22	2.31	1	15	1259
Male	2.69	1.59	0	10	1259
Female	2.53	1.36	0	9	1259
No. of HH Members(0-6)	1.1	1.02	0	4	1259
No. of HH Members(7-11)	0.81	0.82	0	4	1259
No. of Males(12-18)	0.57	0.75	0	4	1259
No. of Females(12-18)	0.45	0.68	0	4	1259
No. of Males(19-29)	0.41	0.6	0	4	1259
No. of Females(19-29)	0.4	0.51	0	3	1259
No. of Males(30-41)	0.27	0.42	0	2	1259
No. of Females(30-41)	0.35	0.45	0	2	1259
No. of Males(42-60)	0.34	0.46	0	2	1259
No. of Females(42-60)	0.3	0.43	0	1.33	1259
No. of HH Members(> 61)	0.24	0.48	0	2	1259

Table 10: Summary Statistics on Out-Migration of Members by Age, Gender and Reason

Variable	Mean	Std. Dev.	Min.	Max.	N
Dummy for Out-Migrants	0.33	0.47	0	1	1259
Dummy for Male Out-Migrants	0.22	0.41	0	1	1203
Dummy for Female Out-Migrants(12-18)	0.17	0.38	0	1	1243
Dummy for Out Mig (0-6)	0.05	0.21	0	1	823
Dummy for Out Mig (7-11)	0.03	0.17	0	1	766
Dummy for Female out-Migrants(12-18)	0.1	0.3	0	1	507
Dummy for Male out-Migrants(12-18)	0.06	0.23	0	1	581
Dummy for Female out-Migrants(19-29)	0.08	0.27	0	1	537
Dummy for Male out-Migrants(19-29)	0.12	0.32	0	1	498
Dummy for Female Out-Migrants(30-41)	0.02	0.12	0	1	509
Dummy for Male Out-Migrants(30-41)	0.03	0.16	0	1	385
Dummy for Female Out-Migrants(42-60)	0.01	0.11	0	1	438
Dummy for Male Out-Migrants(42-60)	0.02	0.12	0	1	466
Dummy for Child Death	0.04	0.19	0	1	1334
Dummy for Marriage	0.08	0.27	0	1	1334
Dummy for To be Close to School	0.03	0.16	0	1	1334
Dummy for Work Related	0.07	0.26	0	1	1334
Dummy for Adult Death	0.04	0.21	0	1	1334
Dummy for Divorced	0.02	0.14	0	1	1334
Dummy for Other Reasons	0.03	0.18	0	1	1334

Table 11: Summary Statistics on In-Migration of Members by Age, Gender and Reason

Variable	Mean	Std. Dev.	Min.	Max.	N
Dummy for Incoming Members	0.46	0.5	0	1	1334
Dummy for Male In-Migrants	0.25	0.44	0	1	1334
Dummy for Female In-Migrants	0.31	0.46	0	1	1334
Dummy for In-Migrants(0-6)	0.36	0.48	0	1	1334
Dummy for In-Migrants(7-11)	0.03	0.17	0	1	1334
Dummy for Female In-Migrants(12-18)	0.04	0.2	0	1	1334
Dummy for Male In-Migrants(12-18)	0.03	0.18	0	1	1334
Dummy for Female In-Migrants(19-29)	0.06	0.23	0	1	1334
Dummy for Male In-Migrants(19-29)	0.04	0.2	0	1	1334
Dummy for Female In-Migrants(30-41)	0.01	0.11	0	1	1334
Dummy for Male In-Migrants(30-41)	0.01	0.11	0	1	1334
Dummy for Female In-Migrants(42-60)	0.01	0.07	0	1	1334
Dummy for Male In-Migrants(42-60)	0	0.05	0	1	1334
Dummy for Birth	0.34	0.47	0	1	1334
Dummy for Marriage	0.07	0.26	0	1	1334
Dummy for To be Close to School	0.01	0.09	0	1	1334
Dummy for To be Close to Work	0.02	0.14	0	1	1334
Dummy for To live with Parent/Relat.	0.04	0.2	0	1	1334
Dummy for Other Reasons	0.05	0.22	0	1	1334
Dummy for Divorced	0.01	0.09	0	1	1334

Table 12: Summary Statistics on Levels of Treatment Related Variables

Variable	Mean	Std. Dev.	Min.	Max.	N
Lagged Payment ('000 birr)	1.34	1.44	0	14.52	272
Lagged Per Cap. Pay ('000 birr)	0.28	0.28	0	2.56	272
Lagged Total PW Labor Months	3.71	3.05	0	20.23	271
Lagged Per. Cap PW Labor Months	0.15	0.14	0	0.84	266
Duration of Participation	3.04	1.56	1	6	272
No. Years Before and After	2.75	1.66	1	6	1334

Table 13: Pre-Program Mean Comparison of Covariates across Beneficiary and Comparison Households

	Comparison	Obs	Beneficiary	Obs	Difference	p-val
Male Headed Household	0.82	371	0.79	248	0.04	0.240
Head Age	46.02	371	44.11	248	1.91	0.128
No Formal Education	0.81	371	0.79	248	0.02	0.524
Primary Education	0.19	371	0.21	248	-0.02	0.548
Secondary Education	0.04	371	0.04	248	0.00	0.995
Per Capita Land_Res (Ha)	0.05	332	0.06	227	-0.00	0.352
Own Land (In Ha.)	1.27	359	1.30	234	-0.03	0.703
Value of Productive Assets	4.96	361	4.97	234	-0.01	0.947
Corrugated Metal Roof	0.20	371	0.13	248	0.07^{**}	0.022
Distress Asset Sell	0.54	371	0.59	248	-0.05	0.235
Shocks to Crop Production	0.69	359	0.69	245	0.00	0.921
Poblems in Crop Production	0.36	365	0.39	248	-0.02	0.570
Shock to Livestocks	0.69	366	0.73	245	-0.04	0.326
Poor Relative to Others	0.53	370	0.59	248	-0.06	0.167
Average Relative to Others	0.42	370	0.47	248	-0.05	0.196

Appendix .2: Additional Results

Table 14: Probability of Observing a Household with an Outgoing Female Member: By Age

	OLS(LPM)				Probit				With HH FE			
	12-18	14-18	16-18	19-29	12-18	14-18	16-18	19-29	12-18	14-18	16-18	19-29
main												
Payment× Post	-0.012	-0.214	-0.222	-0.026	0.029	-0.214	-0.064	-0.040	0.089	-0.049	-0.089	-0.138
	(0.222)	(0.233)	(0.244)	(0.189)	(0.207)	(0.242)	(0.058)	(0.193)	(0.227)	(0.222)	(0.227)	(0.166)
Sq. Payment× Post	-0.263	-0.137	-0.217	0.011	-0.210	-0.202	-0.036	-0.002	-0.389*	-0.322	-0.340*	0.144
	(0.221)	(0.230)	(0.185)	(0.160)	(0.184)	(0.476)	(0.126)	(0.175)	(0.206)	(0.208)	(0.175)	(0.129)
PW Days× Post	-0.325	0.029	-0.089	0.416	-0.356	-0.003	0.002	0.341	-0.526	-0.184	-0.304	0.487*
	(0.436)	(0.481)	(0.652)	(0.327)	(0.393)	(0.465)	(0.126)	(0.330)	(0.420)	(0.421)	(0.553)	(0.283)
Sq. PW Days× Post	1.052	0.580	0.973	-0.464	2.838	3.332	12.864***	1.411	1.088	0.655	1.299	-0.442
	(0.885)	(0.862)	(1.117)	(0.635)	(2.919)	(4.773)	(2.895)	(1.357)	(0.897)	(0.786)	(0.910)	(0.597)
Post Period (d)	0.025	0.038	-0.007	0.063	-0.027	-0.046	-0.998***	-0.005	-0.012	0.005	-0.036	0.048
	(0.062)	(0.067)	(0.100)	(0.050)	(0.108)	(0.171)	(0.002)	(0.060)	(0.081)	(0.079)	(0.103)	(0.064)
Number of Obs.	673	567	423	687	673	567	423	687	673	567	423	687
$adjR^2$	.048	.054	.044	.069					.035	.043	.053	.049
Prob > F	0.00	0.00	0.00	0.00					0.04	0.03	0.00	0.02
Household FE	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	-	-	-	-

Standard errors clustered at Woreda Level in Parenthesis
(d) for discrete change of dummy variable from 0 to 1
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 15: Probability of Observing a Household with an Outgoing Male Member: By Age

	OLS(LPM)				Probit				With HH FE			
	12-18	14-18	16-18	19-29	12-18	14-18	16-18	19-29	12-18	14-18	16-18	19-29
main												
Payment× Post	-0.204	-0.162	-0.123	0.162	-0.197*	-0.140	-0.110	0.218	-0.214	-0.185	-0.150	0.045
	(0.133)	(0.139)	(0.144)	(0.190)	(0.108)	(0.108)	(0.095)	(0.198)	(0.131)	(0.140)	(0.138)	(0.199)
Sq. Payment× Post	0.137	0.092	0.053	0.127	-0.369	-0.497	-0.767*	0.007	0.127	0.116	0.069	0.225
	(0.094)	(0.092)	(0.095)	(0.171)	(0.467)	(0.438)	(0.444)	(0.181)	(0.089)	(0.095)	(0.094)	(0.165)
PW Days× Post	0.363	0.203	0.207	-0.326	0.198	0.054	0.000	-0.375	0.546*	0.400	0.444	-0.222
	(0.288)	(0.299)	(0.381)	(0.419)	(0.193)	(0.170)	(0.181)	(0.468)	(0.302)	(0.322)	(0.387)	(0.433)
Sq. PW Days× Post	-0.663	-0.358	-0.473	-1.300	-0.308	-0.023	0.208	-1.973	-1.348*	-1.006	-1.723	-1.728**
	(0.728)	(0.757)	(1.143)	(1.102)	(0.692)	(0.668)	(0.589)	(1.636)	(0.799)	(0.911)	(1.277)	(0.858)
Post Period (d)	-0.009	-0.012	-0.003	0.060	-0.012	-0.009	0.000	0.090	0.015	-0.020	0.021	0.064
	(0.048)	(0.045)	(0.066)	(0.063)	(0.034)	(0.033)	(0.036)	(0.073)	(0.056)	(0.054)	(0.072)	(0.071)
Number of Obs.	714	621	491	709	714	621	491	709	714	621	491	709
$adjR^2$	.041	.045	.024	.051					.026	.046	.033	.041
Prob > F	0.00	0.00	0.00	0.00					0.63	0.62	0.55	0.00
Household FE	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	-	-	-	-

Standard errors clustered at Woreda Level in Parenthesis
(d) for discrete change of dummy variable from 0 to 1
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 16: Probability of Observing a Household with an Incoming Female Member: By Age

	OLS(LPM)				Probit				With HH FE			
	12-18	14-18	16-18	19-29	12-18	14-18	16-18	19-29	12-18	14-18	16-18	19-29
main												
Payment× Post	0.041	-0.023	-0.051	0.069	0.018	-0.040	-0.071	0.071	0.071	-0.015	-0.038	0.024
	(0.100)	(0.060)	(0.061)	(0.107)	(0.091)	(0.060)	(0.051)	(0.082)	(0.121)	(0.069)	(0.061)	(0.108)
Sq. Payment× Post	0.003	0.009	0.030	-0.004	0.018	0.003	0.036	0.446**	-0.011	0.008	0.029	0.006
	(0.046)	(0.032)	(0.030)	(0.056)	(0.046)	(0.044)	(0.026)	(0.205)	(0.054)	(0.039)	(0.032)	(0.059)
PW Days× Post	-0.061	0.038	0.026	-0.282	-0.040	0.067	0.049	-0.093	-0.066	0.035	0.021	-0.241
	(0.161)	(0.104)	(0.087)	(0.186)	(0.139)	(0.107)	(0.082)	(0.109)	(0.200)	(0.121)	(0.084)	(0.209)
Sq. PW Days× Post	0.143	0.007	0.100	0.741	0.865	0.379	0.443	0.727	0.037	0.248	0.509**	1.046**
	(0.316)	(0.203)	(0.201)	(0.445)	(0.664)	(0.484)	(0.461)	(0.458)	(0.475)	(0.307)	(0.230)	(0.505)
Post Period (d)	0.004	0.003	-0.007	-0.045	-0.020	-0.007	-0.019	-0.045	0.020	0.010	-0.001	-0.072
	(0.032)	(0.029)	(0.024)	(0.042)	(0.033)	(0.030)	(0.025)	(0.028)	(0.045)	(0.039)	(0.030)	(0.044)
Number of Obs.	1137	1137	1137	1137	1137	1137	1137	1137	1137	1137	1137	1137
$adjR^2$	.0044	.0036	-.0042	.012					.016	.015	.016	.012
Prob > F	0.00	0.00	0.06	0.00					0.02	0.19	0.18	0.56
Household FE	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	-	-	-	-
Mean of the Dep. Var	0.04	0.04	0.03	0.05	0.04	0.04	0.03	0.05	0.04	0.04	0.03	0.05

Standard errors clustered at Woreda Level in Parenthesis
(d) for discrete change of dummy variable from 0 to 1
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 17: Probability of Observing a Household with an Incoming Male Member: By Age

	OLS(LPM)				Probit				With HH FE			
	12-18	14-18	16-18	19-29	12-18	14-18	16-18	19-29	12-18	14-18	16-18	19-29
main												
Payment× Post	-0.041	-0.003	-0.008	-0.065	-0.092	-0.057*	-0.033*	0.001	-0.014	0.009	0.008	-0.071
	(0.054)	(0.034)	(0.028)	(0.074)	(0.068)	(0.030)	(0.017)	(0.045)	(0.059)	(0.037)	(0.023)	(0.082)
Sq. Payment× Post	0.042	0.010	0.005	0.038	0.209	0.053	0.024	-0.012	0.040	0.010	0.007	0.039
	(0.031)	(0.015)	(0.012)	(0.037)	(0.138)	(0.058)	(0.066)	(0.047)	(0.033)	(0.019)	(0.011)	(0.039)
PW Days× Post	-0.030	0.043	0.090	-0.328**	0.070	0.094**	0.070*	-0.263***	-0.124	-0.034	-0.002	-0.341**
	(0.135)	(0.082)	(0.074)	(0.157)	(0.115)	(0.047)	(0.042)	(0.102)	(0.120)	(0.049)	(0.035)	(0.151)
Sq. PW Days× Post	0.337	-0.064	-0.167	1.101***	0.485	-0.092	-0.096	1.225***	0.697	0.251*	0.154*	1.080***
	(0.367)	(0.155)	(0.134)	(0.372)	(0.390)	(0.085)	(0.071)	(0.374)	(0.423)	(0.128)	(0.088)	(0.389)
Post Period (d)	-0.031	-0.009	-0.004	-0.107***	-0.054	-0.007	-0.003	-0.108***	-0.048**	-0.014	-0.010	-0.111***
	(0.021)	(0.010)	(0.010)	(0.030)	(0.034)	(0.009)	(0.006)	(0.031)	(0.023)	(0.016)	(0.011)	(0.039)
Number of Obs.	1137	1137	1137	1137	1137	1137	1137	1137	1137	1137	1137	1137
$adjR^2$	.016	.012	-.0026	.016					.046	.035	.007	.029
Prob > F	0.00	0.22	0.52	0.00					0.07	0.23	0.44	0.34
Household FE	No	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	-	-	-	-
Mean of the Dep. Var	0.04	0.02	0.01	0.04	0.04	0.02	0.01	0.04	0.04	0.02	0.01	0.04

Standard errors clustered at Woreda Level in Parenthesis
(d) for discrete change of dummy variable from 0 to 1
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 18: Probability of Observing a Household with an Incoming Female Member Due to Marriage

	Probit					With HH FE				
	12-18	14-18	15-18	16-18	19-29	12-18	14-18	15-18	16-18	19-29
main										
Payment× Post	0.112***	0.015	0.002	-0.000	-0.010	0.115	0.008	0.013	0.020	-0.060
	(0.039)	(0.016)	(0.025)	(0.018)	(0.027)	(0.098)	(0.029)	(0.029)	(0.028)	(0.060)
Sq. Payment× Post	0.171	-0.104	-0.097	-0.087	0.217***	-0.020	0.009	0.007	0.003	0.036
	(0.202)	(0.127)	(0.151)	(0.128)	(0.079)	(0.038)	(0.022)	(0.022)	(0.021)	(0.038)
PW Days× Post	-0.106	-0.026	-0.048	-0.038	0.051*	-0.215	0.006	-0.020	-0.025	0.144
	(0.083)	(0.054)	(0.050)	(0.042)	(0.029)	(0.168)	(0.101)	(0.088)	(0.087)	(0.106)
Sq. PW Days× Post	0.370	0.209	0.166	0.150	-0.053	0.524*	0.207	0.290	0.245	-0.282
	(0.240)	(0.173)	(0.132)	(0.123)	(0.124)	(0.297)	(0.277)	(0.231)	(0.215)	(0.184)
Post Period (d)	-0.005	0.003	-0.004	-0.003	-0.006	0.010	0.020	0.013	0.019	0.001
	(0.016)	(0.018)	(0.013)	(0.011)	(0.009)	(0.039)	(0.035)	(0.028)	(0.028)	(0.025)
Number of Obs.	899	899	899	899	899	899	899	899	899	899
$adjR^2$						.024	.012	.011	.013	.027
Prob > F	.	.	.	.	.	0.00	0.20	0.16	0.06	0.58
Household FE	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	-	-	-	-	-
Mean of the Dep. Var	0.03	0.02	0.02	0.01	0.02	0.03	0.02	0.02	0.01	0.02

Standard errors clustered at Woreda Level in Parenthesis
(d) for discrete change of dummy variable from 0 to 1
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

References

Andersson, Camilla, Alemu Mekonnen, and Jesper Stage, "Impacts of the Productive Safety Net Program in Ethiopia on livestock and tree holdings of rural households," *Journal of Development Economics*, 2011, *94* (1), 119 – 126.

Barrientos, Armando, "Social Transfers and Growth: What Do We Know? What Do We Need to Find Out?," *World Development*, 2012, *40* (1), 11 – 20.

Berhane, Guush, John Hoddinott, Neha Kumar, Alemayehu Seyoum Taffesse, Michael Tedia Diressie, Yisehac Yohannes, Rachel Sabates-Wheeler, Mulugeta Handino, Jermey Lind, Mulugeta Tefera, and Feyera Sima, "Evaluation of Ethiopia's Food Security Program: Documenting Progress in the Implementation of the Productive Safety Nets Programme and the Household Asset Building Programme," Mimeo, Washington, D.C:International Food Policy Research Institute 2011.

_ , _ , _ , and _ , "The impact of Ethiopia's Productive Safety Nets and Household Asset Building Programme: 2006-2010," International Food Policy Research Institute 2011.

Bertrand, Marianne, Sendhil Mullainathan, and Douglas Miller, "Public Policy and Extended Families: Evidence from Pensions in South Africa," *The World Bank Economic Review*, 2003, *17* (1), 27–50.

Deaton, Angus and Christina Paxson, "Economies of Scale, Household Size, and the Demand for Food," *Journal of Political Economy*, 1998, *106* (5), pp. 897–930.

Edmonds, Eric, Kristin Mammen, and Douglas Miller, "Rearranging the Family?: Income Support and Elderly Living Arrangements in a Low-Income Country," *Journal of Human Resources*, 2005, *40* (1).

Fafchamps, Marcel and Agnes R. Quisumbing, "Chapter 51 Household Formation and Marriage Markets in Rural Areas," in T. Paul Schultz and John A. Strauss, eds., , Vol. 4 of *Handbook of Development Economics*, Elsevier, 2007, pp. 3187 – 3247.

Foster, Andrew D., "Household Partition in Rural Bangladesh," *Population Studies*, 1993, *47* (1), 97–114.

_ and Mark R. Rosenzweig, "Household Division and Rural Economic Growth," *The Review of Economic Studies*, 2002, *69* (4), 839–869.

Gilligan, Daniel O., Hoddinott John, and Alemayehu Seyoum Taffesse, "The impact of Ethiopia's Productive Safety Net Program and Its Linkages," International Food Policy Research Institute (IFPRI) Discussion Paper 00839 2008.

Hoddinott, John and Tseday Mekasha, "Social Protection Programs and Household Demographic Dynamics," Unpublished 2015.

—, **Guush Berhane, Daniel O. Gilligan, Neha Kumar, and Alemayehu Seyoum Taffesse**, "The Impact of Ethiopia's Productive Safety Net Programme and Related Transfers on Agricultural Productivity," *Journal of African Economies*, 2012, *21* (5).

Hosegood, Victoria, Anne Case, and Cally Ardington, "Labor Supply Responses to Large Social Transfers: Longitudinal Evidence from South Africa," *American Economic Journal: Applied Economics*, 2009, *1* (1), 22–48.

Lanjouw, Peter and Martin Ravallion, "Poverty and Household Size," *The Economic Journal*, 1995, *105* (433), pp. 1415–1434.

Maitra, Pushkar and Ranjan Ray, "Resource inflows and household composition: evidence from South African panel data," *Journal of International Development*, 2003, *15* (8), 1037–1047.

Nino-Zarazúa, Miguel, Armando Barrientos, Samuel Hickey, and David Hulme, "Social Protection in Sub-Saharan Africa: Getting the Politics Right," *World Development*, 2012, *40* (1), 163 – 176.

Posel, Dorrit, James A. Fairburn, and Frances Lund, "Labour migration and households: A reconsideration of the effects of the social pension on labour supply in South Africa," *Economic Modelling*, 2006, *23* (5), 836 – 853. Ninth Annual Conference on Econometric Modelling for Africa, School of Economics, University of Cape Town, 2004.

Rosenzweig, Mark R., "Risk, Implicit Contracts and the Family in Rural Areas of Low-Income Countries," *The Economic Journal*, 1988, *98* (393), 1148–1170.

— **and Oded Stark**, "Consumption Smoothing, Migration, and Marriage: Evidence from Rural India," *Journal of Political Economy*, 1989, *97* (4), 905–926.

Schubert, Bernd and Rachel Slater, "Social Cash Transfers in Low-Income African Countries: Conditional or Unconditional?," *Development Policy Review*, 2006, *24* (5), 571–578.

Stark, Oded and David E. Bloom, "The New Economics of Labor Migration," *The American Economic Review*, 1985, *75* (2), pp. 173–178.

Stecklov, Guy, Paul Winters, Marco Stampini, and Benjamin Davis, "Do Conditional Cash Transfers Influence Migration? A Study Using Experimental Data from the Mexican Progresa Program," *Demography*, 2005, *42* (4), pp. 769–790.

Winters, Paul, Guy Stecklov, and Jessica Todd, "Household Structure and Short-Run Economic Change in Nicaragua," *Journal of Marriage and Family*, 2009, *71* (3), 708–726.

Chapter 3

A Closer Look at Exchange Rate Induced Inflation in Ethiopia

A CLOSER LOOK AT EXCHANGE RATE INDUCED INFLATION IN ETHIOPIA

Tseday J. Mekasha* Kiflu G. Molla†

Abstract

In this paper, we examine exchange rate pass-through to domestic prices with a particular focus on the Ethiopian economy. We have employed a Structural Vector Autoregressive (SVAR) model where identification is achieved based on a combination of short and long-run restrictions. The long-run identifying restrictions are derived from a simple small open economy macro model. The results from the impulse response analysis points to a pass-through rate that starts out small but steadily increases in subsequent quarters becoming close to complete in about two years. According to the variance decomposition results, however, other macroeconomic shocks, including shocks to aggregate demand, output supply as well as shocks to foreign financial flows and trade balance appear to play important role in explaining the variation in inflation in Ethiopia.

Keywords: Exchange Rate, Pass-Through, Inflation, Ethiopia
JEL Classification: F14, F31, F23

*Copenhagen University, Department of Economics. Copenhagen, Denmark; tseday.mekasha@econ.ku.dk; Tel +45 35 32 37 83

†Stockholm University, Department of Economics. SE-106 91 Stockholm, Sweden; kiflu.molla@ne.su.se; Tel +46 8 16 32 15

1 Introduction

The degree, speed and duration in which domestic price level responds to changes in exchange rate are crucial factors in understanding inflation dynamics and determining the effectiveness of monetary policy. In particular, whether the price level responds proportionately or less proportionately to changes in the exchange rate and the speed in which this occurs are among the crucial issues in Exchange Rate Pass-Through (ERPT) analysis.[1] For instance, if ERPT happens to be quick and complete, any induced nominal depreciation (devaluation) will immediately be matched by a proportional rise in prices which will cut short any potential change in the real value of the currency. In such cases monetary policy is less likely to be effective in changing the real exchange rate and enhancing competitiveness, even in the short run. On the other hand, in an environment where the rate of pass-through is low or incomplete, depreciations (devaluations) of the nominal exchange rate is highly likely to lead to depreciation of the real exchange rate and hence help to prevent loss of competitiveness. It is therefore important to have a quantitative estimate of ERPT to predict if nominal depreciations (devaluations) lead to increased domestic inflation or help to improve (at least maintain) export competitiveness of a country.

For countries in sub-Saharan Africa (SSA) in general and Ethiopia in particular, preventing loss of export competitiveness is a pressing policy challenge and the more so when faced with inflationary environments. In this regard the Ethiopian economy is a case in point where, in recent years, observing double digit inflation has become the norm than the exception.[2] In such inflationary instances, policy makers often resort to currency devaluation measures so as to improve/maintain export competitiveness. For example, following the recent inflationary pressures, the National Bank of Ethiopia (NBE) has taken various devaluation measures the major one being on September 2010 where the Ethiopian currency (Birr) was devalued by 20 percent. Moreover, both the World Bank and IMF are still calling for further devaluation of the Birr arguing that the overvalued exchange rate is the main factor dragging the country's export competitiveness. (See IMF (2014), World Bank (2014)). Although these devaluation measures seem unavoidable in the face of alarming inflation rates, excess demand for foreign currency and deterioration of the trade balance, it is not clear whether such measures will add up more to the inflationary pressure rather than helping to improve export competitiveness.[3] This is a contentious issue for other SSA countries as well and it begs an empirical exploration.

[1]The degree of pass-through (complete vs incomplete) to local currency price of imports by and large depends on the pricing behavior of foreign exporting firms. If foreign firms follow Producer Currency Pricing (PCP) strategy, then the exchange rate change will be fully reflected in local currency price of imports (case of complete pass-through). On the other hand, firms might pursue Local Currency Pricing (LCP) strategy where they absorb all (large part) of the exchange rate change into their markups. In this case, assuming price rigidity, local currency price of imports will not change with movements in exchange rate (case of low or incomplete pass-through). Another popular explanation often used in the literature to explain the issue of incomplete pass-through, without assuming price rigidity, is Pricing To Market (PTM) strategy of foreign firms. That is, firms adjust price of exports (in their home currency) by the same proportion as the change in exchange rate and thus exercising some degree of price discrimination across destinations. In this case, the local currency price of imports will not be affected by the change in exchange rate implying incomplete or low ERPT.

[2]The inflation rate reached a peak of 40 percent in quarter three of 2011, stayed above 20 percent for most of 2012 before it gets to a single digit in the second quarter of 2013 (see Figure 18, see also Geiger and Goh (2012)).

[3]For example, in its exchange rate assessment for the 2013/14 fiscal year, the fund indicated an REER overvaluation of the Birr in the range of 10 to 13 percent. It is further reported that policy makers in Ethiopia resist IMF's call for further devaluations on concerns of potential feedback effects on inflation. IMF (2014)

Against this background, the main objective in this paper is to empirically analyze the the extent of exchange rate induced inflation in Ethiopia. In particular, this paper makes an effort to answer important questions including how big is the rate of ERPT, how quickly it leads to inflation and for how long the effect will last for? In addition, for each consecutive year after the exchange rate shock, attempt is made to estimate the percentage variation in prices that is attributable to exchange rate shock and shocks to other macroeconomic variables. To get further insight on the results, the main analysis is further complemented with a disaggregated level evidence where we have estimated the degree of pass-through for food and non-food prices.

To achieve the aforesaid objectives, the paper employs a Structural Vector Auto Regressive (SVAR) approach where identification is achieved using a combination of long-run and short run restrictions. The long-run restrictions are derived from a simple small open economy macro model in the spirit of Clarida and Gali (1994). As discussed below this is unlike the common practice in the literature where studies rely on either single equation method or SVAR approach based on short run restrictions. As presented in Section 3, combining long-run and short run restrictions has the merit of addressing the apparent identification challenge in ERPT analysis, which arises due to the simultaneous relationship between price and exchange rate. To the best of our knowledge this paper is the first to use a combination of long-run and short run restrictions to overcome the identification challenge in ERPT analysis. This is one major contribution of the current paper.

Although the issue of ERPT is one of the extensively researched areas in international macroeconomics, the existing literature for the most part focuses on developed countries and the evidence in the case of developing countries, particularly those in SSA, is relatively limited. Among the few studies, Choudhri and Hakura (2006) using a number of developed and developing countries including some SSA countries find a positive relationship between the rate of pass-through and the average level of inflation. This is in line with the hypothesis suggested by Taylor (2000) that pass-through is low in low inflation environments. Similarly, for SSA countries, Razafimahefa (2012) investigates exchange rate pass-through and its determinants for all SSA countries and finds an average ERPT rate of about 40 percent and indicated a considerable heterogeneity in ERPT rates across SSA countries. Apart from this most of the remaining few papers on the region are country level studies. These include, Younger (1992) and Frimpong and Adam (2010) for Ghana, Mwase (2006) for Tanzania and Melesse (2014) for Ethiopia. The recent paper by Melesse (2014) studies sectoral consumer price inflation in Ethiopia using SVAR approach.

As indicated above the current paper uses a different identification strategy from the above papers. Specifically, these papers use differing methodologies, ranging from single equation method to SVAR approach mainly using short run restrictions. Aron et al. (2014) Unlike single equation model, which has been quite common in ERPT literature, SVAR modeling has the merit of allowing for endogenous interaction between exchange rate and other macroeconomic variables including prices. Although single equation models can also be informative, their treatment of movements in exchange rate as exogenous is problematic. This implicit assumption of single equation models ignores the possible reverse causation that runs from prices to exchange rate or the fact that prices and exchange rates potentially respond to the same shocks.

For example, as is also discussed in Ito and Sato (2007), in the case of floating exchange rate regimes, an increase in the domestic price level is highly likely to lead to nominal depreciations. On the other hand, in the case of fixed, adjustable pegged or managed floating exchange rate regimes, devaluation measures are in most cases policy responses to domestic inflation and/or to the underlying fundamentals of the economy (See also Hossain (2005), Shambaugh (2008) and Younger (1992)). In view of this, it is important to treat both exchange rate and prices as endogenous variables and this warrants the use of system of equations approach.

Although SVAR approach is preferable to single equation method for the reason discussed above, identification of the structural shocks is still a major concern. Papers that employ SVAR approach in analyzing ERPT (among others, see Ito and Sato (2008), McCarthy (2000, 2007) and papers cited above) identify the structural shocks using restrictions on the contemporaneous reactions of the endogenous variables to the structural shocks. Following the seminal work by Sims (1986), these papers impose restrictions on the short run effects of the structural shocks so as to get a recursive ordering of the variables in the system, that is, apply Choleskey decomposition to identify the structural shocks. Although this approach has generally been useful, its application in ERPT analysis is commonly criticized for imposing short run restrictions that do not have strong theoretical support. Given that monetary policy, prices and exchange rates do potentially affect each other, the recursive ordering of the variables in the system implied by the short run restrictions is hard to justify theoretically.

Cognizant of this problem, this paper mainly relies on theory guided long-run restrictions (neutrality properties) to recover the structural shocks following Blanchard and Quah (1989) (BQ) approach. Specifically, unlike the existing literature which is entirely dependent on short run restrictions, in this paper identification is achieved based on a combination of short run and long-run restrictions. Guided by a simple small open economy model, we put restrictions on the long-run effects of the structural shocks on the endogenous variables in the system and we combine this with two additional restrictions on the contemporaneous relationship between the variables in the system. This is particularly important in the case of developing countries where it is relatively difficult to determine the short run relationship between macro variables. Apart from its advantage in that it does not entirely rely on theoretically controversial restrictions imposed on the contemporaneous relationship between the variables in the model ,this approach as noted earlier, has its merit in dealing with the simultaneity between domestic prices and movements in nominal exchange rate. Further details of this methodology is discussed in Section 3 of the paper.

In order to derive economically plausible restrictions on the long-run relationships between the variables in the system, a modified version of the small open macroeconomic model of Clarida and Gali (1994) is used as a guideline. A variant of this model has recently been used by Shambaugh (2008) and Barhoumi (2009) in the context of ERPT analysis. Unlike these papers, however, the model applied in current paper is modified to include, among others, the role of external financial flows and shocks to trade balance.[4] These extensions are particularly relevant for developing countries and the more so for countries in SSA where these factors play a major role in determining movements of both exchange rates and prices. In the case of the Ethiopian economy, for example, running current account deficit is the norm and external financial flows (development aid and remittance) play a major role in financing this deficit. For instance, in 2013/14 Ethiopia's current account deficit reached 7.1 percent of GDP and

[4]We also differ from these paper in how we measure the rate of pass-through.

earnings from remittances that reach 7 percent of GDP and official transfers of 1.3 Billion USD in 2013/14 help to offset the current account deficit. (see IMF (2014)). Therefore, in addition to our identification strategy used in this paper, an explicit inclusion of shocks to trade balance and external financial flows in an open economy model is another contribution of the paper.[5]

Coming to the results, the response of CPI to a one standard deviation nominal exchange rate shock is found to be positive and statistically significant starting from the second quarter and onwards. The response cumulates over subsequent quarters and leaves the domestic price at a permanently higher level. In particular, even if the corresponding estimate of the rate of exchange rate pass-through to aggregate CPI starts out low, it quickly increases to more than 0.8 in about four quarters and becomes near complete in about two years after the shock. In terms of the relative importance of shocks, only about 7 percent of the variation in inflation in three to four years is explained by shocks to nominal exchange rate. Other macroeconomic shocks including shocks to BOP, aggregate supply and aggregate demand are found to explain the lion's share of the variation in inflation, particularly in longer horizons.

The rest of the paper is structured as follows: While Section 2 presents the theoretical model, Section 3 discusses the econometric methodology and identification strategies. Section 4 presents the data and relevant tests. The estimation results and related discussions are presented in Section 5. Finally, Section 6 gives concluding remarks.

2 Theory

This section presents the small open economy model which is used as a guideline for deriving the long-run identifying restrictions that we apply in the empirical analysis. In particular, the theoretical model for the most part draws on a stochastic version of the two country rational expectations open macroeconomic model presented in Clarida and Gali (1994). This model was originally developed by Obstfeld (1985) and has been extensively used to explain sources of real exchange rate fluctuations. In the context of exchange rate pass-through, however, application of this model has been quite limited. Only recently Shambaugh (2008) and Barhoumi (2009) used this model to explicitly study exchange rate pass-through.

As indicated earlier, we modify this model to make it more compatible with the underlying macroeconomic realities of developing countries in general and Ethiopia in particular. Specifically, we have added shocks to trade balance and external financial flows to this an otherwise standard rational expectation open macroeconomic model. The inclusion of trade balance shocks is important as the widening of current account deficit puts pressure on the level of exchange rate that policy makers want to maintain. For instance, according to IMF (2014), a REER depreciation of about 10 percent is needed in order to close the current account gap in Ethiopia. Similarly, given that Ethiopia is highly dependent on external assistance and earnings from remittances, shocks to external financial flows are also crucial in determining movements in the exchange rate and other macro variables in the economy. The above IMF report, for

[5]Even if there are extensions of open economy models that include shocks to oil price, commodity price and external flows, we have not come across papers that apply such extensions in the context of exchange rate pass-through applications.

example, points out that the 3.7 Billion USD average remittance inflow during the three years period 2011/12-2013/14 was higher than earnings from exports of goods in the same period. This, together with official transfers that reached 1.3 Billion USD in 2013/14, largely support the country's current account deficit. Shocks to these resource inflows will thus undoubtedly influence the level of price and exchange rate that prevail in the economy. For this reason, the aforesaid extension is important in the context of the Ethiopian economy. On top of the above extensions, we have also taken into account the limited capital mobility (closed capital account) nature of the Ethiopian economy. For a similar extension in the case of SSA countries with limited capital mobility, see the paper by Sissoko and Dibooglu (2006). These authors apply a similar extension to examine the sources of macroeconomic fluctuations with a particular focus on the exchange rate system in SSA countries.

In what follows, we first list out the necessary equations that comprise the system and move on to solving the long-run equilibrium of the model. In doing so, we start by describing the aggregate demand (IS-equation) given in Equation 1 below.[6] In this equation all the variables, except the real interest rate, are expressed in their log forms. Apart from explicitly writing out the autonomous components of net exports as well as the components of consumption, investment and government expenditure that are financed by foreign financial flows, Equation 1 represents the standard IS equation. Unless otherwise stated, all of the parameters used in this model are assumed to be positive.

$$y_t^d = \gamma[d_t + x_t + f_t - \rho[i_t - \mathbb{E}(p_{t+1} - p_t)] + \eta(s_t - p_t)] \tag{1}$$

While x_t represents autonomous net export, f_t stands for foreign financial flows (foreign aid and remittances) used to finance consumption, investment and government expenditure. On the other hand, d_t captures autonomous component of aggregate demand financed through domestic (internal) resources.

We have expressed the real interest rate in nominal terms using the relation $R_t = i_t - \Pi_{e,t}$, where i_t and $\Pi_{e,t}$ denote the nominal interest rate and expected inflation, respectively. We have also used the following approximation to define expected inflation ($\Pi_{e,t}$). That is, $\Pi_{e,t} = \mathbb{E}\frac{\Delta P_{t+1}}{P_t} \approx \mathbb{E}ln(1 + \frac{\Delta P_{t+1}}{P_t}) = \mathbb{E}ln(\frac{P_{t+1}}{P_t}) = \mathbb{E}(p_{t+1} - p_t)$ and hence the real interest rate can be given by $i_t - \mathbb{E}(p_{t+1} - p_t)$. Finally, $q_t = s_t - p_t$ is log of the real exchange rate where the foreign price level is normalized to one. The parameters γ, ρ and η respectively capture the Keynesian multiplier, response of investment to real interest rate and response of net export to real exchange rate.

The autonomous (exogenous) part of aggregate demand financed through domestic resources (d_t), autonomous net exports (x_t) and foreign financial flows (f_t) are assumed to follow a ran-

[6] In Appendix A, we list out the different equations that form aggregate demand in order to show how we have introduced the different shocks. While the aggregate demand equation we have in the appendix is expressed in levels of the variables, here we have expressed the same equation using the logarithms of the variables for the sake of convenience in the empirical model estimation and interpretation. In short, Equation 1 is not a log transformation of Equation 39.

dom walk process as given below (see equations 2-4):

$$d_t = d_{t-1} + \mu_t^d \tag{2}$$

$$x_t = x_{0,t-1} + \mu_t^x \tag{3}$$

$$f_t = f_{t-1} + \mu_t^f \tag{4}$$

Where μ_t^d is aggregate demand shock which can be taken as fiscal expansion or shifts in consumption of investment functions. On the other hand, μ_t^x and μ_t^f are assumed to capture shocks to net exports (trade balance) and foreign financial flows, respectively. In standard IS equation representation, μ_t^x is entered as an aggregate demand shock and is assumed to capture exogenous increases in exports or decreases in imports of the country concerned and represents a right ward shift in the IS curve.

Turning to the supply side of the goods market, the output supply equation in this small open economy model is given by Equation 5. As in Clarida and Gali (1994) output supply is assumed to be determined by the productive capacity of the economy ($\hat{y}_t$). We have also assumed that the productive capacity economy follows a random walk process as given in Equation 6.

$$y_t^s = \hat{y}_t \tag{5}$$

$$\hat{y}_t = \hat{y}_{t-1} + \mu_t^s \tag{6}$$

Where μ_t^s denotes aggregate supply shock and it can be considered as technology or labor supply shock.

The market clearing condition in the goods market $y_t^d = y_t^s = y_t$ establishes equilibrium in the goods market and we can thus rewrite the IS equation as:

$$y_t = \gamma[d_t + x_t + f_t - \rho[i_t - E_t(p_{t+1} - p_t)] + \eta(s_t - p_t)] \tag{7}$$

In the above IS equation, output (y_t) is positively related to real exchange rate $(s_t - p_t)$ and shocks to domestic demand (d_t), net exports (x_t) and foreign financial flows (f_t) and it is negatively related to the real interest rate.

Turning to the money market, demand for money (real money balance) is given by:

$$m_t^d - p_t = y_t - \lambda i_t \tag{8}$$

where λ is the response of demand for real money balances to nominal interest rate.

Money supply (m_t^s) on the other hand is assumed to be determined by the national bank of Ethiopia and, for simplicity, it is assumed to follow a random walk process given in Equation 9:

$$m_t^s = m_{t-1}^s + \mu_t^m \tag{9}$$

The market clearing condition in the money market, $m_t^d = m_t^s$, is given by the standard LM condition indicated in Equation 10.

$$m_t = p_t + y_t - \lambda i_t \tag{10}$$

Finally, we introduce a separate equation for nominal exchange rate, which is given as a function of the domestic money supply (m_t) and a disturbance term (e_t). The disturbance term captures other determinants of the nominal exchange rate like foreign prices (foreign money supply) and other exogenous disturbances to the nominal exchange rate. For a similar way of defining the nominal exchange rate, see also Wehinger (2000). This definition of the nominal exchange rate is partly based on the monetary approach to exchange rate determination which establishes that increases in domestic money supply leads to an equiproportionate depreciation of the home currency (see Dornbusch (1980)).[7] Similarly, Bacchetta and van Wincoop (2005) also derive the equilibrium nominal exchange rate as the ratio of home to foreign money supplies $(S_t = M_t/M_t^*)$. Assuming that the disturbance to the nominal exchange rate (e_t) follows a random walk process given by Equation 12, we have:

$$s_t = m_t + e_t \tag{11}$$

$$e_t = e_{t-1} + \mu_t^e \tag{12}$$

Long Run Equilibrium

As can be noted from above, the standard Uncovered Interest Rate Parity (UIP) equation is not used in the determination of interest rate. Given the underdevelopment of capital market in

[7] According to the monetary approach to exchange rate determination, the nominal exchange rate is given by $s_t = m_t - m_t^* - a(y_t - y_t^*) + b(i_t - i_t^*)$.

the Ethiopian economy or that the country has a closed capital account, UIP, which relies on the existence of perfect capital mobility, is unlikely to hold. Thus, before solving the long-run equilibrium solutions of this aggregate demand-aggregate supply model, we eliminate the interest rate variable from the goods market equilibrium (IS) condition using the relation described in the money market equilibrium (LM) condition. That is:

$$y_t = \gamma\left[d_t + x_t + f_t - \rho\left(\frac{1}{\lambda}(p_t + y_t - m_t) - \mathbb{E}[p_{t+1} - p_t]\right) + \eta(s_t - p_t)\right] \tag{13}$$

Using the definition of the nominal exchange rate given in Equation 11, we can re-write this Equation 13 as:

$$y_t = \gamma\left[d_t + x_t + f_t - \rho\left(\frac{1}{\lambda}(p_t + y_t - m_t) - \mathbb{E}[p_{t+1} - p_t]\right) + \eta(m_t + e_t - p_t)\right] \tag{14}$$

Re-arranging this we get a first-order difference equation in price level (Equation 15) which is used to solve for the long-run equilibrium price.

$$p_t = \frac{\lambda}{\rho + \rho\lambda + \eta\lambda}\left[d_t + x_t + f_t - \frac{\lambda + \gamma\rho}{\gamma\lambda}y_t + \frac{\rho + \eta\lambda}{\lambda}m_t - \rho\mathbb{E}p_{t+1} + \eta e_t\right] \tag{15}$$

Solving this first-order difference equation in price (see Appendix B for the derivation), the long-run equilibrium price level is given by:

$$p_t = \frac{\lambda}{\rho + \eta\lambda}[d_t + x_t + f_t] - \frac{\lambda + \gamma\rho}{\rho + \eta\lambda}y_t + m_t + \frac{\lambda\eta}{\rho + \eta\lambda}e_t \tag{16}$$

The above equation can also be used to define demand for equilibrium real money balances which is given by:

$$m_t - p_t = -\frac{\lambda}{\rho + \eta\lambda}[d_t + x_t + f_t] + \frac{\lambda + \gamma\rho}{\rho + \eta\lambda}y_t - \frac{\lambda\eta}{\rho + \eta\lambda}e_t \tag{17}$$

The equilibrium real exchange rate is derived as a rate that, given the existing patterns of trade and financial flows, is compatible with a financeable trade deficit and it is given by Equation 18. That is, in the long-run, a country's current account is said to be sustainable if the sum of trade deficit and foreign financial inflows become zero.

$$s_t - p_t = \frac{\theta}{\eta}y_t - \frac{1}{\eta}(x_t + f_t) \tag{18}$$

where θ and η in Equation 18 respectively capture the propensity to import out of disposable income and response of net exports to real exchange rate.

To show the long-run impact of the five structural shocks on the system's five endogenous variables, below we express the long-run solution to the model using the first difference of the endogenous variables:

$$\Delta y_t = \mu_t^s \tag{19}$$

$$\underbrace{\Delta(s_t - p_t)}_{\Delta q_t} = \frac{\theta}{\eta}\mu_t^s - \frac{1}{\eta}(\mu_t^x + \mu_t^f) \tag{20}$$

$$\Delta(m_t - p_t) = -\frac{\lambda}{\rho + \eta\lambda}[\mu_t^d + \mu_t^x + \mu_t^f] + \frac{\lambda + \gamma\rho}{\rho + \eta\lambda}\mu_t^s - \frac{\lambda\eta}{\rho + \eta\lambda}\mu_t^e \tag{21}$$

$$\Delta p_t = \frac{\lambda}{\rho + \eta\lambda}[\mu_t^d + \mu_t^x + \mu_t^f] - \frac{\lambda + \gamma\rho}{\rho + \eta\lambda}\mu_t^s + \mu_t^m + \frac{\lambda\eta}{\rho + \eta\lambda}\mu_t^e \tag{22}$$

$$\Delta s_t = \frac{\lambda}{\rho + \eta\lambda}\left(\mu_t^d - \frac{\mu_t^x + \mu_t^f}{\eta}\right) + \left(\frac{\theta}{\eta} - \frac{\lambda + \gamma\rho}{\rho + \eta\lambda}\right)\mu_t^s + \mu_t^m + \frac{\lambda\eta}{\rho + \eta\lambda}\mu_t^e \tag{23}$$

Equation 19 to 23 define the long-run restrictions that we use to identify the structural shocks in the system. In particular, assuming that prices are flexible in the long-run, in this model, output is considered to be supply determined. That is, demand and nominal shocks do not affect the level of output in the long-run. Thus, only supply shocks are assumed to have a long-run impact on the level of output, in the long-run. [8]

On the other hand, the real exchange rate is determined by shocks to output supply as well as shocks to balance of payments ($\mu_t^z = \mu_t^x + \mu_t^f$). Moreover, in the long-run, demand for real money balances is affected by all shocks except shocks to money supply. In this model, shocks to demand for real money balances are captured by shocks to aggregate demand. Furthermore, in the long-run, the domestic price level and the nominal exchange rate are affected by all shocks in the system. While shocks to domestic prices are captured by shocks to money supply, shocks to nominal exchange rate capture changes in nominal exchange rate that is not caused by changes in the other endogenous variables in the system.

3 Empirical Model and Identification

As pointed out in Section 1, the structural shocks in this paper are identified using a combination of long-run and short run restrictions. The long-run restrictions are done in the spirit of Blanchard and Quah (1989) approach. Accordingly, guided by the theory model presented in the previous section, we put restrictions on the long-run effects of the structural shocks on

[8] Blanchard and Quah (1989) also make similar simplifying assumption arguing that even if demand shocks do also affect long-run output, their effect is small relative to that of supply disturbances.

the endogenous variables in the system. We combine this with a short run restriction where we put two restriction on the contemporaneous relationship between some of the variables in the system. Identification using combination of short-run and long-run restrictions is first introduced by Gali (1992). For a similar approach see also Alexius and Post (2008). This approach is important to deal with the simultaneity between domestic prices and movements in nominal exchange rate. Another advantage of this method is that we do not need to entirely rely on restrictions imposed on the short run (contemporaneous) relationship between the variables in the model.

In this section we, therefore, specify a VAR model with five variables including output supply (y_t), real exchange rate (q_t), demand for real money balances ($m_t - p_t$), consumer prices (CPI_t) and nominal exchange rate (s_t). The corresponding five structural shocks to be identified are: supply shocks, shocks to the balance of payments (that is, shocks to trade balance and financial flows), shocks to aggregate demand, shocks to money supply as well as shocks to nominal exchange rate.

As can be seen from the previous section, supply shocks are found in changes in output whereas shocks to balance of payments are found in changes in the real exchange rate that are not caused by changes in output supply. Moreover, shocks to aggregate demand are given by changes in the demand for real money balances that are not caused by changes in output supply and real exchange rate. Similarly, shocks to money supply are given by changes in domestic prices that is not caused by changes in output supply, demand for real balances or the real exchange rate. Finally, shocks to nominal exchange rate are given by changes in nominal exchange rate that are not caused by changes in any of the four endogenous variables in the system.

Let μ_t be a vector that defines these structural shocks and ΔY_t be vector of the first difference of the endogenous variables in the VAR model:

$$\Delta Y_t = (\Delta y_t, \Delta q_t, \Delta(m_t - p_t), \Delta s_t, \Delta p_t)'$$

where Δy_t is output supply which is measured by gross domestic product in constant prices, Δq_t is real exchange rate, $\Delta(m_t - p_t)$ is real money balances, Δp_t is domestic price level as captured by consumer price index or alternatively food and non-food prices and Δs_t is nominal exchange rate.[9] All variables are expressed in logs. Since all the variables in the model are found to be non-stationary (as shown in Section 4), we have used these variables in their first differences to define the VAR model.

To start with the reduced form Vector Autoregressive (VAR) model:

$$\Delta Y_t = B_1 \Delta Y_{t-1} + B_2 \Delta Y_{t-2} + \ldots + B_p \Delta Y_{t-p} + u_t \quad \text{Or} \quad B(L)\Delta Y_t = u_t \tag{24}$$

[9] An increase in s_t is indicates a depreciation of the Ethiopian birr and vice versa.

Where B_i are 5×5 matrix of coefficients, $B(L) = B_0 + B_1 L + B_2 L^2 + \ldots + B_p L^p$, L is the lag order operator and $i = 1, 2, \ldots p$ represents the lag order. Equation 24 and the corresponding variance-covariance matrix (Σ_u) of the reduced form residuals (u_t) can be estimated using available information on the endogenous variables in the system. Let the variance-covariance matrix of the reduced form residuals be given by:

$$E(u_t u_t') = \Sigma_u \tag{25}$$

Assuming that the matrix polynomial B_i has all its roots inside the unit circle and is invertible, the above reduced form VAR(p) process will have the following infinite order moving average (Wold) representation:

$$\Delta Y_t \quad = \quad u_t + \Phi_1 u_{t-1} + \Phi_2 u_{t-2} + \ldots = \sum_{j=0}^{\infty} \Phi_j u_{t-j} = \Phi(L) u_t \tag{26}$$

Where $\Phi_0 = I_5$, $\Phi_s = \sum_{j=1}^{s} \Phi_{s-j} A_j$ for $s = 1, 2, \ldots$ and $\Phi(L) = B(L)^{-1}$

In its structural moving average form, the above VAR model can be expressed as:[10]

$$\Delta Y_t \quad = \quad C_0 \mu_t + C_1 \mu_{t-1} + C_2 \mu_{t-2} + \ldots = \sum_{i=0}^{\infty} C_i \mu_{t-i} = C(L) \mu_t \tag{27}$$

Where $C(L) = C_0 + C_1 L^1 + C_2 L^2 + \ldots$.

Using Equation 26 and 27 and noting that $\Phi_0 = I_5$, we have that:

$$u_t \quad = \quad C(0) \mu_t \qquad \text{for} \quad L = 0 \tag{28}$$

$$C(L) \quad = \quad \Phi(L) C(0) \qquad \text{for} \quad L = 0, 1, 2 \ldots \tag{29}$$

Using the relationship between u_t and μ_t given above (that is, using Equation 25 in Equation 28) and the assumption that structural innovations summarized in μ_t are mutually orthogonal and have a unit variance, we get:

$$\Sigma_u = C(0) E(\mu_t \mu_t') C(0)' = C(0) C(0)' \tag{30}$$

[10] Inverting $C(L)$, in Equation 27, we can also write the structural vector autoregressive (SVAR) representation of the model as $A(L) \Delta Y_t = \mu_t$.

Where $C(0)$ is a matrix of contemporaneous effects of the structural shocks and it models the short run relationship between the 5 variables in the system. Identification based on short run restrictions requires putting zero restrictions on the $C(0)$ matrix. On the other hand, identification of the structural shocks using long-run restrictions requires putting restrictions on the cumulative impulse responses which is given as: $\mathbf{C(1)} = C_0 + C_1 + C_2 + C_3 + \ldots$. As we have indicated in the introduction, in this paper we use a combination of short run and long-run restrictions to achieve identification.

Likewise, denoting the cumulative effect of the reduced form residuals as $\mathbf{\Phi(1)} = \Phi_0 + \Phi_1 + \Phi_2 + \Phi_3 + \ldots$, we can define the matrix of cumulative impulse responses $(\mathbf{C(1)})$ in terms of $\mathbf{\Phi(1)}$ and $C(0)$. To do this, we use repeated substitution into Equation 29 to arrive at $C_i = \Phi_i C_0$; and summing both sides of this equation over i gives us at the key equation for identification:[11]

$$\mathbf{C(1)} \quad = \quad \mathbf{\Phi(1)}C(0) \tag{31}$$

The aim is to identify the contemporaneous relationship ($C(0)$), the structural system dynamics as defined in ($\mathbf{C(1)}$) and the time-series of structural shocks (μ_t). With these, we can estimate the structural impulse responses functions (SIRFs) and compute the structural forecast error variance decomposition (SFEVD). What we need here is therefore to estimate the matrix $C(0)$, which, as indicated in Equation 30, is given by $C(0)C(0)' = \Sigma_u$. Using $C(0)$ together with $\mathbf{\Phi(1)}$, which can be computed from the reduced form VAR model, we can get the cumulative effects of the structural shocks using Equation 31.

Since the structural shocks are not observed or that one cannot directly recover the parameters of the structural moving average model in Equation 27, the reduced form representation is instead used to recover the structural shocks. Specifically, the structural shocks are obtained by transforming the reduced form residuals using Equation 28 and 30. However, since the reduced form model is itself under identified, we need to impose restrictions in the system. That is, the variance covariance matrix of the reduced form residuals in Equation 30 represents 25 equations in 25 unknowns, where, given that Σ_u is symmetric, 10 of the equations are redundant. We are therefore left with 15 equations to identify the 25 parameters, which implies that we need 10 additional restrictions to arrive at a just identified system.

To achieve identification one can put short-run, long-run or combination of short-run and long-run restrictions on the reactions of the endogenous variables to the structural shocks. Identification based on short-run (as in Sims (1986)) and long-run restrictions (as in Blanchard and Quah (1989)), require putting the remaining ten restrictions on the $C(0)$ and $\mathbf{C(1)}$ matrices, respectively. As indicated above, this paper uses a combination of short run and long-run restrictions and the specific restrictions we are imposing on the system are outlined below.

Using the predictions derived from the theory model in the previous section (see Equation 19 to 23) we obtained eight long-run restrictions. Adding two short run restrictions, which are specified below, we get the 10 restrictions needed for a just identified system. This amounts to

[11] For $L = 0, C_0 = C_0$, for $L = 1, C_1 = \Phi_1 C_0$, for $L = 2, C_2 = \Phi_2 C_0$, $\ldots$ $C_i = \Phi_i C_0$.

putting zero restrictions on the $\mathbf{C(1)}$ and $C(0)$ matrices shown in Equation 31. The long-run restrictions are:

1. Only supply shocks (μ_t^s) have long-run effects on output supply;[12]
2. In the long-run real exchange rate is determined by shocks to balance of payments ($\mu_t^z = \mu_t^x + \mu_t^f$) as well as shocks to output supply (μ_t^s);
3. Demand for real money balances, in the long-run, is affected by all shocks in the system (μ_t^s, μ_t^z, μ_t^d and μ_t^e) but shocks to money supply (μ_t^m).
4. In the long-run, the domestic price level is affected by all shocks in the system; specifically it is affected by shocks to output supply, aggregate demand, balance of payments as well as shocks to money supply and nominal exchange rate;
5. Finally, in the long-run the nominal exchange rate is also affected by all shocks in the system.

In addition to the above long-run restrictions, we have also introduced two additional restrictions on the short run (contemporaneous) relationship between the variables. Specifically, we assume that shocks to nominal exchange rate will not have short run (contemporaneous) effects on domestic prices. This is consistent with the idea that prices adjust sluggishly to changes in the nominal exchange rate. Similarly, money supply shocks will not have short-run (contemporaneous) effects on the nominal exchange rate. Given Ethiopia's managed floating exchange rate regime, it is highly likely that the nominal exchange rate will remain the same within the quarter-of the money supply shock. We impose these two additional short-run restrictions on the $C(0)$ matrix and we arrive at a just identified system.

Taking the above long-run restriction into account and using Equation 27, the long-run response of the endogenous variables to the structural shocks, with the expected signs predicted from the theory model, can be summarized in the following matrix:

$$\begin{bmatrix} \Delta y_t \\ \Delta q_t \\ \Delta(m_t - p_t) \\ \Delta s_t \\ \Delta p_t \end{bmatrix} = \overbrace{\begin{bmatrix} + & 0 & 0 & 0 & 0 \\ + & - & 0 & 0 & 0 \\ + & - & - & - & 0 \\ + & - & + & + & + \\ - & + & + & + & + \end{bmatrix}}^{\mathbf{C(1)}} \begin{bmatrix} \mu_t^s \\ \mu_t^z \\ \mu_t^d \\ \mu_t^e \\ \mu_t^m \end{bmatrix}$$

Given the above restrictions on the $\mathbf{C(1)}$ and $C(0)$ matrix, the relation indicated in Equation 31 can be summarized in the following matrix.

[12]One could reasonably argue about adding shocks to oil and/or commodity prices (capturing terms of trade shocks) as this could potentially affect output in the long-run. To limit the size of the system of equations we are dealing with, we instead add these variables as purely exogenous shocks to the system in the estimation of the structural VAR model.

$$
\overbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ UR & UR & 0 & 0 & 0 \\ UR & UR & UR & UR & 0 \\ UR & UR & UR & UR & 1 \\ UR & UR & UR & UR & 1 \end{bmatrix}}^{\mathbf{C(1)}} = \Phi(1) \overbrace{\begin{bmatrix} UR & UR & UR & UR & UR \\ UR & UR & UR & UR & UR \\ UR & UR & UR & UR & UR \\ UR & UR & UR & UR & 0 \\ UR & UR & UR & 0 & UR \end{bmatrix}}^{C(0)}
$$

where *UR* refers to unrestricted parameters.

In addition to the ten zero restrictions indicated above, we also have three identified parameters (indicated with ”1”s in the above matrices) giving us an overidentified system. However, since the estimation procedure that we are using requires the system to be just identified, we have not used these identified parameters for identification. Another point worth noting here is that, in our identification scheme presented above, the respective ordering of price and exchange rate doesn't change the empirical results. This is unlike identification schemes based entirely on short run or long-run restrictions and can deal with the simultaneous relationship between price and exchange rate in a better way.

4 Data and Pre-estimation Tests

The data used in this paper is quarterly data covering the period 1993q1-2014q4.[13] The main data sources are IMF International Financial Statistics (IFS), World Bank's World Development Indicators data base and national data sources including the National Bank of Ethiopia and Central Statistics office of Ethiopia. Since data on nominal effective exchange rate is not available for the period after 2010q4, we had to use the bilateral nominal exchange rate which is available until 2014q4. However, we have also checked the results using the multilateral exchange rate and the results are not qualitatively different.

Before we do a formal test on the time series properties of the variables, we have started with a preliminary assessment through graphical inspection of the level and difference of each of the series presented in Appendix C.[14] It is clear from these figures that, except the real exchange rate, the levels of all the variables (CPI, Nominal exchange rate, GDP, real money balances) appear to have a clear upward trend. The first difference of all the variables, on the other hand, is close to a stationary series. The plots based on the first difference of the variables do also

[13]The years before 1993 are not included as Ethiopia had a socialist regime that practiced heavy price control and followed a fixed exchange rate regime.

[14]As it is also indicated in Juselius (2006), it is “much easier to detect an outlier observation in the differences of a non-stationary variable than in the levels.”

reveal presence of some outliers in the aforementioned series that we take into account in the main analysis.[15]

Given the subjectivity inherent in the visual inspection, we move on to a formal testing of the behavior of each series so as to determine whether the variables in question are stationary or not. This is done using standard unit root tests including Augumented Dickey Fuller (ADF), Phillip Perron (PP) and Kwiatkowski, Phillips, Schmidt and Shin (KPSS) tests. The results from these tests are reported in Table 4 in the appendix.

As can be seen from Table 4, both the ADF and PP tests indicate that all the variables are non-stationary at levels. Confirming these results, the estimated test statistics from KPSS indicate that the null of stationarity is rejected in all cases. As can be seen from Table 5 in the appendix, the first difference of all the variables, according to all the three tests, appear to be stationary suggesting that all of the variables are integrated of order one, $I(1)$.

The main criticism with the above unit root tests is that these tests may fail to reject unit root in the presence of structural break. In particular, if there is a structural break in the series, the above tests can wrongly suggest unit root while the series might actually be stationary around a structural break. To address this concern, we check the time series properties of the variables by taking the presence of structural break into account. This is done using Clemente et al. (1998) method which endogenously determine the break dates from the data.

The Clemente et al. (1998) test gives two models: Additive outlier (AO) and Innovative outlier (IO) models which respectively capture sudden and gradual changes in the mean of the series. As can be seen from the results reported in Table 6, the null hypothesis of a unit root in each series cannot be rejected in all cases even after taking into account the presence of structural breaks in the series (the critical value of this test associated with a 5 percent level of significance is equal to -5.49). Therefore, based on the above test results we continue our analysis treating all the variables as having a unit root process.[16]

5 Results and Discussion

In this section, we present the results from the estimation of the structural VAR model with five variables: Real GDP, real exchange rate, demand for real money balances, nominal exchange rate and consumer price index or alternatively its disaggregated components. To control for seasonal variations, account for outlier observations and structural breaks, we have respectively included quarterly dummies, impulse dummies as well as a shift dummy in the structural VAR estimation.[17] We have also included world prices of oil and coffee as purely exogenous variables potentially affecting the endogenous variables in the system. Given that Ethiopia is an oil

[15] Food and Non-food price indices are also found to be integrated of order one even after accounting for structural breaks. Trends in the level and first difference of these variables is given in Figure 16.

[16] We have also checked if there is any co-integration between the variables in the system using Johansen maximum likelihood test for co-integration. The results show one co-integrating relationship among the variables in question (see Table 7) for the results.

[17] As can be seen from trends in GDP (Figure 10) and CPI (Figure 12), there is a clear structural break (around 2003q1) and we have captured this using a shift dummy.

importing country and heavily dependent on commodity (mainly coffee) exports, these two variables, capturing terms of trade shocks, are important sources of macroeconomic shocks.

For the structural VAR estimation, a lag length is selected using the standard lag selection methods including LR, FPE, AIC, HQIC and SBIC. Except the LR method which suggests a lag length of four, all the other methods suggest a lag length of two. We thus choose a lag length of two as this is suggested by most of the tests and as it is also found to be sufficient to clear the autocorrelation in the disturbances. We have also run tests for autocorrelation, normality and VAR stability (results are attached in the appendix, see Table 8, Table 9 and Figure 15).

5.1 Response of CPI to Macroeconomic Shocks

One of the main tools in structural VAR analysis is impulse response function which shows the contemporaneous and an overtime response of endogenous variables in the system to one standard deviation increase in the structural shocks. Since the long-run constraints (restrictions) we have used for the SVAR estimation are motivated by economic theory, it is possible to attach causal interpretation to the resulting Structural Impulse Response Functions (SIRFs). In each case, the SIRFs are plotted with the 0.16 and 0.84 percentile (confidence) bands. Following recommendations by Sims and Zha (1999) (see also Estima (2010)), the confidence bands used to test the statistical significance of the impulse responses are computed using Monte Carlo integrations method with 10000 draws.

As indicated above, we estimate the SVAR model in the log first difference of the variables and hence the impulse response functions indicate the responses of the growth rates of the variables rather than their levels. We therefore base our analysis on the cumulated impulse response functions which can be interpreted as responses of the logs of the variables in question. In this case, the estimated impulse response functions will have percentage interpretation; where 100 $\times$ the cumulative impulse response estimates indicates the percentage change in the endogenous (responding variable) following a one standard deviation in the shock variable.

As pointed out in the introduction, the main identification challenge in empirically estimating the rate of exchange rate pass-through to domestic prices is deciding the ordering of the variables in the VAR; specifically the respective ordering of prices and nominal exchange rate. In this paper, using identification strategy that is based on a combination of short and long-run restrictions, we have managed to make the ordering of these two variables irrelevant for the results. It is therefore important to emphasize that, although the results reported below are done based on ordering of the variables that puts the nominal exchange rate last in the system, the results do not change even if we instead order price last in the system.

Response of CPI to other Macroeconomic Shocks in the System

Before turning our focus to the central theme of this paper, that is, exchange rate induced inflation, we start by investigating the impact of the structural shocks on the system's endogenous variables. In doing so our main focus is the response of domestic prices to other macroeconomic shocks, apart from nominal exchange rate, so as to assess the roles these other shocks play in determining the domestic price level in Ethiopia. In the process, we also assess if the impulse response results are in line with the predictions of the theory model presented in this paper and, more generally, of other standard small open economy macro models.

To start with supply shock, as can be seen from Figure 1, on the left panel, aggregate supply shock, as theoretically expected, has a price dampening effect. In particular, a one standard deviation positive shock to aggregate supply appears to have a negative and statistically significant effect on the domestic price level that converges to -1.9 percent in about nine quarters after the shock. On the other hand, all the remaining variables in the system are found to respond positively to supply shock. The increase in the demand for money following a supply shock is in line with standard economic theory (See Equation 8). Moreover, the increase (depreciation) of nominal exchange rate following a supply shock can be justified under certain conditions. In particular, as can be seen from Equation 23, a positive supply shock can induce nominal depreciation if the elasticity of imports to disposable income (θ) is high. Similarly, the real exchange rate is also found to depreciate following a positive supply shock. In view of the observed nominal depreciation coupled with the decline in domestic price following a supply shock, the real depreciation is expected.

Moving to shocks to balance of payments (BOP), which captures shocks to trade balance and external financial flows, as can be seen from the right panel of Figure 1, BOP shock has a positive and statistically significant impact on CPI, both on impact and in subsequent quarters. A one standard deviation BOP shock is found to have a persistent positive impact on CPI that starts out below 2 percent, on impact and converges to around 3 percent on the eighth quarter after the shock. Since BOP shocks are in a way capturing aggregate demand shocks that are financed through external financial flows, their positive impact on domestic price is in line with a priori expectation. Apart from affecting domestic prices, BOP shocks are also found to decrease (appreciate) nominal and real exchange rate and to decrease the demand for money balances as is also predicted by the theory.The lower money demand in response to a positive shock to BOP can plausibly be explained by agents fear of expected inflation. On the other hand, a positive shock to BOP is not found to have a statistically significant effect on supply.

Similarly, aggregate demand shocks are also found to have a positive impact on the domestic price level as expected. In particular, following a one standard deviation shock to aggregate demand, the increase in domestic price level converges to 1.7 percent in about seven quarters after the shock. Regarding the impact on other endogenous variables, while a positive shock to demand does not have a permanent effect on output and real exchange rate, nominal exchange rate is found to respond positively both on impact and subsequent quarters and all the estimates are statistically significant. Given that both domestic price level and nominal exchange rate respond or adjust to a positive shock in demand, the lack of impact on real exchange rate is expected. On the other hand, contrary to expectation, the impact of demand shock on real money demand is not statistically significant in all horizons.

Finally, the estimated effect of shocks to money supply, which captures the variation in the price level that is not explained by the above three shocks, is found to have a positive impact on domestic price as can be seen from the cumulative impulse responses depicted in Figure 2. In particular, the effect of money supply shock on domestic price level starts out high, at 0.45 percent, on impact, declines in subsequent quarters and converges to about 0.43 percent nine quarters after the shock. Looking at the impact of money supply shock on other variables in the system, shocks to money supply do not have impact on output. In the case of real exchange rate, however, monetary shocks induces appreciation of the real exchange rate until the first quarter after the shock. This is mainly following the restrictions we put on the model. In particular, since nominal exchange rate in our model is restricted not to respond to monetary shock in the short run, it is inevitable for the real exchange rate to fall (appreciate) for a short while, until nominal exchange rate adjusts and offsets the rise in domestic price induced by money supply shocks.

Overall, as can be noted above, the estimated impulse response functions in almost all cases make sensible predictions or concur with theoretical predictions giving credence to the estimated model. In particular, the direction of the response of domestic price level to supply, demand, BOP and money supply shocks is all in line with both the theory model presented in Section 2 and/or is broadly consistent with other open economy macro models. Moreover, in all cases, the impact of the shocks on the domestic price level is statistically significant as can be seen from the confidence bands of the respective impulse responses in Figure 1 and Figure 2. Below we turn our attention to answering the main question of this paper, that is, to investigate the extent of exchange rate induced inflation, or the degree and speed of exchange rate pass-through to domestic prices, in Ethiopia.

Figure 1: Cumulative Impulse Responses to Supply and BOP Shocks

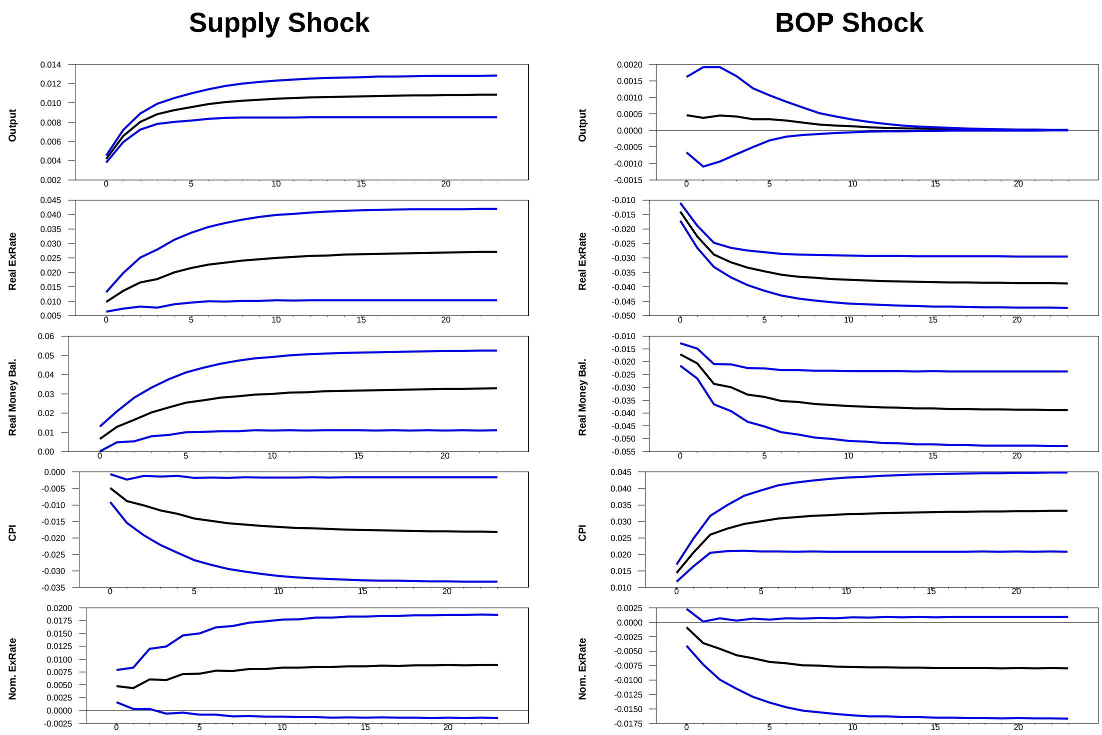


Figure 2: Cumulative Impulse Responses to Aggregate Demand and Money Supply Shocks

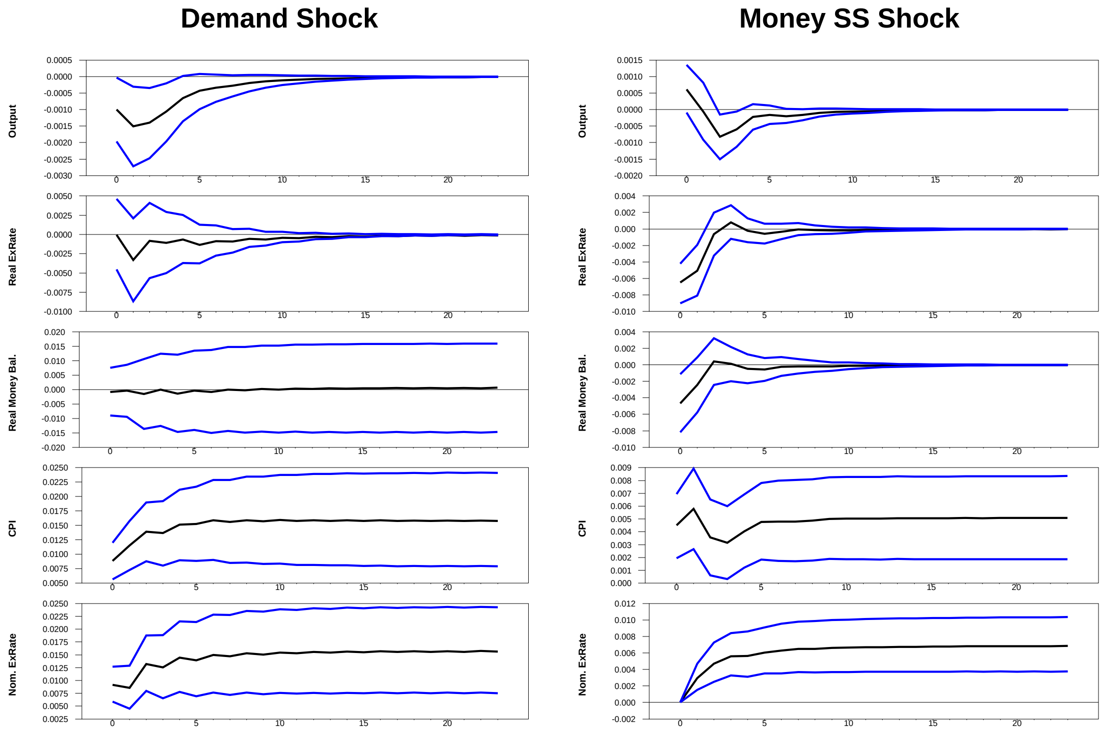
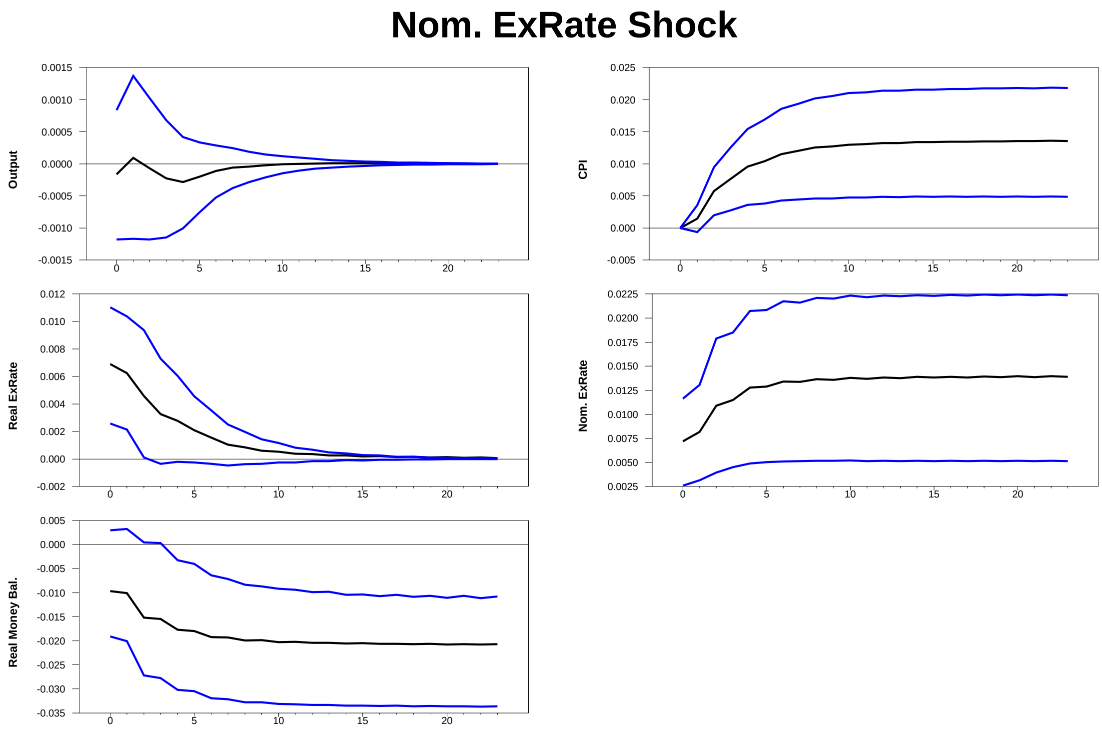


Figure 3: Cumulative Impulse Responses to Nominal Exchange Rate Shocks

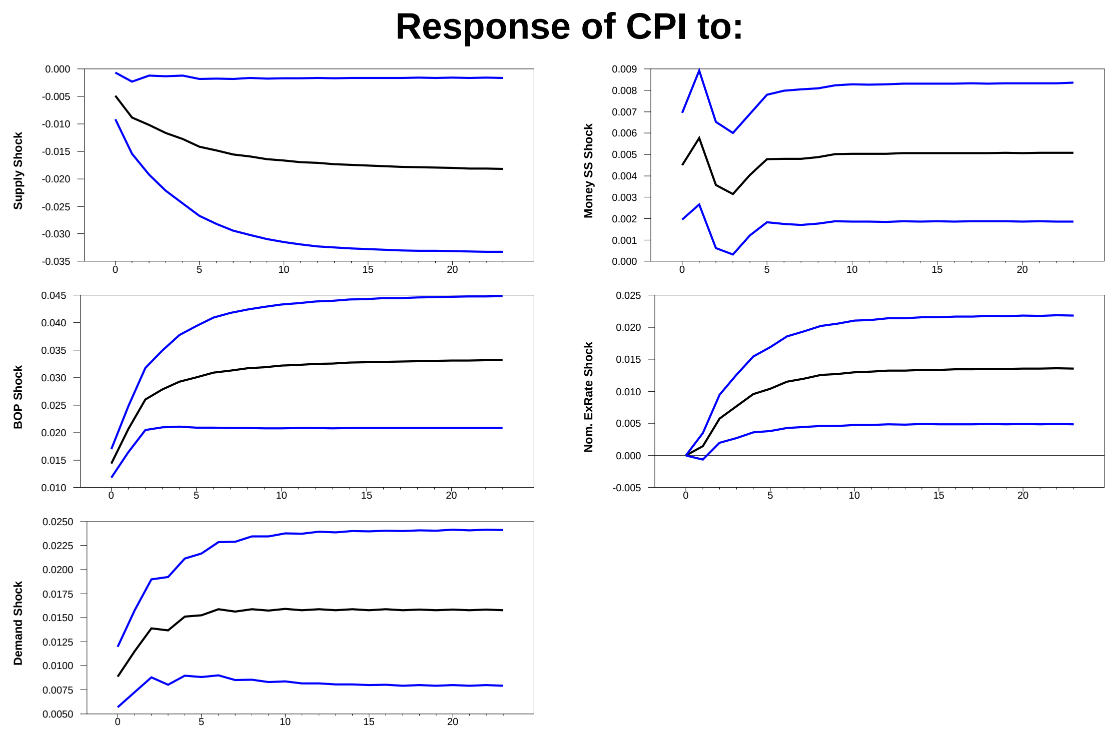


Evidence on Exchange rate induced Inflation

Figure 4 summarizes the cumulative structural impulse responses of the domestic price level to the different macroeconomic shocks in the system. Our focus in this section is on the impact of shocks to nominal exchange rate in domestic price level. Response of CPI following a shock (a depreciation) in the nominal exchange rate can be seen from the second plot on the right side of Figure 4.

On impact, the structural impulse response of CPI following a nominal exchange rate shock appears to be zero. This is expected as we have restricted the contemporaneous impact of a shock in the nominal exchange rate on CPI to be zero. In the first quarter after the nominal exchange rate shock, the impulse response of CPI becomes positive, albeit, statistically insignificant. From the second quarter onwards, however, the estimated impulse response of CPI to shocks in nominal exchange rate appears to be positive and statistically significant.

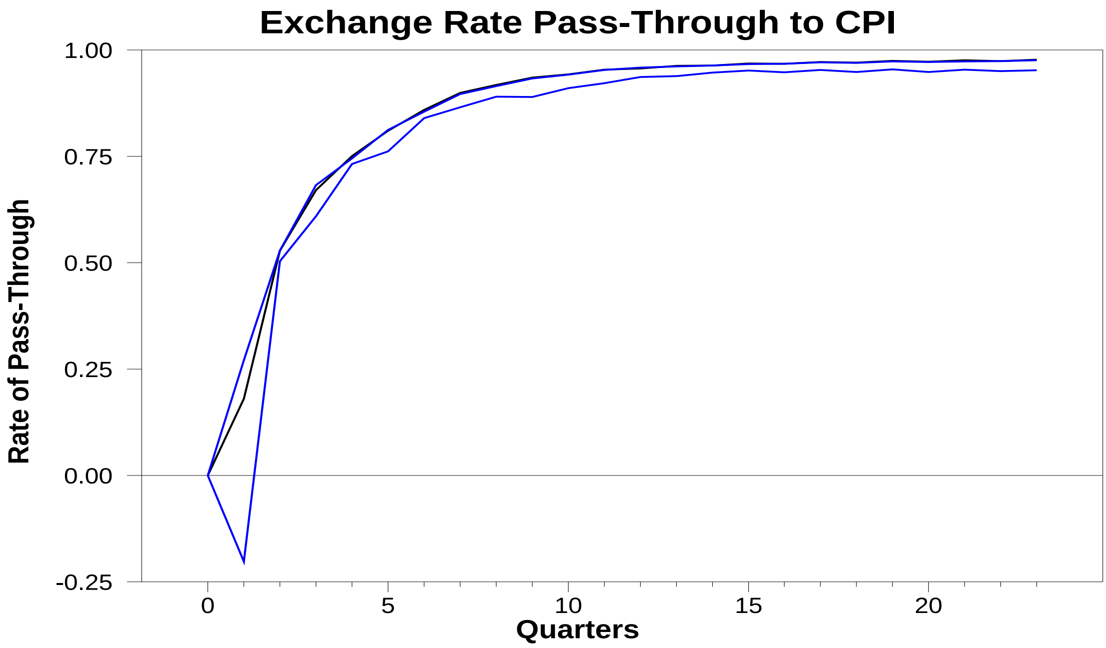
Figure 4: Cumulative Response of CPI to Nominal Exchange Rate and other Shocks



The estimated response of domestic price to a nominal exchange rate shock cumulates over periods leaving it at a permanently higher level. Specifically, as can be noted from Figure 4, a one standard deviation shock to nominal exchange rate is associated with a 0.60 percent and 1.0 percent increase in the price level in the second and fifth quarters after the exchange rate shock, respectively. Overtime the price response converges to 1.3 percent in about three years after the exchange rate shock.

To give the above results a more meaningful interpretation, below, we have calculated the rate of exchange rate pass-through (ERPT) to the domestic price level as the percentage change

Figure 5: Estimated Rates Exchange Rate Pass-Through



in the price level that is due to a one percent change in the nominal exchange rate. To this end, we have followed the standard practice in the literature and divide the cumulative response of CPI to a one standard deviation exchange rate shock by the respective cumulative response of the nominal exchange rate to the same shock. The estimated rates of ERPT we have calculated in this way for each quarter after the exchange rate shock along with its confidence bound is depicted in Figure 5.

As can be noted from Figure 5, ERPT to domestic prices in Ethiopia starts low (about 17 percent in the first quarter after the exchange rate shock and grows quickly to 53 percent in the second quarter after the shock. About 60 percent of the price change already takes place in the second quarter and in about two years (eight quarters) after the exchange rate shock, the rate of ERPT reaches 90 percent and it latter converges to 97 percent. Although ERPT is known to be larger for developing countries in general, the estimate that we find here for the Ethiopian economy can be considered even larger, the more so for ERPT to CPI. Different factors, including the persistence of the exchange rate in Ethiopia, the country's heavy reliance on imports as well as the inflationary environment in the country can potentially explain the high exchange rate pass-through to domestic price observed for Ethiopia. Together with the potentially low importance of the Ethiopian economy to foreign exporters, the high rate of pass-through estimated in this paper is in line with the the predictions of producer currency pricing models (among others, see McCarthy (2007), Taylor (2000)).

In sum, the above evidence from the impulse response shows that exchange rate pass-through to consumer prices in Ethiopia, although starts out at a lower level, is found to increase in subsequent quarters and becomes near complete two years after the exchange rate shock; leaving the domestic price at a permanently higher level.

Apart from shocks to nominal exchange rate, we have also seen that domestic price level is found to respond positively to shocks to BOP, aggregate demand and money supply (See Figure 4). In particular, while the domestic price level responds negatively to aggregate supply shocks, the price level responds positively to BOP, demand and money supply shocks. This shows that other macroeconomic shocks are also important in defining movements of CPI in Ethiopia.

To get a better insight on how the different macroeconomic shocks contribute to variations in domestic inflation (price level) in Ethiopia, below we have presented results on variance decomposition of CPI.

5.2 Decomposing the Variation in Consumer Price

A large exchange rate pass-through estimate does not necessarily mean that depreciation (devaluation) of the nominal exchange rate is the only or the most important source of variation in inflation in Ethiopia. Despite observing a very large exchange rate pass-through, the exchange rate shock might have small contribution in explaining variation in inflation if the size of the shock is small. Determining the percentage variation in domestic price level that can be attributed to shocks to nominal exchange rate as well as shocks to other macroeconomic variables in the system is therefore the next relevant exercise. Table 1 presents the structural forecast error variance decomposition (SFEVD) which shows the relative importance of each structural shock in explaining variations in growth rates of consumer prices in Ethiopia.[18]

To start with our variable of interest, as can be seen from Table 1, shocks to nominal exchange rate are found to play no role in explaining the variance in inflation in the first quarter. This is expected given the identification restriction we imposed on the contemporaneous relationship between these two variables. However, even in the second and third quarters, shocks to nominal exchange rate tend to explain only 0.3 and 1.8 percent of the variation in inflation, respectively. The percentage share of the variance in inflation that is attributable to changes in nominal exchange rate appears to increase in subsequent quarters. In particular, after 12 quarters (three years) and 16 quarters (four years), 7.3 and 7.9 percent of the variation in growth rate of consumer prices in Ethiopia is explained by shocks to nominal exchange rate, respectively. From the fourth quarter onwards, shocks to nominal exchange rate appear to explain larger percentage of the variation in inflation than do money supply shocks.

As can be seen above, despite the high degree of pass-through, nominal exchange rate shock appear to determine only a smaller percentage of the variation in inflation. This can partly be explained by the fact that most of the observed changes in nominal exchange rate (NER) of the Ethiopian birr are very small depreciations of the birr; although large, devaluations are rare. Another reason can be due to the fact that NER shocks are identified in the system through changes in NER that is not explained by other shocks in the system.

Nominal exchange rate shocks do play important role in determining the variance in the growth rate of demand for real money balances as well as the nominal exchange rate itself.

[18] Since we have estimated the model in log differences, it should be noted that we are decomposing the variance in the growth rates of the endogenous variables, and not the variance in the levels of the variables. Moreover, since we have reported the median SFEVDs, the decompositions attributable to the different shocks may not add up to 100%.

Table 1: Decomposition of Variance for Consumer Price Index

Horizon	Supply Shock	BOP Shock	Demand Shock	Money SS Shock	Nom. ExRate Shock
1	7.3	59.2	21.9	5.2	0.0
	[0.8,22.7]	[37.4,77.7]	[8.7,38.6]	[1.0,13.0]	[0.0, 0.0]
2	10.1	59.2	19.2	4.6	0.3
	[1.4,27.7]	[37.8,77.2]	[7.7,34.6]	[1.1,11.1]	[0.0, 1.1]
3	10.0	60.1	18.4	2.8	1.8
	[1.4,28.4]	[39.0,77.4]	[7.3,33.9]	[0.8, 7.2]	[0.4, 4.5]
4	10.4	59.9	17.3	2.0	3.0
	[1.4,29.8]	[39.0,77.2]	[6.6,32.7]	[0.6, 5.4]	[0.6, 7.2]
8	11.5	55.1	16.7	1.5	6.1
	[1.6,33.8]	[34.8,73.1]	[4.9,33.4]	[0.4, 4.2]	[1.2,13.8]
12	12.2	52.6	16.3	1.4	7.3
	[1.6,35.7]	[32.3,71.0]	[4.2,34.0]	[0.4, 4.0]	[1.4,16.4]
16	12.6	51.4	16.0	1.3	7.9
	[1.7,36.9]	[30.8,70.2]	[3.8,34.4]	[0.3, 3.9]	[1.5,17.6]

Over the horizons, about 37 to 23 percent of the variance in the growth rate of demand for real money balances is attributable to nominal exchange rate shocks. Given that the interest rate in the Ethiopian economy is only partially liberalized and that it is not regularly adjusted, expected inflation, exchange rates and foreign financial inflows play important role in explaining the variation in the demand for money (see also Loening et al. (2009))

The variation in inflation that is due to own shock (money supply shock) is found to be relatively small, especially at longer steps. In particular, money supply shocks contributes only to 5.2 and 4.6 percent of the variation in inflation in the first and second quarters, respectively. The contribution of money supply shocks tend to decrease over the horizons and in the fourth year only 1.3 percent of the variation in inflation in Ethiopia is due to innovations of money supply. The relatively smaller contribution of money supply shocks in determining the variance in inflation can be expected given that money supply shocks are identified through changes in the price level that is not caused by the other shocks in the system. In general, money supply shocks do not appear to play any significant role in determining the variation in any of the endogenous variables in the system.

From Table Table 1 it is also clear that, at shorter steps, most of the one-step forecast error variance in domestic inflation is explained by shocks to BOP, which is identified through changes in the real exchange rate that is not explained by output supply shocks. In particular, BOP shocks explain 59 and 60 percent of the variation in CPI in the first and fourth quarters, respectively. Afterwards, the contribution of innovations in the BOP in explaining the variation in inflation appear to decrease slightly and yet BOP shocks remain the prime driver of variation in inflation; after three years 53 percent of the variation in inflation is attributable to BOP shocks. Although this seems an anticipated, since BOP shocks are part of aggregate demand shocks that are associated with trade balance and foreign financial flows, both of which are important in the Ethiopian economy, this result shouldn't come as a surprise. BOP shocks also play important role in explaining variations in the growth rates of the real exchange rate (47 to 72 percent), demand for money (42 to 40) and the nominal exchange rate (2.4 to 8.2 percent) (see D).

Similarly, aggregate demand shocks, identified through changes in the demand for real money balances that is not explained by supply or BOP shocks, are also found to play important role in determining the variation in inflation. Specifically, shock to aggregate demand contribute to 21 and 17.3 percent of the variation in inflation in the first and fourth quarters, respectively. This makes aggregate demand shock the second most important shock deriving variations in CPI in the short-run. Over the horizons, the contribution of shocks to aggregate demand declines only slightly where, even after three years (12 quarters), 16.3 percent of the variation in inflation is attributable to aggregate demand shocks. The important role of aggregate demand shocks in explaining variations in inflation is expected given that these shocks are mainly meant to capture expansionary fiscal policies. Aggregate demand shocks are also found to explain a sizable share of the variation in the growth rate of the nominal exchange rate, where about 50 to 35 percent of the variance in the growth rate of the nominal exchange rate is found to be attributable to aggregate demand shocks.

Moreover, shocks to output supply also appear to have a major role in determining the variation in consumer prices, especially at longer steps. As can be seen from Table 1, even if the importance of shocks to output supply is relatively small in the short run, its contribution increases over time. In particular, in the second year and afterwards, shocks to output supply becomes important factor contributing to about 12 and 13 percent of the variation in inflation in third and fourth years, respectively.

Overall, even if BOP shocks, shocks to demand and aggregate supply are found to be the major deriving forces determining variations in the growth rate of domestic prices in Ethiopia, the importance of shocks to nominal exchange rate cannot be ignored. In the third year and after wards, more than 7 percent of the variation in inflation is explained by shocks to nominal exchange rate. This, together with the strong cumulative impulse response of CPI following shocks to nominal exchange rate, has its own implication for effectiveness of monetary policy. That is, impacts of changes in nominal exchange rate (devaluation measures) on real variables will be lower than anticipated. As can be seen from Table 12, in the first and second quarters 11.3 to 4.3 percent of the variance in the growth rate of the real exchange rate is attributable to nominal exchange rate shocks. In view of this, the devaluation measures aimed at improving (maintaining) the country's export competitiveness may not be able to achieve the desired targets since such measures will be partly matched by a rise in the domestic price level.

We have presented the variance decompositions of the other endogenous variables in Appendix D. As it is theoretically expected, the variance in growth rate of output supply is mainly determined by its own shock and this is the more so at longer horizon where the other shocks do not have meaningful contribution. Specifically, over the long horizon, 98 percent of the variation in output supply is attributable to a shock to output supply itself. Over shorter horizons, on the other hand, shocks to BOP as well as shocks nominal exchange rate and money supply play some role in determining variations in the growth rate of output (see Table 11). While most (71.5 to 47 percent) of the variation in the growth rate of the real exchange rate is due to own (BOP) shocks, 23 to 26 percent of the variation is attributable to output supply shocks. Money supply and nominal exchange rate shocks, on the other hand, appear to have some role in the short run (see Table 12). As indicated above, shocks to BOP (41 to 43 percent) and nominal exchange rate (23 to 37 percent) have important contributions towards explaining the variance in the growth rate of demand for real money balances. At longer steps, supply shocks are also important in explaining the variance in the growth rate of demand for real money balances; after the second year, more than 23 percent of the variation is attributable to output supply shocks (see Table 13). The variance in the growth rate of the nominal exchange rate can be attributed to the different shocks in the system, no one shock is responsible for the majority of the variance in the growth rate of the nominal exchange rate. However, over the horizons, aggregate demand shocks (35 percent to 50 percent) and nominal exchange rate shocks (24 percent 25 percent) appear to have important role (see Table 14).

5.3 Impulse Response of Disaggregated CPI: Food Price and Non-Food Price Indices

In this section, we look at the response of disaggregated price indices to changes in nominal exchange rate. Specifically, we estimate the structural impulse responses of food and non-food price indices to nominal exchange rate as well as other shocks in the system. To this end, we employ the same model as above except that now we had to limit the sample period between 1997q4 to 2014q4 because of lack of data on the food and non-food price indices for the period prior to 1997.

As can be see from Table 10 food items constitute the larger share of the consumer basket. As can also be seen from Figure 16, which shows trends in aggregate, food and non-food price indices over the sample period, trends in the food price index closely mimics trends in aggregate price index. For these reasons we expect food prices to respond to nominal exchange rate shocks in a similar way as aggregate consumer prices and this appears to be the case as can be seen from the impulse response estimates presented in Figure 6.

As can be seen from the cumulative structural impulse response estimates depicted in Figure 6, the response of food prices to changes in nominal exchange rate is positive and statistically significant in the second and subsequent quarters. As in the case of aggregate price index, a one standard deviation shock to nominal exchange rate leaves food prices at a permanently higher level. For each quarter after the exchange rate shock, Figure 7 plots the corresponding exchange rate pass-through rates to food prices which indicates that pass-through to food price in Ethiopia becomes near complete in about five to six quarters after the exchange rate shock. The observed high rate of ERPT to food prices can be explained by not only the relatively high import content of the food price index, but also by the fact that the production and distribution of food items involve imported items.

Similarly, the cumulative structural impulse response of non-food prices to a one standard deviation nominal exchange rate shock is also positive and statistically significant, again starting

Figure 6: Response of Food Price to a Nominal Exchange Rate Shocks:

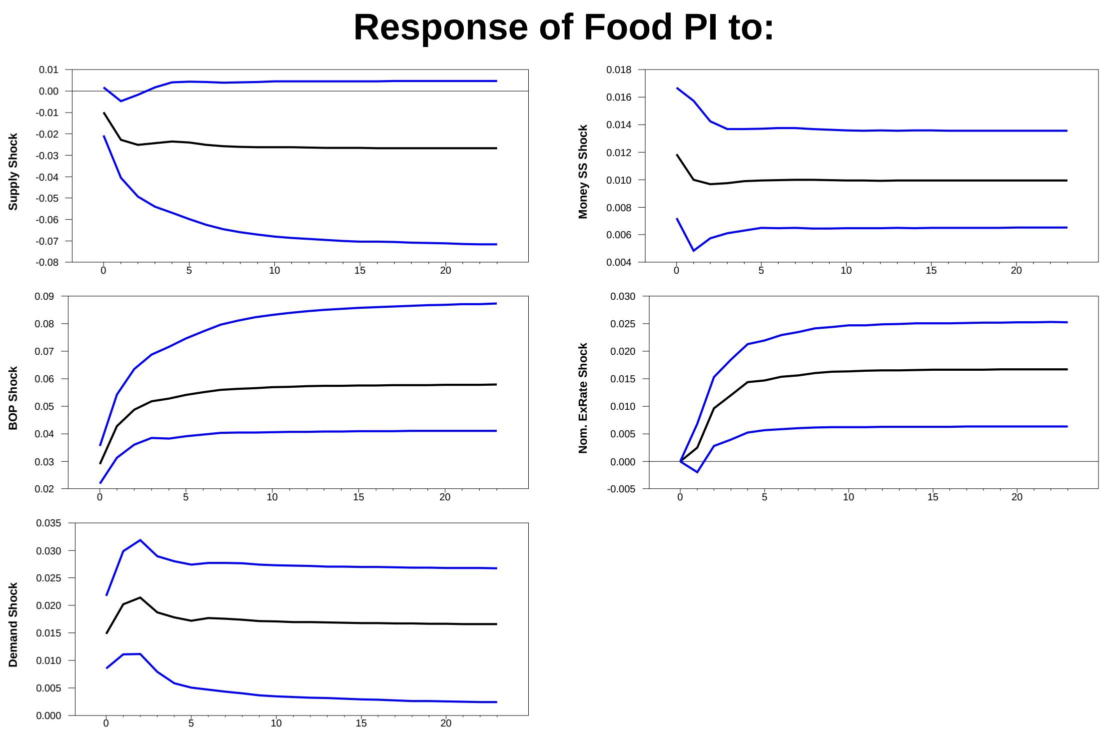
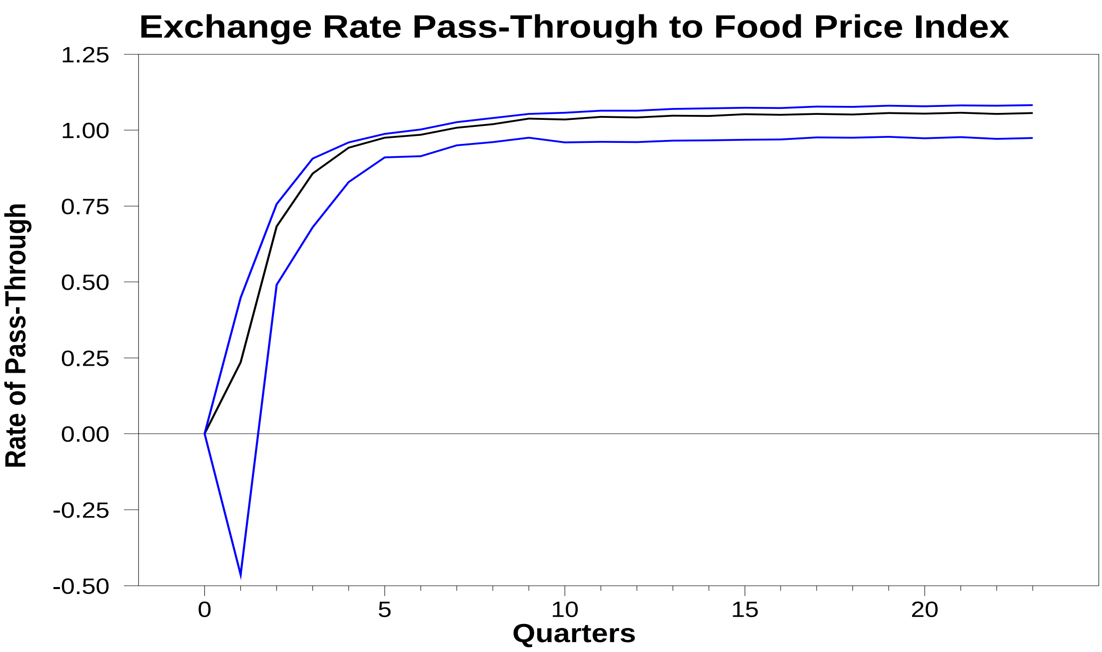
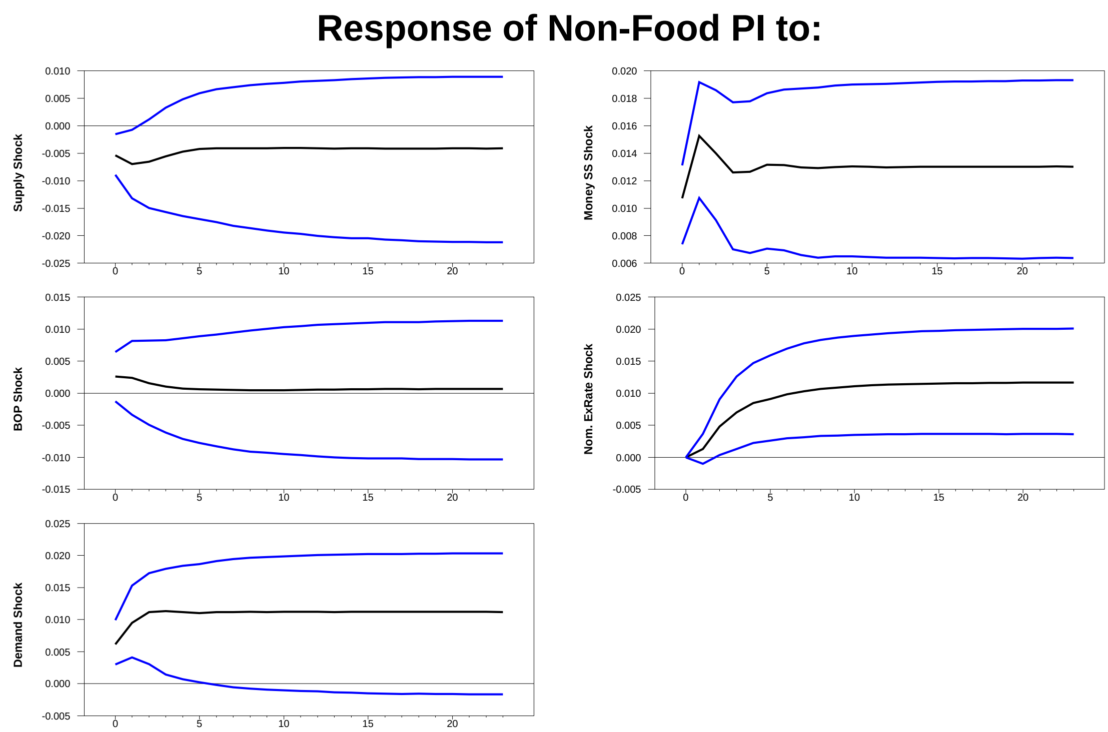


Figure 7: Estimated Rates Exchange Rate Pass-Through to Food Price Index



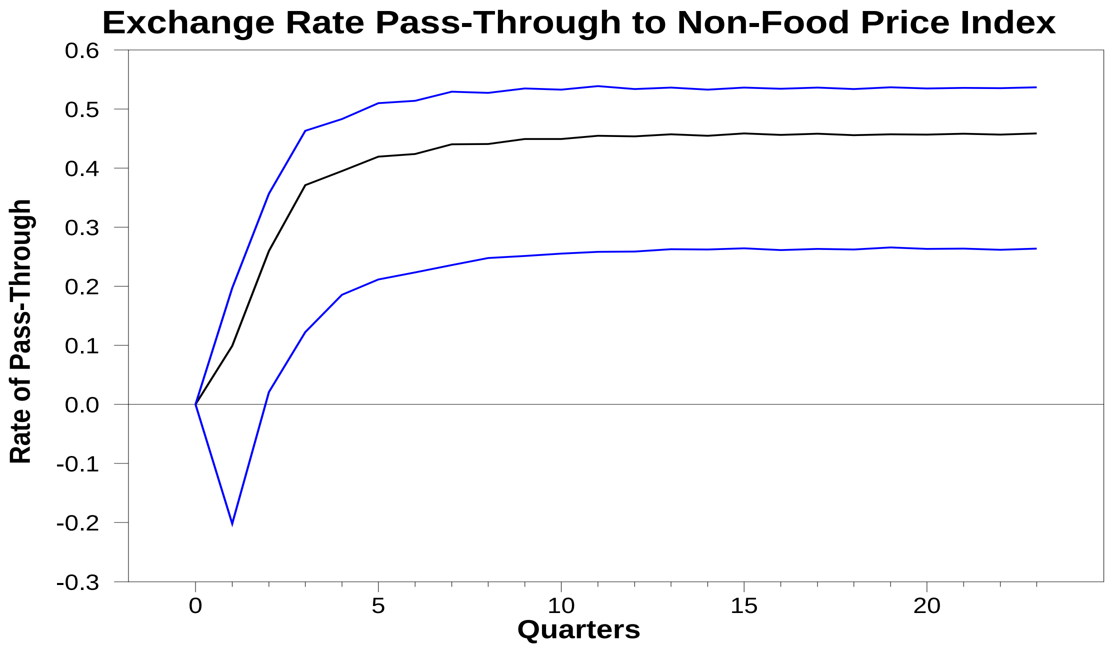
from the second quarter on wards. Compared to the above two cases, the impulse response of non-food prices appears to be a bit lower (see Figure 8). As can also be seen from Figure 9, the estimated exchange rate pass-through rate to non-food prices is found to be relatively lower and it converges to about 47 percent in about 12 quarters after the exchange rate shock. Given that non-food items can be largely imported, the smaller ERPT observed in the case of non-food price index may seem strange. However, a closer look at the composition of the non-food price index indicate that it is composed of non-tradable items (like services) which are less likely to be affected by changes in exchange rate. Besides, the fact that the non-food price index also include items whose prices are highly regulated by the government (e.g. water, electricity and communication) can be another reason for the observed smaller rate of ERPT to non-food price index.

Figure 8: Response of Non-Food Price to Shocks from Nominal Exchange Rate:



Moving to the decomposition of variances for growth rates of food and non-food price indices, we can see from Table 2, in the first quarter, the lion's share of the variation in food inflation is explained by shocks to balance of payments (BOP) which determines about 60 to 62 percent of the variance. However, while the contribution of BOP shocks decline in subsequent quarters, shocks to aggregate supply gained importance at longer horizons. In three to four years, about 18 percent of the variation in food inflation is attributable to shocks to output supply. While shocks to aggregate demand contribute to about 15 (on impact) and 7 percent (in about three years), shocks to nominal exchange rate and money supply account for smaller percentage of the variation in growth rate of food prices. Although money supply shocks explain 9.5 and 5.1 percent of the variance in food inflation in the first and second quarters, its

Figure 9: Estimated Rates Exchange Rate Pass-Through to Non-Food Price Index



contribution declines in subsequent quarters reaching only 2 percent in the third year. After the first quarter, nominal exchange rate shocks have a small but stable contribution to the variation in food prices; in the third year, shocks to nominal exchange rate accounts for about 3.5 percent of the variation in food inflation.

In the case of non-food prices, however, shocks to money supply appear to explain a larger share of the variation in non-food inflation, 53 percent on impact to 26 percent in the third year. Aggregate demand (17 to 19 percent) and aggregate supply shocks (13 to 12 percent) also play important role in determining the variation in the growth rate of non-food prices (see Table 3). Even if nominal exchange rate play little role at shorter horizons, its contribution has gradually increased determining more than 12 percent of the variation in non-food price inflation after 3 years.

Table 2: Decomposition of Variance for Food Price Index

Horizon	Supply Shock	BOP Shock	Demand Shock	Money SS Shock	Nom. ExRate Shock
1	8.4	59.5	15.0	9.5	0.0
	[0.8,28.5]	[35.1,77.3]	[4.9,31.7]	[3.3,19.2]	[0.0, 0.0]
2	14.6	59.4	13.4	5.1	0.3
	[2.3,39.2]	[33.7,78.9]	[4.3,28.3]	[1.6,11.3]	[0.0, 1.1]
3	16.2	59.6	12.2	3.7	1.2
	[2.5,42.7]	[33.9,78.8]	[4.0,26.0]	[1.2, 8.1]	[0.3, 3.2]
4	16.3	60.5	10.7	3.2	1.9
	[2.7,44.0]	[35.1,79.4]	[3.5,23.3]	[1.1, 6.7]	[0.4, 4.7]
8	17.0	61.7	7.7	2.3	3.2
	[2.9,47.5]	[35.4,79.9]	[2.2,18.8]	[0.8, 5.1]	[0.7, 7.8]
12	17.6	61.7	6.5	2.0	3.5
	[2.8,49.5]	[34.9,80.4]	[1.6,17.3]	[0.7, 4.6]	[0.7, 8.8]
16	17.9	61.7	6.0	1.8	3.6
	[2.8,50.9]	[34.5,80.8]	[1.3,16.6]	[0.6, 4.3]	[0.7, 9.2]

Table 3: Decomposition of Variance for Non-Food Price Index

Horizon	Supply Shock	BOP Shock	Demand Shock	Money SS Shock	Nom. ExRate Shock
1	13.4	5.0	16.9	52.6	0.0
	[2.0,34.1]	[0.5,19.1]	[4.3,42.1]	[23.6,76.4]	[0.0, 0.0]
2	12.2	4.5	18.4	51.1	0.5
	[1.9,33.6]	[0.8,17.0]	[4.3,43.9]	[23.6,75.5]	[0.0, 2.2]
3	11.8	4.7	20.6	45.3	2.4
	[2.4,34.0]	[1.1,16.3]	[5.2,46.1]	[20.4,69.8]	[0.3, 8.1]
4	11.7	5.1	21.7	40.2	4.6
	[2.7,34.2]	[1.2,16.4]	[5.5,47.5]	[17.3,64.4]	[0.6,14.5]
8	11.6	5.5	20.8	30.4	10.2
	[2.9,35.4]	[1.3,17.9]	[4.8,46.3]	[10.7,55.1]	[1.4,28.5]
12	11.6	5.6	19.7	26.3	12.4
	[2.7,37.7]	[1.2,19.6]	[4.0,45.8]	[8.2,51.3]	[1.7,33.9]
16	11.5	5.7	18.8	24.2	13.4
	[2.4,39.8]	[1.2,20.5]	[3.5,45.7]	[6.6,49.5]	[1.8,36.4]

6 Conclusion

In this paper, we examine exchange rate pass-through to domestic inflation with a particular focus on the Ethiopian economy. Despite the vast literature on exchange rate pass-through, the existing literature for the most part focuses on developed countries and the evidence for developing countries, particularly for those in SSA, is quite limited. This paper thus aims to contribute to the limited evidence regarding the impact of exchange rate shocks to prices in sub Saharan Africa (SSA) by considering the case of Ethiopia. To this end, we employ a Structural Vector Autoregressive (SVAR) model where identification of the structural shocks is achieved using a combination of short run and long-run restrictions, where the latter is derived from a simple small open economy macro model.

In the theory model, which we use as a guideline in specifying the long-run restrictions needed for identification, effort is made to reflect the existing realities of SSA countries in general and Ethiopia in particular. Specifically, we modified this standard open macroeconomic model to take into account the fact that the Ethiopian economy has a closed capital account (has limited/no private capital inflows) and is heavily dependent on external financial flows (foreign aid and remittances) to finance its norm trade deficit.

The results from estimated cumulative structural impulse response functions are starkly consistent with theoretical predictions giving us confidence in the model and identification strategy we have employed in this paper. The impulse response of domestic prices to a one standard deviation nominal exchange rate shock (depreciation/devaluation) is found to be positive and statistically significant, that leaves domestic prices at a permanently higher level in about two years. The corresponding rate of exchange rate pass-through to aggregate CPI, starts out low but quickly increases to more than 80 percent in about four quarters and approaches to 96 percent in about two years period. While impulse responses of food prices follow the same pattern as aggregate CPI, exchange rate pass-through to non-food prices is found to be relatively lower, approaching to 47 percent. The higher exchange rate pass-through to domestic price that we found for Ethiopia is consistent with the the predictions of producer currency pricing models. Given the potentially low importance of the Ethiopian economy to foreign exporters, this is not a very difficult hypothesis to accept. Moreover, the fact that the Ethiopian economy is characterized by persistent exchange rate, high import dependence and inflationary environment can also explain the large exchange rate pass-through estimate we find in this paper.

Despite of the large exchange rate pass-through estimate, shocks to nominal exchange rate are found to account for only about 7 to 8 percent of the variation in inflation in three to four years. In longer horizons, a larger proportion of the variation in domestic inflation is instead found to be explained by shocks to BOP, aggregate supply and aggregate demand, in their respective order. On the other hand, shocks to money supply is found to play a very limited role in explaining the variation in the general inflation, which is contrary to our expectation. While this remains to be true in the case of variance decompositions of food inflation, the variance decomposition of non-food inflation reveals a totally different picture. In the case of non-food prices, shocks to money supply explain a larger percentage of the variation in non-food inflation in Ethiopia.

Overall the results have their own implication for effectiveness of monetary policy. In particular, the high rate of ERPT found in Ethiopia shows that impacts of changes in nominal exchange rate (devaluation measures) on real variables will be lower than anticipated. In view of this, the devaluation measures aimed at improving (maintaining) the country's export competitiveness may not be able to achieve the desired targets since such measures will be largely matched by a rise in the domestic price level. This is also reflected in the actual overtime trends of CPI, nominal and real exchange rate depicted in Figure 17.

Therefore, even if devaluation measures are sometimes unavoidable, particularly in order to maintain competitiveness in the face of inflationary pressures or excess demand in foreign currency, taking such measures per se cannot be successful in achieving the desired target unless accompanied by other macroeconomic measures that can put inflation in check.

Appendices

Appendix A: *Components of Aggregate Demand*

Below we define components of the aggregate demand equation in order to show how the two shocks (shocks associated with trade balance and external financial flows) are introduced in the standard IS equation:

$$C_t = C_{0,t} + \sigma Y_t + F_t^C \tag{32}$$

$$I_t = I_{0,t} - \rho R_t + F_t^I \tag{33}$$

$$G_t = G_{0,t} + F_t^G \tag{34}$$

$$EX_t = EX_{0,t} + \eta_x Q_t \tag{35}$$

$$IM_t = IM_{0,t} - \eta_m Q_t + \theta Y_t \tag{36}$$

Defining autonomous net exports as $X_{0,t} = EX_{0,t} - IM_{0,t}$ and $\eta_x + \eta_m = \eta$, we can define the current account as:

$$CA_t = \underbrace{\eta Q_t - \theta Y_t^d + X_{0,t}}_{\text{Net exports}=NX} + F_t = \Delta Res_t \tag{37}$$

Where: $F_t = F_t^C + F_t^I + F_t^G$ and the respective components F_t^C, F_t^I, F_t^G show the amount (in Ethiopian Birr) of external financial flows used to finance consumption, investment and government expenditure. Moreover, σ and ρ in Equation 32 and 33 respectively refer to propensity to consume out of disposable income and response of investment to changes in the real interest rate R. On the other hand, η and θ in Equation 37 capture the response of trade balance to movements in the real exchange rate (Q_t) and propensity to import from disposable income, respectively.[19] Lastly, $X_{0,t}$ and F_t in equation 37, respectively, refer to the autonomous level of

[19]The real exchange rate is given by $Q = \frac{S_t P_t^*}{P_t}$ where S_t is nominal exchange rate defined here as the price of the foreign currency (USD) in terms of the local currency (Birr) and a rise in S_t implies depreciation/devaluation of the Ethiopian Birr; P_t is the domestic price level and P_t^* is the foreign price level and is normalized to one.

net exports (NX) and exogenous financial flows needed to finance trade deficit. Since Ethiopia has a closed capital account, the trade deficit is by and large financed through foreign currency earnings from foreign aid and remittances. In the short to medium run, this will be equal to a change in the country's foreign reserves ΔRes_t. That is, when foreign financial flows F_t exceed the country's trade deficit there will be a positive change in foreign reserves and vice versa.

Given its components listed above, the aggregate demand equation, assuming a good's market equilibrium ($Y_t^d = Y_t^s = Y_t$) can be given as:

$$Y_t^d = C_{0,t} + \sigma Y_t^d + I_{0,t} - \rho R_t + G_{0,t} + F_t + \underbrace{\eta Q_t - \theta Y_t^d + X_{0,t}}_{NX = \Delta Res_t - F_t} \tag{38}$$

$$Y_t^d = \frac{1}{1-\sigma+\theta}[\underbrace{C_{0,t} + I_{0,t} + G_{0,t}}_{D_t} + X_{0,t} + F_t - \rho R_t + \eta Q] \tag{39}$$

Appendix B: *Long run equilibrium*

Long run Price Level

$$p_t = \frac{\lambda}{\rho + \rho\lambda + \eta\lambda}\left[d_t + x_t + f_t - \frac{\lambda + \gamma\rho}{\gamma\lambda} y_t + \frac{\rho + \eta\lambda}{\lambda} m_t + \eta e_t - \rho \mathbb{E}[p_{t+1}]\right] \tag{40}$$

$$p_t = \frac{\lambda(d_t + x_t + f_t)}{\rho + \rho\lambda + \eta\lambda} - \frac{(\lambda + \gamma\rho)y_t}{\rho\gamma + \rho\lambda\gamma + \eta\lambda\gamma} + \frac{(\rho + \eta\lambda)m_t}{\rho + \rho\lambda + \eta\lambda} + \frac{\lambda\eta(e_t)}{\rho + \rho\lambda + \eta\lambda} - \frac{\lambda\rho\mathbb{E}[p_{t+1}]}{\rho + \rho\lambda + \eta\lambda} \tag{41}$$

Forwarding the first order difference equation in price given in Equation 41 one period, we get:

$$\begin{aligned} p_{t+1} &= \frac{\lambda(d_{t+1} + x_{0,t+1} + f_{t+1})}{\rho + \rho\lambda + \eta\lambda} - \frac{(\lambda + \gamma\rho)y_{t+1}}{\rho\gamma + \rho\lambda\gamma + \eta\lambda\gamma} + \frac{(\rho + \eta\lambda)m_{t+1}}{\rho + \rho\lambda + \eta\lambda} \\ &+ \frac{\lambda\eta(e_{t+1})}{\rho + \rho\lambda + \eta\lambda} - \frac{\lambda\rho\mathbb{E}[p_{t+2}]}{\rho + \rho\lambda + \eta\lambda} \end{aligned} \tag{42}$$

Substituting p_{t+1} back into p_t, we get:

$$
\begin{aligned}
p_t \;=\;& \frac{\lambda\left[d_t + (1-\eta)(x_t + f_t) + \eta \Delta res_t\right]}{\rho(1+\lambda)} + \frac{m_t}{1+\lambda} - y_t \left(\frac{\lambda + \gamma\rho - \lambda\gamma\eta\theta}{\gamma\rho(1+\lambda}\right) \\
+\;& \mathbb{E}\left[\frac{\lambda^2\left[d_{t+1} + (1-\eta)(x_{0,t+1} + f_{t+1}) + \eta\Delta res_{t+1}\right]}{\rho(1+\lambda)^2}\right] + \frac{\lambda \mathbb{E} m_{t+1}}{(1+\lambda)^2} \\
-\;& \lambda\left(\frac{\lambda + \gamma\rho - \lambda\gamma\eta\theta}{\gamma\rho(1+\lambda)^2}\right)\mathbb{E}y_{t+1} + \left(\frac{\lambda}{1+\lambda}\right)^2 \mathbb{E}[p_{t+2}] \qquad (43)
\end{aligned}
$$

$$
\begin{aligned}
p_t \;=\;& \frac{\lambda(d_t + x_t + f_t)}{\rho + \rho\lambda + \eta\lambda} - \frac{(\lambda + \gamma\rho)y_t}{\rho\gamma + \rho\lambda\gamma + \eta\lambda\gamma} + \frac{(\rho + \eta\lambda)m_t}{\rho + \rho\lambda + \eta\lambda} + \frac{\lambda\eta(e_t)}{\rho + \rho\lambda + \eta\lambda} \\
-\;& \frac{\lambda\rho}{\rho + \rho\lambda + \eta\lambda}\left[\frac{\lambda(d_{t+1} + x_{0,t+1} + f_{t+1})}{\rho + \rho\lambda + \eta\lambda} - \frac{(\lambda + \gamma\rho)y_{t+1}}{\rho\gamma + \rho\lambda\gamma + \eta\lambda\gamma} + \frac{(\rho + \eta\lambda)m_{t+1}}{\rho + \rho\lambda + \eta\lambda}\right. \\
+\;& \left.\frac{\lambda\eta(e_{t+1})}{\rho + \rho\lambda + \eta\lambda} - \frac{\lambda\rho\mathbb{E}[p_{t+2}]}{\rho + \rho\lambda + \eta\lambda}\right] \qquad (44)
\end{aligned}
$$

$$
\begin{aligned}
p_t \;=\;& \frac{\lambda}{\rho + \rho\lambda + \eta\lambda}\left[d_t + x_t + f_t + \frac{\lambda\rho}{\rho + \rho\lambda + \eta\lambda}(d_{t+1} + x_{0,t+1} + f_{t+1})\right] \qquad (45) \\
-\;& \frac{\lambda + \gamma\rho}{\rho\gamma + \rho\lambda\gamma + \eta\lambda\gamma}\left[y_t + \frac{\lambda\rho}{\rho + \rho\lambda + \eta\lambda}y_{t+1}\right] + \frac{\rho + \eta\lambda}{\rho + \rho\lambda + \eta\lambda}\left[m_t + \frac{\lambda\rho}{\rho + \rho\lambda + \eta\lambda}m_{t+1}\right] \\
+\;& \frac{\lambda\eta}{\rho + \rho\lambda + \eta\lambda}\left[e_t + \frac{\lambda\rho}{\rho + \rho\lambda + \eta\lambda}e_{t+1}\right] + \left(\frac{\lambda\rho}{\rho + \rho\lambda + \eta\lambda}\right)^2 \mathbb{E}[p_{t+2}] \qquad (46)
\end{aligned}
$$

A repeated iteration of this can be summarized as:

$$
\begin{aligned}
p_t \;=\;& \frac{\lambda}{\rho + \rho\lambda + \eta\lambda}\lim_{T\to\infty}\sum_{k=t}^{T-1}\left[\left(\frac{\lambda\rho}{\rho + \rho\lambda + \eta\lambda}\right)^{k-t}\mathbb{E}[d_k + x_{0,k} + f_k]\right] \\
-\;& \frac{\lambda + \gamma\rho}{\rho\gamma + \rho\lambda\gamma + \eta\lambda\gamma}\lim_{T\to\infty}\sum_{k=t}^{T-1}\left[\left(\frac{\lambda\rho}{\rho + \rho\lambda + \eta\lambda}\right)^{k-t}\mathbb{E}y_k\right] \\
+\;& \frac{\rho + \eta\lambda}{\rho + \rho\lambda + \eta\lambda}\lim_{T\to\infty}\sum_{k=t}^{T-1}\left[\left(\frac{\lambda\rho}{\rho + \rho\lambda + \eta\lambda}\right)^{k-t}\mathbb{E}m_k\right] + \frac{\lambda\eta}{\rho + \rho\lambda + \eta\lambda} \\
& \lim_{T\to\infty}\sum_{k=t}^{T-1}\left[\left(\frac{\lambda\rho}{\rho + \rho\lambda + \eta\lambda}\right)^{k-t}\mathbb{E}e_k\right] + \lim_{T\to\infty}\sum_{k=t+1}^{T}\left[\left(\frac{\lambda\rho}{\rho + \rho\lambda + \eta\lambda}\right)^{k-t}\mathbb{E}p_k\right] \qquad (47)
\end{aligned}
$$

A bounded solution requires that:

$$\lim_{T\to\infty}\sum_{k=t+1}^{T}\left[\left(\frac{\lambda\rho}{\rho+\rho\lambda+\eta\lambda}\right)^{k-t}\mathbb{E}p_k\right]=0 \quad \text{and} \tag{48}$$

$$\begin{aligned}
&\left|\frac{\lambda}{\rho+\rho\lambda+\eta\lambda}\lim_{T\to\infty}\sum_{k=t}^{T-1}\left(\frac{\lambda\rho}{\rho+\rho\lambda+\eta\lambda}\right)^{k-t}\mathbb{E}[d_k+x_{0,k}+f_k]\right.\\
&-\frac{\lambda+\gamma\rho}{\rho\gamma+\rho\lambda\gamma+\eta\lambda\gamma}\lim_{T\to\infty}\sum_{k=t}^{T-1}\left(\frac{\lambda\rho}{\rho+\rho\lambda+\eta\lambda}\right)^{k-t}\mathbb{E}y_k+\frac{\rho+\eta\lambda}{\rho+\rho\lambda+\eta\lambda}\\
&\lim_{T\to\infty}\sum_{k=t}^{T-1}\left(\frac{\lambda\rho}{\rho+\rho\lambda+\eta\lambda}\right)^{k-t}\mathbb{E}m_k+\frac{\lambda\eta}{\rho+\rho\lambda+\eta\lambda}\\
&\left.\lim_{T\to\infty}\sum_{k=t}^{T-1}\left(\frac{\lambda\rho}{\rho+\rho\lambda+\eta\lambda}\right)^{k-t}\mathbb{E}e_k\right|<\infty
\end{aligned} \tag{49}$$

Both of the above conditions will be satisfied given that $\left(\frac{\lambda\rho}{\rho+\rho\lambda+\eta\lambda}\right)<1$. Or, more specifically, unless the absolute value of the log of price level grows at an exponential rate of $\left(\frac{\rho+\rho\lambda+\eta\lambda}{\lambda\rho}\right)$ or above. And convergence of the second term requires that the sequences in $\{d_t,\ x_t,\ f_t,\ m_t,\ y_t,\ e_t\}$ do not grow at an exponential rate of $\left(\frac{\rho+\rho\lambda+\eta\lambda}{\lambda\rho}\right)$ or above.

Letting $v_t=\{d_t,\ x_t,\ f_t,\ m_t,\ y_t,\ e_t\}$

$$\begin{aligned}
v_t &= v_{t-1}+\mu_t^v\\
v_{t+1} &= v_t+\mu_{t+1}^v\\
v_{t+2} &= v_{t+1}+\mu_{t+2}^v\\
v_{t+2} &= v_t+\mu_{t+1}^v+\mu_{t+2}^v\\
etc...\ and &\\
\mathbb{E}(\mu_{t+i}^v) &= 0 \quad \text{for } i\geq 1\\
\lim_{T\to\infty}\sum_{k=t}^{T-1}\left(\frac{\lambda\rho}{\rho+\rho\lambda+\eta\lambda}\right)^{k-t} &= \frac{\rho+\rho\lambda+\eta\lambda}{\rho+\eta\lambda}
\end{aligned}$$

The long-run equilibrium price level can therefore be given by:

$$p_t=\frac{\lambda}{\rho+\eta\lambda}[d_t+x_t+f_t]-\frac{\lambda+\gamma\rho}{\rho+\eta\lambda}y_t+m_t+\frac{\lambda\eta}{\rho+\eta\lambda}e_t \tag{50}$$

Appendix C: *Variables in Log Levels and Differences*

Figure 10: Trends in Log GDP: In Log Levels and Differences

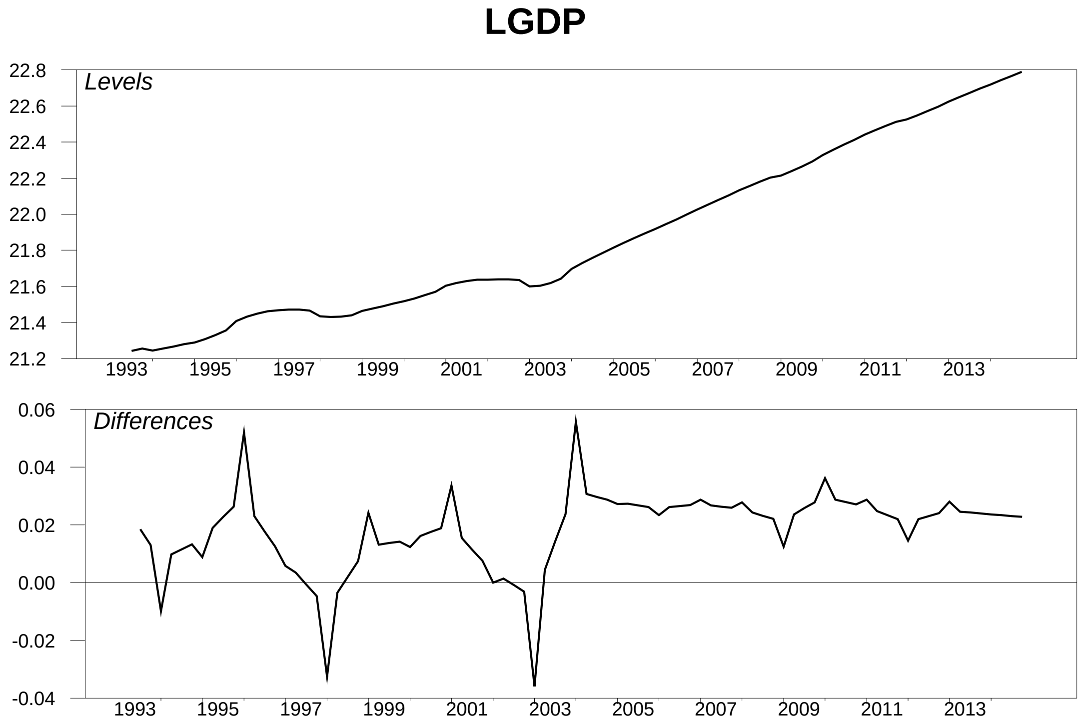


Figure 11: Trends in Log Real Exchange Rate: In Log Levels and Differences

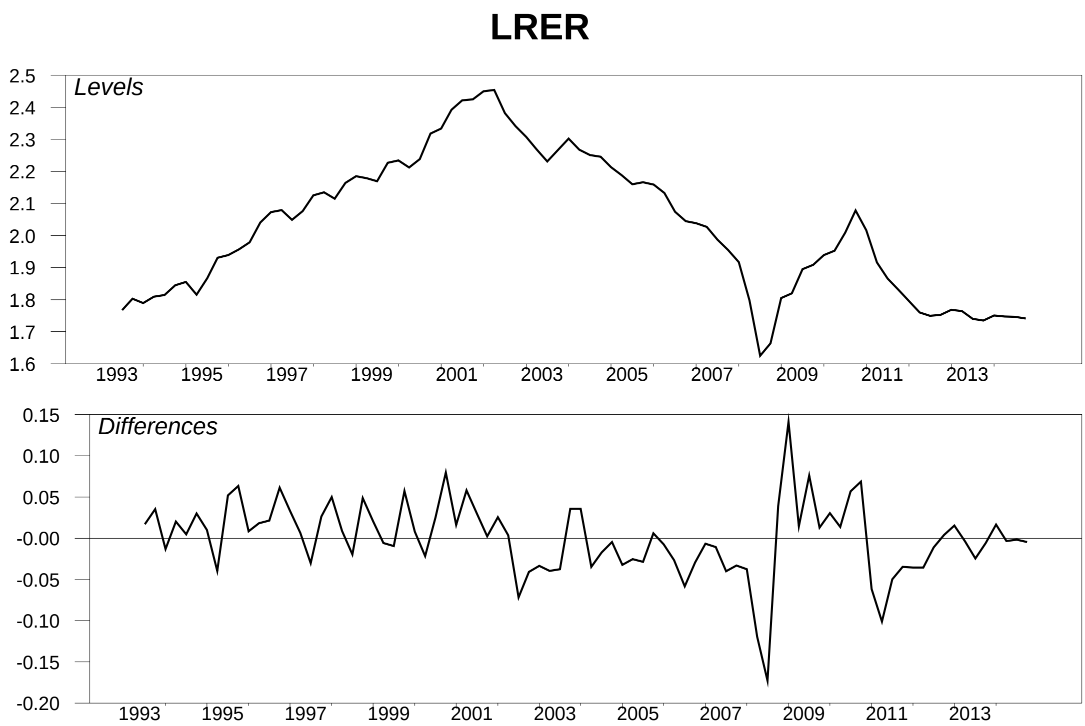


Figure 12: Consumer Price: In Log Levels and Differences

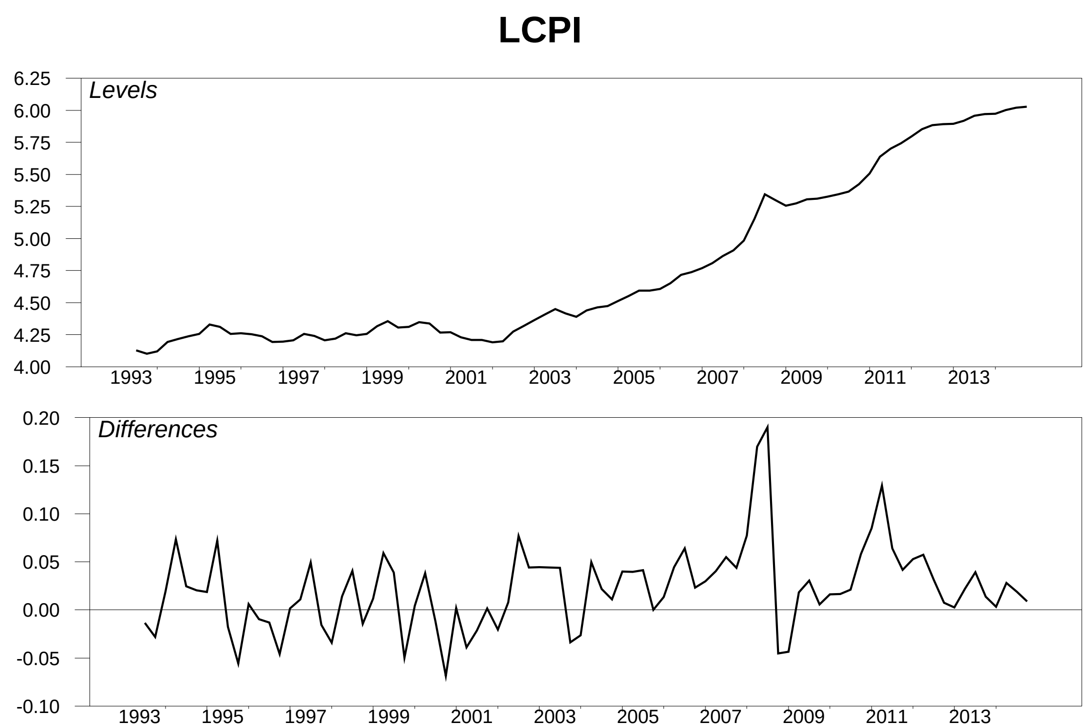


Figure 13: Trends in Log Nominal Exchange Rate: In Log Levels and Differences

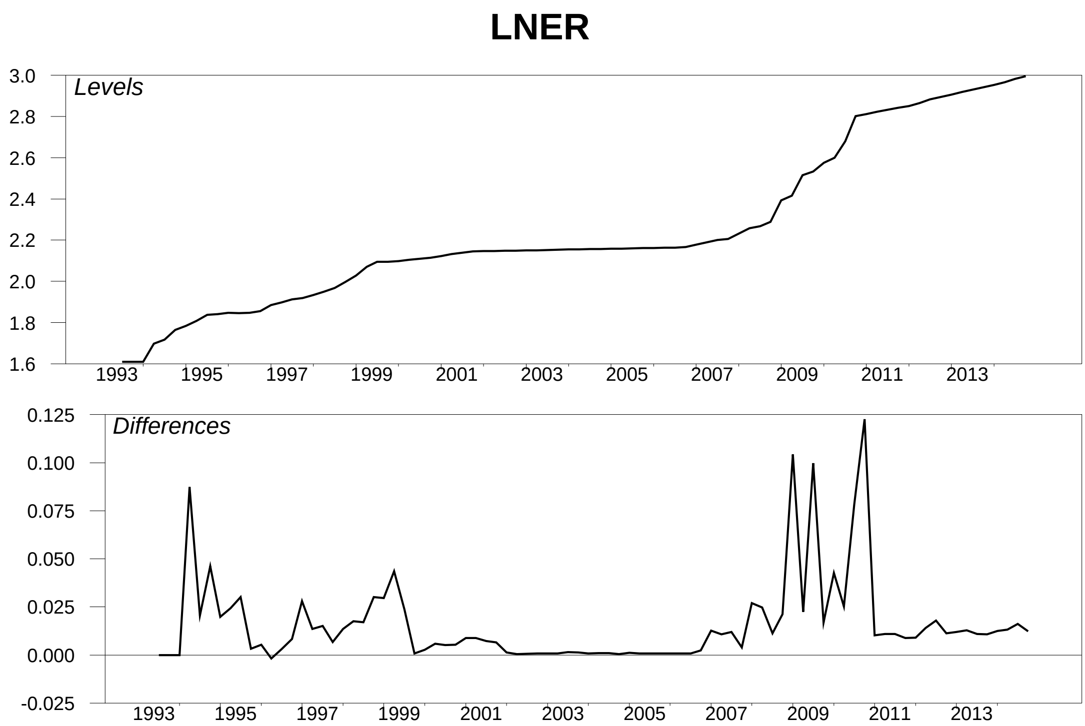


Figure 14: Trends in Log Real Money Demand: In Log Levels and Differences

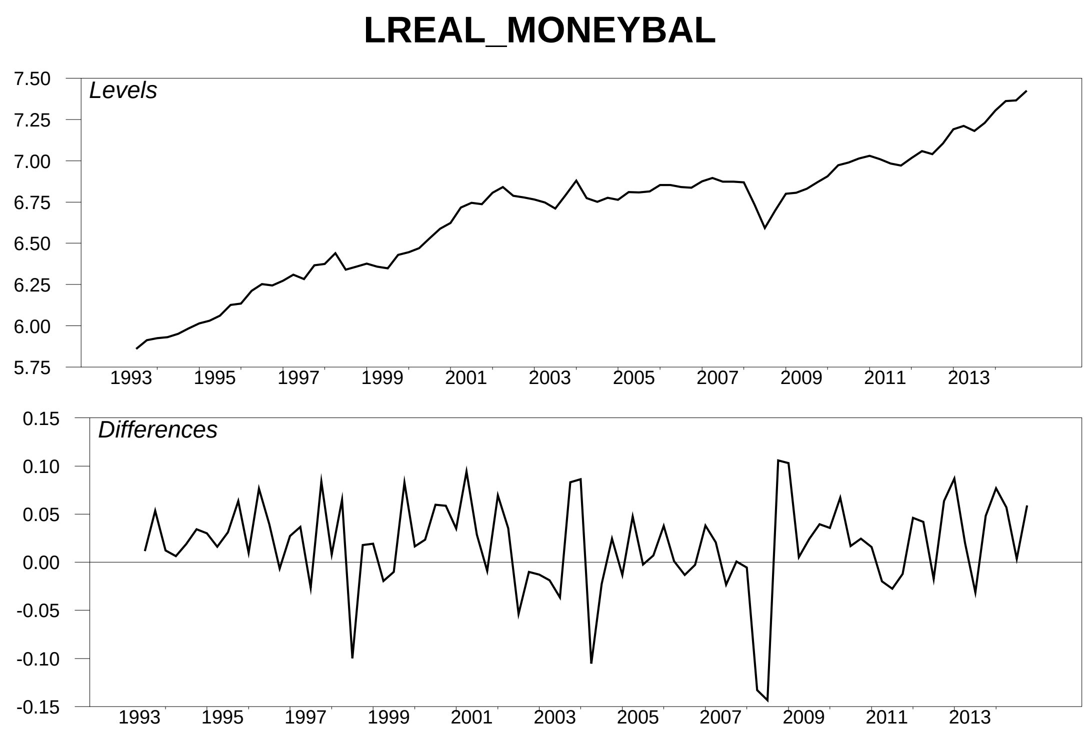


Table 4: Summary of Unit Root Tests: ADF, Philip Perron and KPSS

	ADF				Philip Perron				KPSS		
Variables	Obs	Lags	Z(t)	p_val	Obs	Lags	Z(t)	p_val	Obs	Lags	Test_Stat
Log of Consumer Prices	82	5	1.355	0.997	87	3	1.619	0.998	88	6	1.198
Log Nominal Exchange Rate	83	4	-0.388	0.912	87	3	0.456	0.983	88	6	1.218
Log Output	84	3	-1.392	0.863	87	3	-0.827	0.963	88	6	0.443
Log Real Exchange Rate	82	5	-1.242	0.655	87	3	-1.165	0.689	88	6	0.326
Log Real Money Demand	81	5	-0.832	0.810	86	3	-0.827	0.811	88	6	0.399
Log Coffee Price	82	5	-1.312	0.624	87	3	-1.922	0.322	87	6	1.245
Log Oil Price	83	4	-1.205	0.671	87	3	-1.098	0.716	88	6	1.271

KPSS Crit. values for H0: X=level stationary: 10%: 0.347,5%: 0.463, 1%: 0.739
KPSS Crit. values for H0: X=trend stationary: 10%: 0.119,5%: 0.146, 1%: 0.216

Table 5: Summary of Unit Root Tests in Differences: ADF, Philip Perron and KPSS

	ADF				Philip Perron				KPSS		
Variables	Obs	Lags	Z(t)	p_val	Obs	Lags	Z(t)	p_val	Obs	Lags	Test_Stat
Diff. Log Consumer Prices	82	4	-3.561	0.007	86	3	-5.591	0.000	87	5	0.520
Diff. Log Nominal Exchange Rate	83	3	-3.150	0.023	86	3	-6.662	0.000	87	6	0.174
Diff. Output	81	5	-2.887	0.047	86	3	-4.159	0.001	87	6	0.066
Diff. Real Exchange Rate	82	4	-3.575	0.006	86	3	-5.798	0.000	87	5	0.364
Diff. Real Money Demand	81	4	-3.972	0.002	85	3	-8.031	0.000	86	11	0.143
Diff. Coffee Price	82	4	-5.120	0.000	86	3	-7.088	0.000	87	3	0.073
Diff. Oil Price	83	3	-5.341	0.000	86	3	-6.597	0.000	87	9	0.116

KPSS Crit. values for H0: X=level stationary: 10%: 0.347,5%: 0.463, 1%: 0.739
KPSS Crit. values for H0: X=trend stationary: 10%: 0.119,5%: 0.146, 1%: 0.216

h!]

Table 6: Clemente-Montañés-Reyes unit-root test with double mean shifts

	CLEMAO2					CLEMIO2			
Variables	Obs	Lags	t-stat (Br1)	t-stat (Br2)	t-stat$(\rho-1)$	Lags	t-stat (Br1)	t-stat (Br2)	t-stat$(\rho-1)$
Log of Consumer Prices	80	2	15.345[a]	3.776[a]	-2.833	3	3.512[a]	2.330[b]	-2.941
Log of Nominal Exchange Rate	80	4	12.025[a]	20.271[a]	-3.550	12	1.631	5.751[a]	-4.660
Log of Output	80	0	15.258[a]	9.988[a]	-2.822	12	3.666[a]	1.044	-0.762
Log of Real Exchange Rate	80	12	6.855[a]	-8.848[a]	-2.384	1	1.922[c]	-3.655[a]	-3.698
Log of Real Money Demand	79	0	16.762[a]	7.662[a]	-3.170	0	2.168[b]	3.569[a]	-3.298
Log of Coffee Price	80	1	-6.153[a]	9.554[a]	-4.493	9	-3.033[a]	3.761[a]	-2.931
Log of Oil Price	80	1	7.942[a]	16.139[a]	-5.270	1	3.333[a]	3.800[a]	-4.825

Note: a, b and c indicate statistical significance at 1%, 5%, and 10% level, respectively.
t-stat (Br1) and t-stat (Br2) are t-statistics values for the two break points.
t-stat $(\rho - 1)$ is the test statistics for the unit root test $(\rho - 1)$, 5 % critical value= -5.490

Table 7: Johansen test for co-integration

Rank	Trace Stat.	5% Crit.	1% Crit.
0	124.89	87.31	96.58
1	**54.26**	**62.99**	**70.05**
2	18.31	42.44	48.45
3	10.42	25.32	30.45
4	3.60	12.25	16.26

Deterministic terms: Restricted Trend

Table 8: Test for Autocorrelation (Lagrangian Multiplier Test)

lag	chi2	df	$Prob > chi2$
1	33.1260	25	0.12795
2	32.7648	25	0.13705
3	26.3637	25	0.38840
4	29.5620	25	0.24111
5	35.5554	25	0.07862
6	29.6241	25	0.23867

H0: no autocorrelation at lag orders

Figure 15: VAR Stability Test

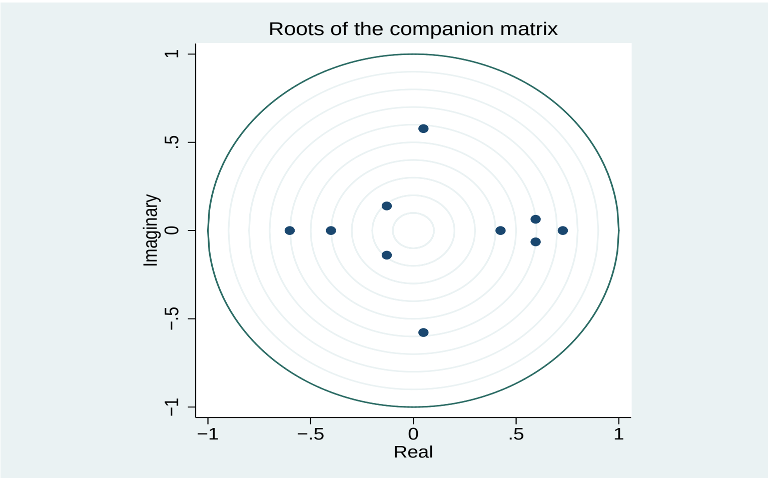


Table 9: Tests for Normality

	Jarque-Bera test			Skewness test				Kurtosis test			
Equation	chi2	df	$Prob > chi2$	Skewness	chi2	df	$Prob > chi2$	Kurtosis	chi2	df	$Prob > chi2$
Diff LnGDP	34.769	2	0.000	-.41633	2.427	1	0.11929	6.0398	32.342	1	0.000
Diff LnRER	2.807	2	0.24572	.18167	0.462	1	0.49666	2.1815	2.345	1	0.12568
Diff LnReal Money Bal.	4.834	2	0.08920	-.29024	1.179	1	0.27749	1.9782	3.654	1	0.05592
Diff LnNER	20.769	2	0.00003	.91339	11.680	1	0.00063	4.6114	9.089	1	0.00257
Diff LnCPI	0.429	2	0.80701	-.15932	0.355	1	0.55108	2.8551	0.073	1	0.78637
All	63.607	10	0.000		16.103	5	0.00655		47.504	5	0.000

Figure 16: Trends in Aggregate CPI, Food and Non-Food Price Indices

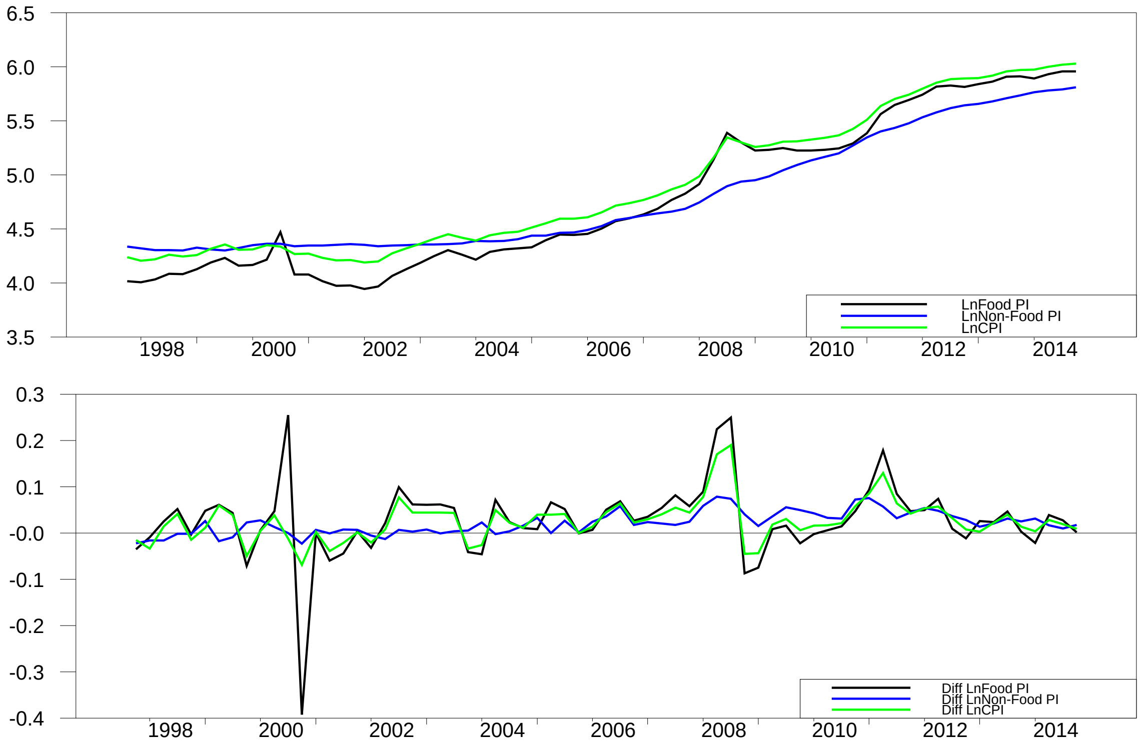


Table 10: Major Groups in the 2006 and 2011 based CPI and their Weights at Country Level

Items	Weights	
	December 2006=100	December 2011=100
1. Food and Non-Alcoholic Beverage	0.57	0.53
2. Non-Food	0.43	0.47
Alcoholic Beverages and Tobacco, Narcotics	2.5	4.85
Clothing and Footwear	8.32	6.62
Housing, Water, Electricity, Gas and Other Fuels	20.56	16.34
Furnishing, Housing Equipment and maintenance of the house	3.75	5.41
Health	1.11	1.08
Transport	2.5	2.8
Communication	-	1.08
Recreation and Culture	1.1	0.6
Education	-	0.45
Restaurant and Hotels	-	5.46
Miscellaneous Goods and Services	3.2	2.56
Total	100	100

Figure 17: Trends in CPI as well as Nominal and Real Exchange Rates in Ethiopia: 1993Q1-2014Q4

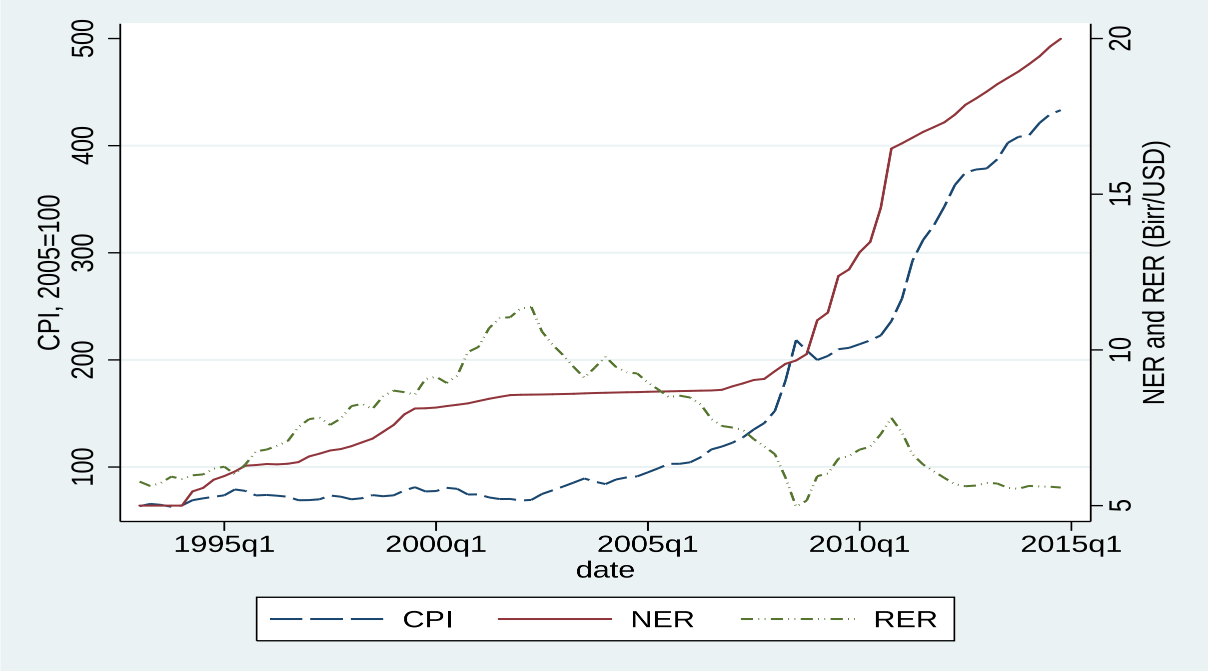
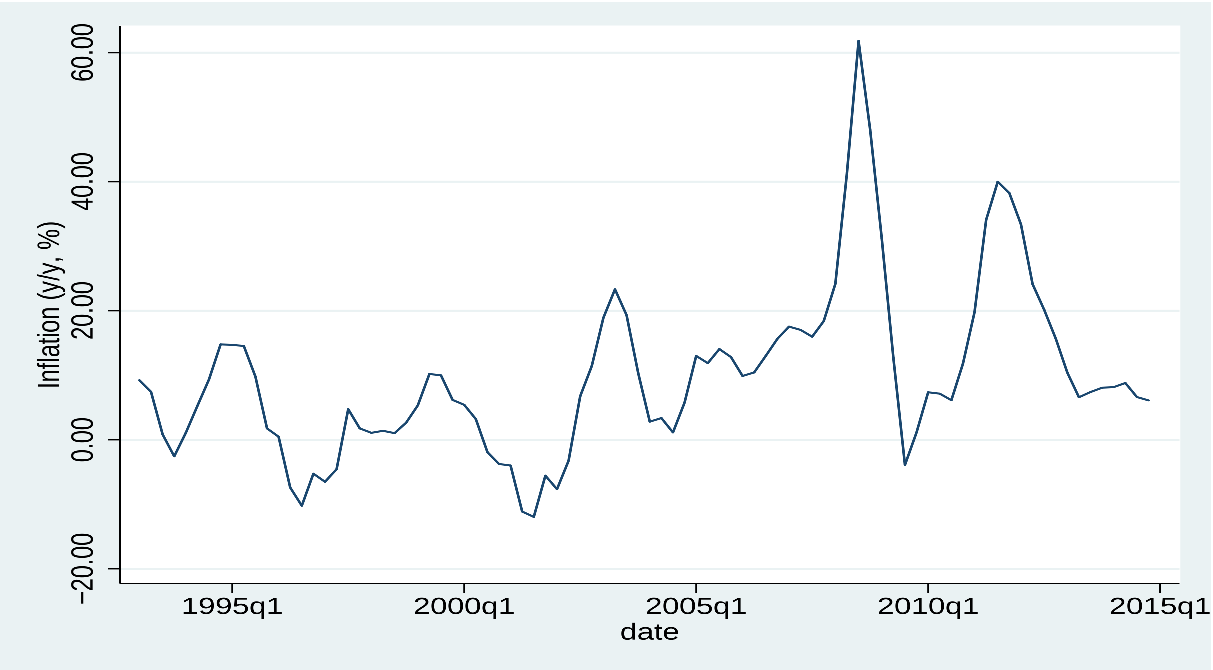


Figure 18: Trends in Inflation in Ethiopia (Percent change, corresponding period previous year): 1993Q1-2014Q4



Appendix D: *Variance Decompositions*

Table 11: Decomposition of Variance for Aggregate Output Supply

Horizon	Supply Shock	BOP Shock	Demand Shock	Money SS Shock	Nom. ExRate Shock
1	79.0	3.2	5.0	2.9	2.1
	[62.4,88.9]	[0.3,13.7]	[0.5,16.8]	[0.3, 8.7]	[0.2, 9.0]
2	83.9	2.4	4.9	1.5	1.7
	[69.1,92.3]	[0.4,10.7]	[0.8,14.9]	[0.5, 4.3]	[0.4, 6.7]
3	87.5	1.9	3.9	1.5	1.3
	[74.8,94.3]	[0.3, 8.6]	[0.7,11.7]	[0.7, 3.1]	[0.3, 5.0]
4	90.2	1.6	3.0	1.2	1.1
	[79.6,95.7]	[0.3, 6.9]	[0.6, 9.1]	[0.5, 2.4]	[0.3, 4.0]
8	95.6	0.8	1.4	0.5	0.6
	[90.4,98.1]	[0.2, 3.3]	[0.3, 4.0]	[0.2, 1.1]	[0.2, 1.9]
12	97.3	0.5	0.8	0.3	0.4
	[94.2,98.8]	[0.1, 2.0]	[0.2, 2.4]	[0.1, 0.7]	[0.1, 1.2]
16	98.1	0.3	0.6	0.2	0.3
	[96.0,99.1]	[0.1, 1.4]	[0.1, 1.7]	[0.1, 0.5]	[0.1, 0.9]

Table 12: Decomposition of Variance for Real Exchange Rate

Horizon	Supply Shock	BOP Shock	Demand Shock	Money SS Shock	Nom. ExRate Shock
1	22.8	46.6	2.2	9.8	11.3
	[9.4,38.8]	[26.4,65.3]	[0.2, 9.4]	[4.0,18.3]	[1.5,27.4]
2	22.5	57.1	2.3	5.2	7.0
	[8.1,40.7]	[36.1,74.7]	[0.6, 8.5]	[1.9,11.0]	[1.2,17.2]
3	22.8	64.1	1.7	2.9	4.3
	[7.4,42.4]	[42.5,80.6]	[0.5, 6.2]	[1.2, 5.9]	[0.9,11.9]
4	22.6	67.8	1.3	2.0	3.0
	[7.0,43.2]	[45.7,84.0]	[0.4, 4.8]	[0.9, 4.0]	[0.7, 8.7]
8	25.0	70.9	0.6	0.8	1.3
	[7.4,47.9]	[47.5,88.4]	[0.2, 2.2]	[0.3, 1.5]	[0.3, 3.6]
12	26.2	71.4	0.4	0.4	0.7
	[7.6,50.4]	[47.1,89.8]	[0.1, 1.3]	[0.2, 0.9]	[0.2, 2.1]
16	26.8	71.5	0.3	0.3	0.5
	[7.8,51.8]	[46.6,90.4]	[0.1, 0.9]	[0.1, 0.7]	[0.1, 1.5]

Table 13: Decomposition of Variance for Demand for Real Money Balances

Horizon	Supply Shock	BOP Shock	Demand Shock	Money SS Shock	Nom. ExRate Shock
1	6.6	41.4	4.3	3.3	36.9
	[0.7,23.1]	[21.9,61.5]	[0.4,17.6]	[0.5, 9.0]	[12.9,50.8]
2	12.8	42.2	4.0	2.0	31.5
	[2.5,32.3]	[22.8,61.9]	[0.6,15.8]	[0.5, 5.5]	[10.8,45.6]
3	14.0	43.3	3.7	1.2	30.0
	[2.6,35.1]	[24.0,62.5]	[0.6,15.2]	[0.4, 3.0]	[11.3,43.7]
4	16.3	43.0	3.6	0.8	28.2
	[3.2,38.7]	[23.9,62.5]	[0.6,14.4]	[0.3, 2.0]	[10.7,42.5]
8	20.8	41.5	3.3	0.3	25.0
	[4.6,45.4]	[22.5,62.2]	[0.5,13.4]	[0.1, 0.8]	[10.3,40.4]
12	22.6	40.6	3.1	0.2	23.9
	[5.1,48.2]	[21.7,61.7]	[0.5,13.2]	[0.1, 0.5]	[9.8,39.9]
16	23.4	40.2	3.0	0.1	23.3
	[5.4,49.4]	[21.2,61.6]	[0.5,13.3]	[0.0, 0.3]	[9.5,39.7]

Table 14: Decomposition of Variance for Nominal Exchange Rate

Horizon	Supply Shock	BOP Shock	Demand Shock	Money SS Shock	Nom. ExRate Shock
1	11.5	2.4	50.3	0.0	25.1
	[1.8,29.4]	[0.2,10.2]	[18.3,75.5]	[0.0, 0.0]	[3.2,63.5]
2	9.9	4.8	43.6	2.2	26.9
	[1.8,27.0]	[1.2,15.3]	[14.6,68.4]	[0.7, 4.7]	[4.0,65.2]
3	8.9	5.1	43.3	3.7	24.9
	[1.5,26.0]	[1.0,16.9]	[15.0,67.8]	[1.4, 7.1]	[3.7,63.8]
4	8.5	5.7	41.0	4.6	25.4
	[1.4,25.7]	[1.1,18.7]	[13.6,65.7]	[1.9, 8.7]	[4.0,63.5]
8	8.4	7.2	37.5	5.1	25.1
	[1.3,27.0]	[1.2,21.9]	[11.7,63.3]	[2.1, 9.7]	[4.2,62.4]
12	8.7	7.8	36.0	5.2	24.8
	[1.3,28.3]	[1.2,23.6]	[11.1,62.3]	[2.1,10.0]	[4.2,61.5]
16	8.8	8.2	35.3	5.3	24.3
	[1.2,29.0]	[1.2,24.7]	[10.6,62.0]	[2.1,10.0]	[4.1,60.9]

References

Alexius, Annika and Erik Post, "Exchange rates and asymmetric shocks in small open economies," *Empirical Economics*, 2008, *35* (3), 527–541.

Aron, Janine, Ronald Macdonald, and John Muellbauer, "Exchange Rate Pass-Through in Developing and Emerging Markets: A Survey of Conceptual, Methodological and Policy Issues, and Selected Empirical Findings," *The Journal of Development Studies*, 2014, *50* (1), 101–143.

Bacchetta, Philippe and Eric van Wincoop, "A theory of the currency denomination of international trade," *Journal of International Economics*, December 2005, *67* (2), 295–319.

Barhoumi, Karim, "How Structural Macroeconomic Shocks Can Explain Exchange Rate Pass-Through in Developing Countries? A Common Trend Approach," *Global Economy Journal*, 2009, *9* (2).

Blanchard, Olivier Jean and Danny Quah, "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *The American Economic Review*, 1989, *79* (4), pp. 655–673.

Choudhri, Ehsan U. and Dalia S. Hakura, "Exchange rate pass-through to domestic prices: Does the inflationary environment matter?," *Journal of International Money and Finance*, 2006, *25* (4), 614 – 639.

Clarida, Richard and Jordi Gali, "Sources of real exchange-rate fluctuations: How important are nominal shocks?," *Carnegie-Rochester Conference Series on Public Policy*, December 1994, *41* (1), 1–56.

Clemente, Jess, Antonio Montas, and Marcelo Reyes, "Testing for a unit root in variables with a double change in the mean," *Economics Letters*, 1998, *59* (2), 175 – 182.

Dornbusch, Rdiger, "Exchange Rate Economics: Where Do We Stand?," *Brookings Papers on Economic Activity*, 1980, *11* (1, Tenth Anniversary Issue), 143–206.

Estima, *RATS Version 8: User's Guide*, Estima, 2010.

Frimpong, Siaw and Anokye Mohammed Adam, "Exchange Rate Pass-Through in Ghana," *International Business Research*, April 2010, *3* (2).

Gali, Jordi, "How Well Does the IS-LM Model Fit Postwar U.S. Data?," *The Quarterly Journal of Economics*, 1992, *107* (2), pp. 709–738.

Geiger, Michael and Chorching Goh, "Ethiopia Economic Update Overcoming Inflation,Raising Competitiveness," Technical Report 20026, The World Bank 2012.

Hossain, A. Akhand, "Granger-Causality Between Inflation, Money Growth, Currency Devaluation and Economic Growth in Indonesia, 1951-2002," *International Journal of Applied Econometrics and Quantitative Studies*, 2005, *2* (3), 45–68.

IMF, "The Federal Democratic Republic of Ethiopia: Staff Report for the 2014 Article IV Consultation," IMF Staff Country Reports 14/303, International Monetary Fund October 2014.

Ito, Takatoshi and Kiyotaka Sato, "Exchange Rate Pass-Through and Domestic Inflation: A Comparison between East Asia and Latin American Countries," Discussion papers 07040, Research Institute of Economy, Trade and Industry (RIETI) June 2007.

_ **and** _ , "Exchange Rate Changes and Inflation in Post-Crisis Asian Economies: Vector Autoregression Analysis of the Exchange Rate Pass-Through," *Journal of Money, Credit and Banking*, 2008, *40* (7), 1407–1438.

Juselius, K., *The Cointegrated VAR Model: Methodology and Applications* Advanced Texts in Econometrics, OUP Oxford, 2006.

Loening, Josef L., Dick Durevall, and Yohannes A. Birru, "Inflation dynamics and food prices in an agricultural economy: the case of Ethiopia," Policy Research Working Paper Series 4969, The World Bank 2009.

McCarthy, Jonathan, "Pass-through of exchange rates and import prices to domestic inflation in some industrialized economies," Technical Report 2000.

_ , "Pass-Through of Exchange Rates and Import Prices to Domestic Inflation in Some Industrialized Economies," *Eastern Economic Journal*, Fall 2007, *33* (4), 511–537.

Melesse, Wondemhunegn Ezezew, "Exchange Rate Transmission into Sectoral Consumer Price Inflation in Ethiopia-SVAR Approach," *Journal of Economic and Financial Modelling*, 2014, *2* (1), 37–54.

Mwase, Nkunde, "An Empirical Investigation of the Exchange Rate Pass-Through to Inflation in Tanzania," IMF Working Papers 06/150, International Monetary Fund June 2006.

Obstfeld, Maurice, "Floating Exchange Rates: Experience and Prospects," *Brookings Papers on Economic Activity*, 1985, *16* (2), 369–464.

Razafimahefa, Ivohasina Fizara, "Exchange Rate Pass-Through in Sub-Saharan African Economies and its Determinants," IMF Working Papers 12/141, International Monetary Fund June 2012.

Shambaugh, Jay, "A new look at pass-through," *Journal of International Money and Finance*, June 2008, *27* (4), 560–591.

Sims, Christopher A., "Are forecasting models usable for policy analysis?," *Federal Reserve Bank of Minneapolis Quarterly Review*, 1986, (Win), 2–16.

_ **and Tao Zha**, "Error Bands for Impulse Responses," *Econometrica*, September 1999, *67* (5), 1113–1156.

Sissoko, Yaya and Sel Dibooglu, "The exchange rate system and macroeconomic fluctuations in Sub-Saharan Africa," *Economic Systems*, 2006, *30* (2), 141 – 156.

Taylor, John B., "Low inflation, pass-through, and the pricing power of firms," *European Economic Review*, June 2000, *44* (7), 1389–1408.

Wehinger, GertD., "Causes of Inflation in Europe, the United States and Japan: Some Lessons for Maintaining Price Stability in the EMU from a Structural VAR Approach," *Empirica*, 2000, *27* (1), 83–107.

World Bank, "Third Ethiopia Economic Update : Strengthening Export Performance through Improved Competitiveness," World Bank Other Operational Studies 20026, The World Bank June 2014.

Younger, Stephen D., "Testing the Link Between Devaluation and Inflation: Time Series Evidence from Ghana," *Journal of African Economies*, 1992, *1* (3), 369–394.

Chapter 4

Aid and Growth: *What Meta-Analysis Reveals*

Aid and Growth: What Meta-Analysis Reveals*

Tseday Jemaneh Mekasha and Finn Tarp

ABSTRACT *Recent literature in the meta-analysis category where results from a range of studies are brought together throws doubt on the ability of foreign aid to foster economic growth and development. This paper assesses what meta-analysis has to contribute to the literature on the effectiveness of foreign aid in terms of growth impact. We re-examine key hypotheses, and find that the effect of aid on growth is positive and statistically significant. This significant effect is genuine, and not an artefact of publication selection. We also show why our results differ from those published elsewhere.*

Keywords: aid and growth, meta-analysis, heterogeneity and publication bias

JEL classification: F35, O1, O4

I. Introduction

The literature on the potential impact of aid on growth is large and multifaceted.[1] Hansen and Tarp (2000) identify three generations of literature, and more recently, a fourth generation has emerged (see Arndt et al., 2010). A distinctive aspect of this generation is the view that aid's aggregate impact on economic growth is non-existent. Doucouliagos and Paldam (2008) (henceforth DP08) reach a similar pessimistic conclusion in their various papers based on a meta-analytic approach and a database including 68 studies on the aid-growth link.

---

* This paper is published in The Journal of Development Studies.

1 See, for example, Mosley (1986), White (1992), Tsikata (1998), Burnside and Dollar (2000), Morrissey (2001), Dalgaard et al. (2004), Tarp (2006), McGillivray et al. (2006), and Rajan and Subramanian (2008), Arndt et al. (2010), among many others.

More specifically DP08 ask (i) whether the aid effectiveness literature has established that aid has an impact on economic growth and if so how large is the impact; and (ii) what explains the heterogeneity in reported aid-growth effects? DP08 apply different meta-analysis techniques,[2] and conclude that the aid effectiveness literature has failed to show that the effect of development aid on growth is positive and statistically significant. They also attribute the variation in the reported effect of aid on growth to different study characteristics (DP08, pp 13-18).

In relation to the aid-growth literature, DP08 is an example where studies which have emerged over a long time period and which rely on differing methodologies and data sets are analysed. The DP08 analysis has attracted attention in policy debates about aid so we decided to re-examine their core aid-growth analytical result.[3] This was motivated by three underlying concerns: (i) the need to specify and justify the underlying *econometric* model used; (ii) *statistical* choices related to measurement of the effect estimates and calculation of the weighted average (both in terms of methodology and choice of precision of coefficient estimates); and (iii) time consuming and tedious data entry and coding work that is not always straightforward to replicate for those interested in the results.

This study reports what we uncovered in the process, and expands the DP08 meta-analysis in various ways that better reflect the econometric, statistical and data challenges faced in this type of research. In doing so, we address two main research questions that are common to any standard meta-analysis: (i) whether the empirical effect (in our case the impact of aid on growth) is different from zero when one combines the existing empirical evidence; and (ii) if so, whether the effect is genuine or an artefact of so-called publication bias (also referred to as the ‘file drawer’ problem).

Meta-analysis, or regression of regression analysis, is normally used with the aim of synthesizing the results from a group of studies while controlling for heterogeneity among studies. This methodology is usually applied in medical science research to assess the effectiveness of a well-defined healthcare intervention by combining data from primary studies that use randomized controlled trials. In recent years, however, meta-analysis has also been applied in economics and other fields of social science. One advantage of meta-analysis is that it can potentially address the subjectivity associated with traditional narrative literature surveys, and it may indeed provide a more systematic and objective (quantitative) assessment of an existing body of findings. Yet, the meta-methodology is by no means flawless (Stanley, 2001). Even if one accepts that meta-analysis is relatively more objective than narrative literature reviews, considerable room for subjectivity remains. For instance, in identifying the appropriate population of studies, authors often exercise personal judgment. Hence, bias from systematic selection of studies may follow. Moreover, decisions regarding data entry and coding, choice of a common metric for the effect size, statistical weighting of the effect estimates, model selection to conceptualize the meta-analysis and choice of moderator

[2] These include Funnel Asymmetry Test (FAT), Meta Significance Test (MST), and a meta-regression analysis (MRA). As regards the MRA both fixed and random model effects results are reported by DP08, who opt for relying on the fixed effects (see DP08: 13)

[3] Doucouliagos and Paldam (2011) expand the dataset and provide a brief update of DP08 but their focus as well as basic methodological choices and conclusions are the same.

variables all involve different levels of personal judgment. Such judgment calls potentially lead to misleading conclusions and hence may jeopardize the relevance of meta-analysis as a quantitative tool for literature review (e.g., Bullock and Svyantek, 1985; Wanous et al. 1989).

In this paper, we bypass the general bias issue in literature search and rely on the exact same 68 studies as DP08. Though these studies are by no means an exhaustive list of papers in the aid-growth literature, we decided to stick to these 68 papers for the sake of comparison.[4] Besides, we use the same common metric to measure the effect size (i.e. partial correlation) and the same set of moderator variables as DP08.[5] Despite these similarities we have relied on a number of different analytical choices.[6] First, we differ from DP08 in relation to model selection for the meta-analysis. In meta-analysis, one can rely on either a 'fixed effects' or a 'random effects' model depending on the assumption the meta-analyst makes regarding the nature of the true effect. DP08 argue that there is a single 'true' effect of aid on growth, which is common to all the 68 studies. This implies that they assume that random sampling error is the only factor behind the variation in reported effect estimates among studies. As a result of this assumption of 'effect homogeneity', DP08 mainly focus on a fixed effects meta-analysis. Our expectation is, in contrast, that the impact of aid on growth across the 68 studies is heterogeneous, and using both statistical tests and graphical tools we reject the effect homogeneity assumption.

One can also rule out the effect homogeneity assumption on theoretical grounds as the effect of aid on growth is a function of other factors. For instance, Burnside and Dollar (2000), Dalgaard et al. (2004), Hansen and Tarp (2001), and Chauvet and Guillaumont (2004), among many others, use interaction terms by which the partial effect of aid on growth is a function, not a constant. Moreover, the fact that the type of aid, the way it is delivered and the donor-recipient relationships differ across countries and have changed over time implies that the primary studies will target different population effect estimates. In sum, the effect homogeneity assumption of the fixed effects model cannot be expected to hold for the aid growth literature. Consequently, we conclude that random effects meta-analysis is more appropriate and show that the underlying model choice does matter for the conclusions drawn.

Second, a major concern with the DP08 approach is the way the partial effect estimate is measured for papers that include non-linear terms like aid squared, aid-policy and aid-institutions. For papers that include one or more of these interaction terms the partial effect of aid will not be measured correctly if one ignores the coefficient of the non-linear term(s). To see this, consider the following growth regression:

4 This should not be taken as an approval of the list of papers identified by DP08. The literature is large and complex.

5 Even if we consider the entire moderator variable set used in DP08 to begin with, we eventually focus on the relevant ones using a General-to-Specific modeling approach to reduce the set.

6 In addition to showing how these differences matter for the results and conclusions, we began by fully replicating the results of DP08.

$$G = \beta_o + \beta_1 * aid + \beta_2 * (aid * X) + \beta_3 * Z + \varepsilon \quad (1)$$

where X may be aid, policy or institutions and Z is a vector of other explanatory variables. In this case, the partial effect of aid on growth is given by $(\beta_1 + \beta_2 * X)$.[7] However, the data in DP08 relies on β_1 as the partial effect of aid.

In the meta-analysis, this problem matters in particular for regressions that use the partial effect as a dependent variable. One case in point is in calculating the weighted average effect of aid on growth. We have recognized this issue by separately estimating the weighted average effect for papers that include one of the aforementioned interaction terms and for those that do not include any of these terms. As shown in Section 3 this choice matters for the results. DP08, on the other hand, ignored the issue.[8]

Third, we differ from DP08 in the method used to calculate the weighted average effect of aid on growth and in our choice of the measure of statistical precision of coefficient estimates. In DP08 the weighted average aid-growth effect is calculated using sample size as weights under the assumption that studies with large sample size are more accurate. Accordingly, DP08 tune in on sample size as the preferred measure of statistical precision of parameter estimates. This choice is not, however, in line with established best-practice in standard fixed and random effects meta-analysis, which recommends that the inverse of the variance of estimates should be used as weights (i.e., as the measure of statistical precision) when calculating the weighted average effects (pooled estimates) from an existing body of empirical literature. Sterne and Harbord (2009) also note that the precision of an effect estimate cannot be fully captured by sample size. Other data characteristics are important in determining standard errors. Studies with very different sample sizes may have the same standard error and precision and vice versa. Consequently, in our estimations of the weighted average (combined) effect of aid on growth, we use the inverse of the variance of estimates as weights (i.e., as measure of statistical precision). As shown in Section 3, the way the weighted average is estimated matters for the results. Moreover, in plotting the funnel plots used for visual inspection, we use the inverse of the standard error of the estimates as a measure of precision. DP08 use sample size; Sterne and Egger (2001) have demonstrated that this approach to measuring the precision in funnel plots is inappropriate.

---

7 Note that in the case of the aid squared term, the partial effect is $\beta_1 + 2\beta_2 * aid$.

8 We are aware of the conditional aid effectiveness meta-analysis in Doucouliagos and Paldam (2010) (DP10), applying a similar meta-analysis as in DP08. In the 2010 paper, the authors conclude that the aid effectiveness literature has failed to establish non-linear terms like aid squared. This conclusion is, however, only as valid as the meta-methodology employed in the paper. Besides, from a total of 147 regressions that include an aid squared term, 100 show a negative and significant coefficient for the aid-squared term (See DP10: 400, Table 2). In view of this, even if one accepts the conclusion in DP10, it only implies that the coefficients of the interaction terms, on average, 'should be zero' rather than indicating that the coefficient of the interaction term from each paper 'is actually zero'. Therefore, the findings of DP10 do not address/justify the concern that we have pointed out regarding the treatment of interaction terms in DP08.

Fourth, turning to data issues we began by re-entering all DP08 data and found reason for some recoding.[9] As a result, the number of observations used for the multivariate meta-regression-analysis (MRA) is increased from 471 to 519.[10] Nevertheless, we have followed DP08 throughout as closely as possible to make sure results are comparable. Thus, even if our revised data set does not exactly match that of DP08, the correlations between the two sets of data are high (Mekasha and Tarp, 2011, Table A9.1).

Before moving on to our analysis we highlight a general concern, which is a potential threat to the credibility of meta-analysis as a tool for quantitatively assessing an existing body of findings. This relates to differences in the quality of the primary studies (i.e., the observations). Meta-analysis combines results from different studies regardless of their quality, and this problem gets more pronounced in social science research where most studies are based on observational/non-experimental data. In contrast to controlled experiments, observational studies differ substantially in their model specification, econometric techniques, functional forms and research design leading to potential quality differences. Such differences are likely to lead to heterogeneity in effect estimates and unless properly captured, this heterogeneity may wrongly be interpreted as publication bias. It is therefore crucial to allow for quality differences in meta-analysis.

However, measuring (assessing) differences in qualities entails subjective judgment. It is nearly impossible to come up with a single yardstick against which quality of the primary studies is defined. Even if researchers agree on a single quality yardstick, how to take this into account in the meta-analysis is another challenge. One suggestion in this regard is to categorize studies as 'good' and 'bad' quality and do the meta-analysis either focusing only on the 'good quality' studies or undertake a separate meta-analysis for each category. But here subjectivity is an obvious issue. There is no way to objectively categorize studies as of 'good' versus 'bad' quality. Another suggestion is to use quality scoring as weights but this method faces strong criticisms on different grounds.[11] Yet another suggestion is to use 'quality' or 'some components of quality' as moderator variable in the meta-regression analysis (MRA) and see whether there is a systematic difference in effect size between 'well-designed' and 'badly-designed' studies.[12] While defining 'quality' may still introduce subjectivity, controlling for 'some components of quality' can partly address this issue.[13] However, this can only be used in the case of multivariate MRA and in general leaves the problem unsolved in the calculation of the weighted average effect where moderators cannot be controlled for.

9 See Table A9.1 in Mekasha and Tarp (2011). Note also that in our data we do not include the variable 'Danida affiliation'. None of the three authors classified by DP08 as Danida affiliated (studies 12, 13, 33, 34 and 40) fell into this category when the studies were examined.

10 Note that we were able to increase the number to 519 by re-coding the values of the moderator variables which, for some studies, were wrongly coded as missing in DP08 (Table A8 in Mekasha and Tarp, 2011).

11 See Greenland (1994), Higgins and Green (2011) and Jüni et al. (1999).

12 Card (2011) suggests that if differences are found, conclusions must be restricted to those studies that the researcher thinks produce most valid results. But here one can argue that this suggestion cannot be appealing if the sample size for the meta-analysis is small.

13 Controlling for 'some aspects of quality' enables the researcher to tell which aspects of quality affect the reported effect size and hence can guide future design of primary studies, Card (2011).

In sum, even if there is broad consensus regarding the importance of considering the quality of the primary studies, there is hardly agreement on how to measure quality and on the ways to incorporate it in the analysis. This makes the issue of quality differences across the primary studies a major caveat. In general, we need to bear the limitations in mind when trying to draw lessons from a meta-analysis and the more so for non-experimental research where the necessary tools to overcome the above challenges are very limited and are only starting to emerge. Moreover, unlike the case of random control trials where the treatment and its effect are well-defined, this is not always the case in observational macro-level studies like aid and growth. This accentuates the caution one needs to exercise when making inference from aid-growth meta-analysis.[14] So unless due care is taken, meta-analysis cannot per se guarantee an objective assessment of an existing body of findings. Moreover, it has long been understood in the medical profession that it does not follow (in any simple way) from a zero meta-impact result that the medical practitioner should immediately stop 'treatment' and leave the ailing patient alone. Absence of evidence should only with great care be interpreted as evidence of absence (as noted by Temple, 2010).

This paper is structured as follows. Section 2 deals with data and methodology, while detailed results are presented in Section 3. Section 4 concludes that meta-analysis, if applied meticulously, suggests a positive and statistically significant impact of aid on growth and importantly suggests there is no publication bias in the aid-growth literature. Various appendix tables in Mekasha and Tarp (2011) provide further background data and detail.

II. Data and Methodology

The data used here originate from 68 published and unpublished aid-growth studies identified by DP08 covering the period of 1970-2004. Since each of the 68 studies reports one or more regressions, we have a total of 542 observations (regressions) to work with.[15]

The first step in any standard meta-analysis is to establish whether the size of the combined empirical effect in the literature under investigation is significantly different from zero or not. This is done by examining the pooled estimates (i.e., the mean overall effect) of all the studies included. There are two approaches to calculating the pooled estimate, i.e. the fixed effects model and the random effects model.[16]

In the fixed effects model it is assumed that all studies come from a population with a fixed average effect size, meaning that all studies are assumed to share a common true effect.

---

14 Here it should be noted that meta-analysis of micro-level observational studies can be more informative as they have a well-defined treatment and better comparability compared to macro-level primary studies. For an example of micro-level meta-analysis see Havranek and Irsova (2011, 2012).

15 We removed one regression from the study (ID 30) as this regression is already included (coded) in study ID 29. In study ID 30, the author used the regression from study ID 29 purely for comparative purposes. Thus, correcting for this double coding leads to 542 observations rather than 543.

16 The terms fixed and random effects used in meta-analysis are quite different from the ones applied in standard panel data models in econometrics. In meta-analysis the difference between fixed and random effects models originate from the underlying assumption as regards the nature of the 'true' effects.

Accordingly, the observed effect size[17] is assumed to vary from one study to another only because of random sampling error (within study variation). In contrast, in the random effects model, the assumption is that studies were drawn from populations that differ from each other in ways that could affect the treatment effect (Borenstein et al., 2007). In this case, the effect size will vary both due to sampling error (the fixed effects model) and due to true variation in effect size (between study variations).

Furthermore, in calculating the pooled estimate and hence the combined empirical effect, each effect size is weighted, the weight being the inverse of the variance from each study. In the case of the fixed effects model the weight is given by $1/v_i$ where v_i is the within study variance. On the other hand, the weight in the random effects model is given by $1/(v_i + \tau^2)$ where v_i and τ^2 refer to the within and between study variances respectively.

Having estimated the mean overall effect, the next step is to examine whether this observed effect is genuine or an artefact of publication bias (the so-called file drawer problem). The most commonly used tool to make a preliminary examination of the presence of publication bias is funnel plots, which are visual graphical images that illustrate the relationship between treatment effects estimated in individual studies (plotted on the horizontal axis) and a measure of study precision (shown on the vertical axis). The idea is that the precision (accuracy) in estimation of the underlying treatment effect (in our case the impact of aid on growth) increases as the study size grows. Consequently, small studies are expected to scatter widely at the bottom of the graph, while the spread is expected to narrow among larger studies at the top of the funnel. If there is no bias the plot will take the shape of an inverted funnel, and be symmetrical around the expected true effect. As indicated above, since sample size cannot fully capture the precision of reported effect size, our choice of measure of precision for the vertical axis in funnel plots follows Sterne and Egger (2001). They argue that standard errors (or their inverses) are the most appropriate measure of the precision of reported effect size.[18]

Even if funnel plots help in tracing publication bias or in general small study effects in the data, visual assessment of funnel plots is essentially subjective. Moreover, Sterne and Harbord (2009) note that funnel plot asymmetry does not necessarily arise from publication bias. Other potential reasons include, for instance, heterogeneity in underlying effects and/or low methodological quality of smaller studies. So, funnel plots should be seen as a generic means for investigating small study effects (if small studies show a larger treatment effect), not as a tool to diagnose a specific type of bias. It is therefore prudent to complement graphical observations from a funnel plot inspection with statistical tests for funnel plot asymmetry. Egger et al. (1997) provide the most commonly used test in the meta-literature; their test is regression-based to assess skewness in a funnel plot. This test starts by examining

---

17 The term effect size refers to the magnitude of the effect observed in each study. In the meta-literature there are different metrics to measure this; the partial correlation coefficient being the most commonly used. As in DP08 we calculate the partial correlation coefficients of each study using $\sqrt{t^2/(t^2 + df)}$ where t and df refer to t-statistics and degrees of freedom respectively.

18 We also present the funnel plots with sample size for comparison with DP08, but our preferred measure of precision follows Sterne and Egger (2001).

the relationship between study i's reported effect size ($Effect_i$) and its associated standard error (SE_i) as follows:

$$Effect_i = \alpha_0 + \alpha_1 SE_i + \varepsilon_i \qquad (2)$$

According to Stanley (2005), one can divide this equation by SE_i to avoid potential problems of heteroscedasticity, rewriting equation (2) as:

$$t_i = \frac{Effect_i}{SE_i} = \alpha_1 + \alpha_0 \frac{1}{SE_i} + \mu_i \qquad \text{where } \mu_i \text{ is } \varepsilon_i/SE_i \qquad (3)$$

The main idea behind this test is that, assuming a non-zero underlying effect and absence of publication bias, small studies will have a precision ($1/SE_i$) and a standardized effect ($Effect_i/SE_i$) close to zero. Large studies will have high precision and the standardized effects are expected to scatter around a regression line that passes approximately through the origin. The slope of this regression line estimates both the size and direction of the underlying effect. Failure of the regression line to pass through the origin implies publication bias. The size of the intercept gives a measure of asymmetry; the larger the deviation from zero the higher the asymmetry and hence bias in the effect size reported by the literature.

In sum, equation (3) provides a basis for testing both funnel graph asymmetry and the presence of a genuine empirical effect beyond any publication bias. Stanley (2005) insists that the presence of an underlying genuine empirical effect, irrespective of publication bias, must be confirmed by another test. This is the so-called meta-significance test (MST), which verifies the authenticity of empirical effects by analysing the relationship between the natural logarithm of the absolute value of a study's standardized effect (t-statistics) and its degrees of freedom (df). The MST equation can be written as:

$$\ln(|t|) = \beta_0 + \beta_1 \ln(df) + \epsilon_i \qquad (4)$$

Equation (4) provides evidence of a genuine empirical effect if $H_0: \beta_1 \leq 0$ is rejected. This test helps to identify a genuine empirical effect over and above any publication bias, and the line of thinking is clear in the following quote:

> *Observing a positive association between df and the standardized test statistic throughout a given empirical literature is an additional means to confirm the authenticity of the effect in question. Without such a confirmation, seemingly positive findings reported in the literature may be the consequence of fortuitous misspecification or systematic publication biases. Without this or similar validation,*

a theoretical economic proposition should not be regarded as empirically corroborated or 'verified'. Seemingly strong empirical results across an entire literature might easily be the remnants of selected bias. (Stanley, 2005: 329)

III. Results and Discussion

In this section we first present the pooled estimate of the combined effect of aid on growth, and then turn to investigating whether the observed effect is genuine (authentic) or an artefact of publication bias.

The Weighted Average Effect of Aid on Growth

The first (and typically main) aim of any meta-analysis is to combine the available empirical evidence so as to establish whether the impact of an intervention is different from zero or not. Accordingly, in Table 1 we present the combined estimates of the impact of aid on growth (and the associated confidence intervals) from fixed and random effects meta-analysis. Both suggest a positive and significant effect of aid on growth (0.082 and 0.098 respectively) when the empirical evidence from the 68 studies is combined.

Table 1. Meta-Analysis of the Effect of Aid on Growth

Method	No. of Regressions	Pooled Estimate	95% CI Lower	95% CI Upper	P-value H_o: No Effect
Overall					
Fixed	537	0.082	0.076	0.089	0.000
Random	537	0.098	0.085	0.112	0.000

Note: Test for heterogeneity: Q = 1791.745 on 536 degrees of freedom (p = 0.000) and the estimate of between studies variance = 0.015. The number of regressions is 537 instead of 542 as four estimates do not have data on standard errors due to missing data, and we have also removed one regression from the study with ID38 as an outlier. We have also checked the sensitivity of the overall effect to the inclusion of the outlier and the results still hold. That is, 0.081 and 0.097 for the fixed and the random effects respectively.

Source: Authors' estimates.

One difference with DP08 is the way we calculate the weighted average from the aid-growth literature as we follow standard practice and calculate the pooled estimate (in both the fixed and random effects models) using the inverse of the variance as weight. DP08 used the sample size as weight and found a weighted average of 0.08.[19] Such a simple weighted average calculation implicitly assumes away between study heterogeneity and is similar to

19 DP08 (pp 8-10) indicate (but do not report) that the weighted average is statistically insignificant. Applying the standard fixed and random effects model on the *original DP08 data* shows that the aid-growth weighted average effect is positive and statistically significant both in the fixed and random effects model with a magnitude of 0.078 and 0.093 respectively.

fixed effects. The results in Table 1 confirm this as the magnitude of the weighted average fixed effects estimate appears to be similar to DP08.

The fixed effect estimate is based on the assumption that there is a single true effect size (population treatment effect) inherent in all studies. This assumption is empirically testable and the fixed effects result can easily be challenged if there is heterogeneity of true effects across studies. Heterogeneity may not always be an issue, as in tightly controlled medical experiments (Schell and Rathe, 1992). As we rely on a wide-ranging set of 68 different studies with varying foci, quality of research design and analytical approach, heterogeneity is to be expected. This is indeed what the Q-test for heterogeneity reported in Table 1 suggests.[20] The presence of heterogeneity is also clearly confirmed graphically in Mekasha and Tarp (2011). The fixed effects model based on the homogeneity of effects assumption is clearly inappropriate in a meta-analysis of the aid and growth literature. Indeed, the effect homogeneity claim does not appear to be supported by the evidence inherent in the data.[21] It is therefore appropriate to focus on the random effects model.

The weighted average effect of aid on growth is positive and statistically significant with a magnitude of 0.098 in the random effects meta-analysis. As can also be seen from Table 1, the DP08 weighted average effect estimate does not fall in our 95 per cent confidence interval which indicates that we can reject their 0.08 estimate at the 5 per cent level of significance. As shown in equation (1) the partial effect of aid on growth will not be measured correctly for papers that aim to capture non-linear effects of aid on growth. Table 2 shows how this matters for the result, including separately re-estimated weighted average effects by classifying the papers based on their treatment of non-linearity.

To illustrate, for papers that include the aid squared term overlooking $2\beta_2 * aid$ will overstate the weighted average effect of aid reported from these papers, because the expected sign of the coefficient of aid squared in (1) above is negative. This is consistent with the result reported in Table 2: the weighted average effect from papers that include the aid squared term is much higher than papers which do not include the aid squared term. In a similar fashion, for papers that include aid-policy and aid-institution interaction terms, the expected sign of the coefficient of the interaction term is positive. Hence, ignoring the $\beta_2 * X$ term in equation (1) will understate the estimated weighted average effect of aid. Again, this is confirmed by the results in Table 2. Papers that include either aid-policy or aid-institution interaction terms appear to have a lower weighted average effect compared to papers that do not include these terms.

The lower part of Table 2 reports the weighted average effect of aid separately for papers that include at least one of the above interaction terms and for those that do not include any of these interaction terms. The random effect estimate of the weighted average effect of aid for

20 The test involves $Q = \sum_{i=1}^{k} w_i\,(T_i - T)^2$ where T_i is the estimate of the effect magnitude, T is the weighted average and w_i is the weight (the inverse of the variance of T_i). Under the null hypothesis of homogeneity, Q is distributed as chi-square with degrees of freedom equal to the number of studies minus one.

21 Even when one applies the heterogeneity tests on the original DP08 data, there is no ground to accept the effect homogeneity assumption of the fixed effects model.

the latter group appears to be positive and statistically significant with a magnitude of 0.138. This magnitude is higher than the estimate found for papers that include at least one of the interaction terms. Moreover, this estimate is also higher than the one reported in Table 1 where issues with non-linearity are ignored.

Thus, overlooking the coefficients of aid squared, aid-policy and aid-institution interaction terms, in the calculation of the partial effect of aid on growth, leads to a biased weighted average effect from the aid-growth literature. While leaving out the coefficient of the aid squared term leads to an upward bias in the weighted average, the bias in the case of aid-policy and aid-institutions is downward. In light of this, the weighted average effect reported in DP08 is biased. To sum up, when one combines the existing empirical evidence from the 68 studies, the results suggest that the effect of aid on growth is about 0.14 and is statistically significantly different from zero.

Table 2. Meta-Analysis of the Effect of Aid on Growth by Classifying the Studies based on the type of Non-Linear Terms Included in the Papers

Type of Non-linearity used in the papers:	No. of Regressions	Combined Effect Estimate	95% CI Lower	95% CI Upper	P-value H_o: No Effect
Studies with aid square					
Fixed	97	0.124	0.112	0.137	0.000
Random	97	0.131	0.110	0.153	0.000
Studies without aid square					
Fixed	441	0.064	0.056	0.072	0.000
Random	441	0.087	0.071	0.104	0.000
Studies with aid-policy					
Fixed	157	0.044	0.034	0.054	0.000
Random	157	0.044	0.027	0.060	0.000
Studies without aid-policy					
Fixed	381	0.113	0.104	0.122	0.000
Random	381	0.131	0.111	0.150	0.000
Studies with aid-institution					
Fixed	27	-0.112	-0.142	-0.081	0.000
Random	27	-0.112	-0.149	-0.075	0.000
Studies without aid-institution					
Fixed	511	0.091	0.084	0.098	0.000
Random	511	0.108	0.094	0.122	0.000
Studies with at least one of the three interaction terms					
Fixed	232	0.067	0.058	0.075	0.000
Random	232	0.066	0.051	0.082	0.000
Studies without the interaction terms					
Fixed	306	0.109	0.097	0.120	0.000
Random	306	0.138	0.113	0.162	0.000

Note: The Q tests for heterogeneity for studies with and without conditionality are Q = 756.157 on 231 degrees of freedom (p-value = 0.00) and Q = 1106.690 on 305 degrees of freedom (p = 0.000) respectively.

Source: Authors' estimates.

Publication Bias versus Authentic Effect

Publication bias is typically said to exist when researchers, editors and reviewers tend to favour statistically significant findings causing studies that yield relatively small and/or insignificant results to remain unpublished (i.e., remain 'in the file drawer'; see Stanley, 2005).[22] Whether this is indeed a problem in the aid-growth literature is not easy to say. In this literature, small and insignificant results have on several occasions drawn considerable academic and policy attention after which they have been shown not to be robust to even minor changes in data and methodology. Prominent examples include the 'micro-macro' paradox by Mosley (1986); the 'aid only works with good policy' hypothesis by Burnside and Dollar (2000); and the Rajan and Subramanian (2008) 'aid is insignificant' finding.[23] In any case, if a publication/small study bias exists it would tend to bias empirical effects, and as such must be carefully investigated with a view to disentangling any genuine empirical impact from publication effects. In line with established practice in the meta-literature we first use funnel plots to visually examine if the aid-growth literature seems to suffer from such bias.

Figure 1: Funnel plot with pseudo 95% confidence limits

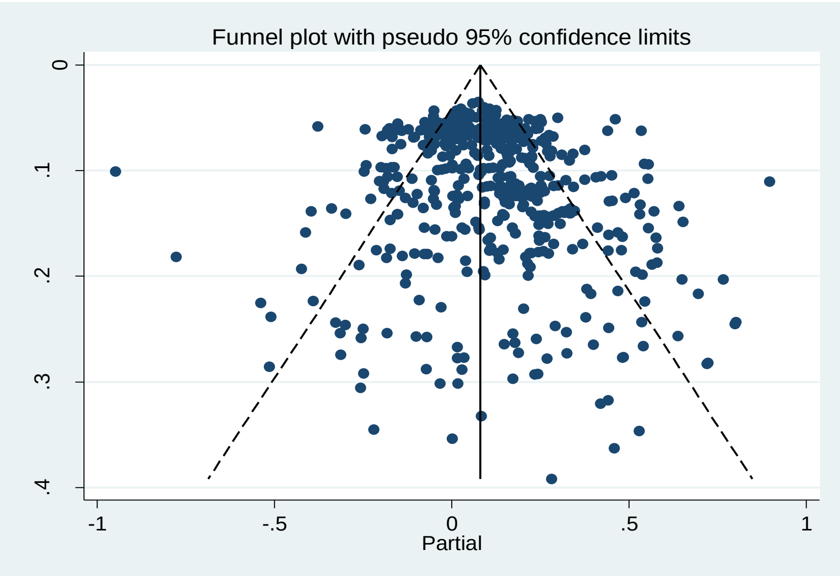


Figure 1 presents a funnel plot using standard error as the measure of precision.[24] The vertical line at the centre of the funnel plot shows a summary estimate of the effect size from the 68 aid-growth studies. When there is no bias, estimates are expected to vary randomly and evenly around this estimate. The diagonal lines in the figure represent the 95 per cent

22 Also, small studies tend to have large standard errors leading to insignificant results. If this leads authors to strive to come up with large-sized effects in order to compensate for the high standard errors such a bias should be detected.

23 See Hansen and Tarp (2000), Hansen and Tarp (2001), Dalgaard et al. (2004), and Arndt et al. (2010).

24 When standard errors are along the vertical axis, the vertical axis is reversed (zero at the top), so as to put large studies at the top of the graph reflecting that larger studies have smaller standard errors.

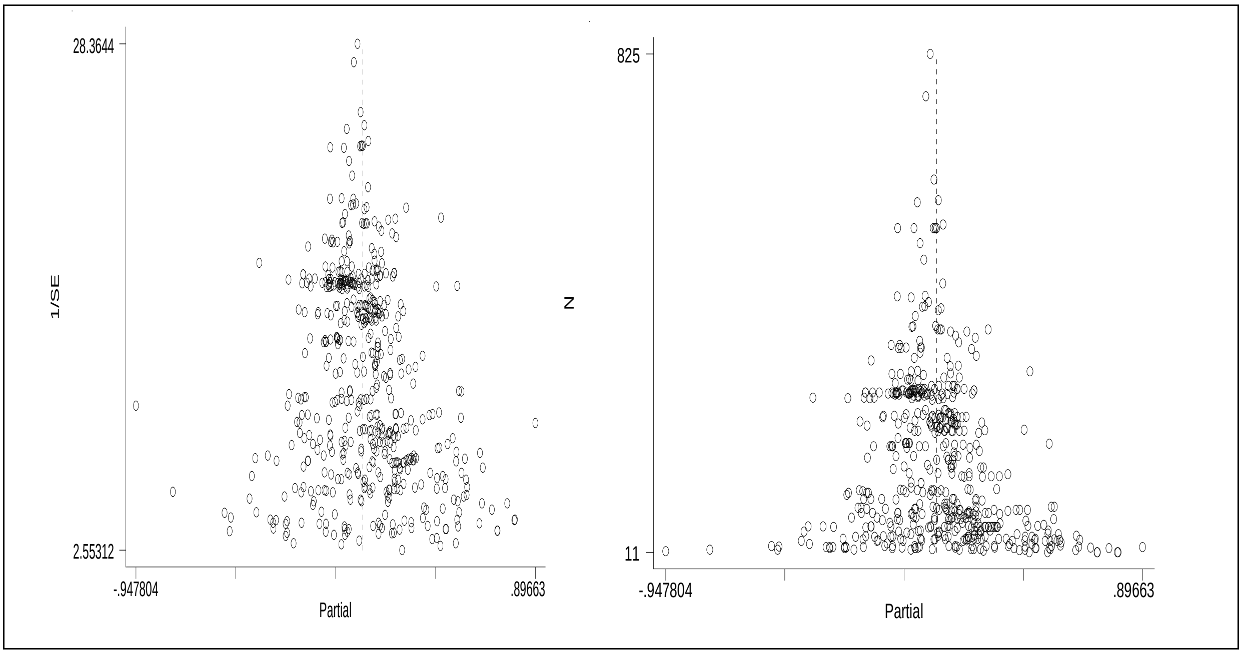
confidence limits around the summary treatment effect for each standard error on the vertical axis.[25] These lines show the expected distribution space of studies in the *absence of heterogeneity*. That is, assuming there is no heterogeneity in the reported effect sizes among studies, 95 per cent of the studies should lie within the funnel defined by the diagonal lines.

As can be seen from the funnel plot in Figure 1, the estimates from the aid-growth literature are fairly randomly distributed around the fixed effect estimate. Although the distribution of the studies to the right of the funnel seems relatively more concentrated, there is no clear asymmetry in the funnel graph. This lack of asymmetry becomes clearly visible in Figure 1.1. Figure 1.1.A relies on the inverse of standard error as the measure of precision and is thus our preferred funnel plot. This figure depicts the clearly symmetrical distribution of the effect of aid on growth as estimated from the 68 studies. A similar impression is also observable in Figure 1.1.B with sample size as the measure of statistical precision (for comparison with DP08). In general, these funnel plots provide no basis to argue for a directional bias once one places the reference line at the correctly estimated overall empirical effect (see also Figure A3 in Mekasha and Tarp, 2011).

Figure 1.1: Funnel plots of the aid–growth literature

A. 1/Se used as precision

B. Sample size used as precision



While the above funnel plot analysis provides no grounds to claim that a publication bias is present, it is premature to draw any firm conclusion about potential publication bias from this evidence. Even though funnel plots may be revealing, their interpretation is subjective and

25 The summary estimate of the effect size in Figure 1 is obtained from the fixed effect model (under the effect homogeneity assumption). This presents one limitation in funnel plot analysis. Vevea and Hedges (1995) explain why one should not necessarily associate asymmetry in the funnel plot with publication bias. Presence of heterogeneity can also potentially lead to such an asymmetry in the funnel plot.

potentially ambiguous so statistical testing is required. The most commonly used statistical test of publication bias is the Egger et al. (1997) test, also known as the funnel asymmetry test (FAT) (Stanley, 2005). FAT basically estimates equation (3), which is then expanded in a next stage to control for more explanatory variables.

The main variables of interest are the constant term and the coefficient of 'precision'. While the coefficient of 'precision' shows the magnitude and direction of any genuine underlying effect over and above any possible bias, the constant term depicts the existence and degree of the bias in the literature surveyed. The results of bivariate and multivariate meta-regression analysis are presented in Tables 3 and 4 respectively. In the bivariate FAT meta-regression-analysis (FAT-MRA) the dependent variable is the standardized effect of aid (t-statistics) regressed on the inverse of the standard error (i.e. precision). Since more than one regression is taken from most of the studies, observations within a study are unlikely to be independent. To address this, standard errors are clustered on publications in all regressions.[26] For the sake of comparison, we also report heteroskedasticity consistent and heteroskedasticity and autocorrelation consistent (HAC) standard errors.

Table 3. Bivariate FAT meta regression analysis dependent variable = standardized effect (t-stat)

	(1)	(2)	(3)
Variables	Robust	HAC	Clustered
Bias Coefficient			
Constant	0.794***	0.794***	0.794***
	(0.164)	(0.223)	(0.297)
Genuine Effect of Aid			
Precision	0.0245*	0.0245	0.0245
	(0.0142)	(0.01998)	(0.0260)
Observations	537	537	537
R-squared	0.005	0.005	0.005

Note: Robust, heteroskedastcity and autocorrelation consistent and clustered standard errors in parentheses ***p<0.01, ** p<0.05, * p<0.1.

Source: Authors' estimates.

Despite the reasonable symmetry in the funnel plot discussed above, the result from the bivariate regression depicted in Table 3 seems to suggest presence of a positive and statistically significant publication bias. The positive sign of the bias suggests that small studies with high standard error tend to report a high partial effect of aid on growth, and

26 In DP08 the results appear to be very sensitive to clustering, see Table A5 and discussion in Mekasha and Tarp (2011)..

hence a statistically significant effect. The FAT-MRA can also be relied on to identify genuine empirical effects of aid on growth regardless of publication bias (Stanley, 2008). In Table 3 this genuine empirical effect is captured by the coefficient of ‘precision’ and the FAT-MRA shows a positive and significant effect in column 1, but this does not appear to be the case when we apply HAC and clustered standard errors. The result reported in column 1 of Table 3 – concerning both publication bias and genuine empirical effect – is fundamentally the same as what DP08 find (they do not use HAC or clustered standard errors). This result is the main basis for their claim that aid is ineffective and publication bias is a problem in the aid-growth literature. Unlike DP08, however, we do not believe that we should conclude and stop the analysis here. Digging deeper is revealing.

The above bivariate Egger et al. (1997) test is commonly criticized for leading to an inflated false-positive rate (high type I error), and such false positive results become a major issue especially when there is between study heterogeneity (see Ioannidis and Trikalinos, 2007). In a similar manner, Stanley (2005) argues that heterogeneity of effects may induce asymmetry into the funnel plots even in the absence of publication bias. This implies that failure to account for factors that can explain heterogeneity in research findings will potentially exaggerate the bias. As heterogeneity is evident in the aid-growth literature one should refrain from making inference about publication bias based on the bivariate regression (Harbord et al, 2009). Instead, one needs to turn to a multivariate analysis.

In the aid-growth literature, drawing conclusions based on bivariate regression will obviously lead to misleading inference for various reasons. The fact that some studies aim to estimate the direct impact of aid on growth while others focus on identifying the transmission channels (such as investment, health, education) makes effect estimates heterogeneous. That is, compared to the former, the direct effect of aid on growth is likely to be smaller in the latter case where the channels are already controlled for. Due to this, the reported effect estimates from the different regression models will obviously vary as a function of the controls included in the regressions, but this is not because of publication bias. Disregarding this fact in the FAT-MRA makes it look as if there is publication bias. Thus, one needs to incorporate information about the controls included in the underlying regressions of the primary studies in the biavariate FAT-MRA regressions. As Stanley (2005) also indicates, if such important information is not controlled for, the FAT-MRA will like any other econometric analysis suffer from omitted variable bias.

Accordingly, we expand the bivariate FAT-MRA model reported above into a more general FAT-MRA by including important explanatory variables that can potentially affect the reported variation (heterogeneity) in research findings.[27] We do not pretend to have insight on this point that goes beyond that of DP08. Accordingly, we first expand the FAT-MRA model by including all the 50 moderator variables they identified. The result from this regression can be found in Mekasha and Tarp (2011), which shows that the magnitude of the precision coefficient improves and becomes significant in two of the three cases. Moreover and importantly, after controlling for factors that can potentially explain heterogeneity in reported effects, the bias coefficient (i.e. the constant term) becomes insignificant in all cases.

27 See for example Rose and Stanley (2005), Abreu et al. (2005) and Stanley (2005, 2008).

This suggests that once the moderator variables (study characteristics) are controlled for then there is no publication bias.

We also note that most of the variables included in the multivariate regression are also statistically insignificant. There is, in other words, a trade-off here between including all the 50 moderator variables in order to control for/explain heterogeneity versus potential multi-collinearity and loss of degrees of freedom. Moreover, all controls are not equally important in contributing to the omitted variable bias and/or explaining heterogeneity. We therefore follow the General-to-Specific (GETS) modelling procedure by Krolzig and Hendry (2001) to systematically reduce the insignificant variables from the multivariate model. By doing so we eliminate 21 of the 50 moderator variables that appear to be non-important; the adjusted R^2 increase from 41 to 43 per cent, supporting the removal of the 21 moderators. The results from the reduced multivariate model are reported in Table 4.

As can be seen from the multivariate FAT-MRA results in Table 4, the genuine impact of aid on growth, as reflected in the coefficient of 'precision', is found to be positive and statistically significant in all three cases with a magnitude of 0.17. To put our results in perspective, we did a back-of-the-envelope calculation based on the estimates in Arndt et al. (2010).[28] This exercise shows that our finding is quite close to their estimate. Compared to the bivariate model, controlling for other variables which can potentially affect the reported variation of the effect of aid on growth greatly improves the magnitude of the genuine effect of aid. Moreover, in all the regressions the constant term, i.e. the parameter used to test for existence of publication bias, becomes statistically insignificant. This is consistent with the result from the funnel plot and indicates lack of evidence to suggest presence of publication bias in the aid-growth literature.

Table 4. Multivariate FAT meta-regression analysis: reduced model dependent variable = standardized effect (t-stat)

	(1)	(2)	(3)
Variables	Robust	HAC	Clustered
Bias Coefficient			
Constant	-0.232	-0.232	-0.232
	(0.321)	(0.308)	(0.350)
Genuine Effect of Aid			
Precision	0.166**	0.166**	0.166*
	(0.0733)	(0.0843)	(0.0924)

28 Note that these estimates are not directly comparable as the estimate in the present paper is given as a partial correlation while the one in Arndt et al. (2010) takes an elasticity interpretation. We make the comparison by first changing the coefficient estimates from Arndt et al. (2010) to a partial correlation using the same formula used to convert the coefficient estimates of the primary studies included in this meta-analysis. Accordingly, we convert a total of 10 regressions from Arndt et al. (2010) to partial correlation and we get a weighted average of 0.173 which is the same as what we get in this meta-analysis. Note, however, that if we focus only on the most preferred regressions from Table 4 in Arndt et al. (2010), this weighted average effect will be 0.26.

Publication Outlet			
Working paper	-0.0697***	-0.0697***	-0.0697***
	(0.0167)	(0.0193)	(0.0184)
Cato	-0.202***	-0.202***	-0.202***
	(0.0324)	(0.0295)	(0.0282)
JDS	-0.0833***	-0.0833***	-0.0833***
	(0.0280)	(0.0271)	(0.0272)
JID	-0.0587***	-0.0587**	-0.0587*
	(0.0196)	(0.0239)	(0.0304)
EDCC	-0.146***	-0.146***	-0.146***
	(0.0389)	(0.0434)	(0.0501)
Applied economics	-0.116**	-0.116**	-0.116**
	(0.0545)	(0.0574)	(0.0519)
Author Detail			
World Bank	-0.0853***	-0.0853***	-0.0853***
	(0.0204)	(0.0198)	(0.0178)
Gender	-0.0737***	-0.0737***	-0.0737**
	(0.0202)	(0.0258)	(0.0293)
Influence	0.0668***	0.0668***	0.0668***
	(0.0164)	(0.0167)	(0.0162)
Data			
Panel	0.105***	0.105***	0.105**
	(0.0379)	(0.0404)	(0.0426)
No. of years	-0.0106***	-0.0106***	-0.0106***
	(0.00162)	(0.00159)	(0.00152)
Asia	0.0303	0.0303	0.0303
	(0.0222)	(0.0222)	(0.0239)
Single country	0.491***	0.491***	0.491**
	(0.160)	(0.170)	(0.191)
y1960s	0.0547**	0.0547**	0.0547
	(0.0270)	(0.0289)	(0.0368)
y1990s	0.103***	0.103***	0.103***
	(0.0318)	(0.0329)	(0.0328)
Sub sample	0.0446**	0.0446***	0.0446**
	(0.0212)	(0.0169)	(0.0187)
Low income	-0.0879***	-0.0879***	-0.0879***
	(0.0284)	(0.0254)	(0.0328)
EDA	-0.0376**	-0.0376**	-0.0376**

	(0.0164)	(0.0176)	(0.0181)
Conditionality			
Aid square	0.0716***	0.0716***	0.0716***
	(0.0125)	(0.01015)	(0.0108)
Interaction institutions	-0.100***	-0.100***	-0.100**
	(0.0248)	(0.0291)	(0.0380)
Specification and Control			
FDI	0.0909***	0.0909***	0.0909**
	(0.0258)	(0.0343)	(0.0417)
Theory	0.0415***	0.0415**	0.0415**
	(0.0155)	(0.0165)	(0.0191)
Average	0.0115***	0.0115***	0.0115***
	(0.00232)	(0.00211)	(0.00206)
Inflation	-0.0510**	-0.0510**	-0.0510***
	(0.0204)	(0.0198)	(0.0173)
Size of government	0.101***	0.101***	0.101***
	(0.0150)	(0.0142)	(0.0151)
Financial development	0.0345***	0.0345**	0.0345**
	(0.0129)	(0.0139)	(0.0142)
Region dummy	-0.0313**	-0.0313**	-0.0313**
	(0.0123)	(0.0130)	(0.0127)
Openness	-0.0706***	-0.0706***	-0.0706**
	(0.0185)	(0.0226)	(0.0274)
Per capita income	-0.0709**	-0.0709**	-0.0709*
	(0.0283)	(0.0318)	(0.0383)
Observations	518	518	518
Adj. R-squared	0.425	0.425	0.425

Note: Q-test for heterogeneity: (chi2(518) = 1000; p>chi2 = 0.000). See Higgins and Thompson (2002) for the test of heterogeneity. Robust, Heteroskedastcity and Autocorrelation Consistent and clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' estimates.

On this basis we suggest that it is highly likely that the DP08 results suffer from omitted variable bias, noting that their conclusions are exclusively dependent on a bivariate FAT-MRA analysis. Mekasha and Tarp (2011, Table A3) also included a variety of robustness check for the FAT-MRA results presented in Table 4. They include considering studies after the 1990s only; excluding studies that did not include African countries; and finally considering published studies only. In all cases the key finding presented in Table 4 holds.

The above evidence should, as Stanley (2005) puts it, be confirmed by a meta-significance test (MST) for authentic effect before firm conclusions are drawn. The MST test uses the relationship between the logarithms of a study's absolute value of t-statistics and the degrees of freedom to examine a genuine empirical effect. A genuine empirical effect is reflected in a positive and statistically significant coefficient of the log of degrees of freedom in equation (4). The bivariate and multivariate results of our MST regressions are reported in Table 5 and Table 6 respectively.

As can be seen from the bivariate regression reported in Table 5, the coefficient of log of degrees of freedom (ln (df)) exhibits a positive sign, but is insignificant in all cases. This should come as no surprise. The results reported in Table 5 are from a bivariate regression, and it is likely that this bivariate MST-MRA suffers from omitted variable bias for reasons similar to those discussed above.

Table 5. Bivariate MST meta regression analysis dependent variable = ln (t-stat)

	(1)	(2)	(3)
Variables	Robust	HAC	Clustered
ln(df)	0.00338	0.00338	0.00338
	(0.0474)	(0.0568)	(0.0635)
Constant	0.0637	0.0637	0.0637
	(0.219)	(0.258)	(0.277)
Observations	538	538	538
R-squared	0.000	0.000	0.000

Note: Robust, heteroskedastcity and autocorrelation consistent and clustered standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' estimates.

We therefore turn again to the DP08 explanatory variables used for the FAT-MRA in Table 4 and run a multivariate MST-MRA. The results from the full model are presented in Mekasha and Tarp (2011); the first three columns of Table 6 report the reduced form MST-MRA model after systematically removing insignificant variables using the GETS modelling procedure. Column 4 of Table 6 checks if the result remains the same when one uses the log of the number of observations (ln(n)) instead of the log of degrees of freedom (ln(df)) as a measure of estimation accuracy.

As Table 6 demonstrates, in all the multivariate MST-MRA regressions, the coefficient of estimation accuracy is positive and significant. This underpins the authenticity of the positive and significant effect of aid on growth observed in the FAT-MRA regressions. Moreover, similar to the FAT robustness checks, the key finding holds under robustness checks for the MST-MRA results (Mekasha and Tarp, 2011).

Table 6. Multivariate MST meta regression analysis: reduced model dependent variable = ln (t-stat)

Variables	(1) Robust	(2) HAC	(3) Clustered	(4) Clustered
Genuine Empirical Effect				
ln(df)	0.328***	0.328***	0.328***	
	(0.0847)	(0.0964)	(0.0820)	
ln(n)				0.365***
				(0.0942)
Publication Outlet				
Working Paper	-0.626***	-0.626***	-0.626***	-0.639***
	(0.145)	(0.138)	(0.140)	(0.139)
CATO	-1.390***	-1.390***	-1.390***	-1.402***
	(0.285)	(0.258)	(0.220)	(0.218)
JDS	-0.606**	-0.606**	-0.606**	-0.611**
	(0.235)	(0.254)	(0.265)	(0.263)
EDCC	-0.877	-0.877**	-0.877***	-0.867***
	(0.541)	(0.354)	(0.316)	(0.316)
AER	-1.029***	-1.029***	-1.029***	-1.035***
	(0.320)	(0.265)	(0.272)	(0.270)
Author Details				
World Bank	-0.496**	-0.496**	-0.496**	-0.504**
	(0.203)	(0.194)	(0.213)	(0.212)
Gender	-0.400**	-0.400**	-0.400**	-0.402**
	(0.178)	(0.159)	(0.155)	(0.155)
Influence	0.334**	0.334**	0.334**	0.330**
	(0.135)	(0.129)	(0.129)	(0.130)
Data				
No. of Years	-0.0357**	-0.0357**	-0.0357**	-0.0356**
	(0.0149)	(0.0168)	(0.0169)	(0.0169)
Africa	-0.286*	-0.286*	-0.286*	-0.297*
	(0.164)	(0.166)	(0.147)	(0.149)
Single Country	1.426***	1.426***	1.426***	1.389***
	(0.300)	(0.298)	(0.252)	(0.249)
y1960s	0.399**	0.399*	0.399*	0.388*

	(0.201)	(0.217)	(0.233)	(0.231)
y1990s	1.016***	1.016***	1.016***	1.004***
	(0.203)	(0.211)	(0.209)	(0.207)
Conditionality				
Aid Square	0.574***	0.574***	0.574***	0.573***
	(0.141)	(0.146)	(0.124)	(0.127)
Interaction Institutions	0.822***	0.822***	0.822***	0.814***
	(0.216)	(0.217)	(0.217)	(0.216)
Specification and Control				
FDI	0.576***	0.576***	0.576***	0.549***
	(0.173)	(0.145)	(0.137)	(0.137)
Gap Model	0.294	0.294	0.294*	0.316**
	(0.211)	(0.185)	(0.149)	(0.151)
Theory	0.612***	0.612***	0.612***	0.618***
	(0.141)	(0.141)	(0.156)	(0.158)
Average	0.0530***	0.0530***	0.0530***	0.0540***
	(0.0128)	(0.0143)	(0.0119)	(0.0124)
Lag used	0.259	0.259	0.259	0.259
	(0.184)	(0.161)	(0.186)	(0.185)
Size of government	0.601***	0.601***	0.601***	0.596***
	(0.137)	(0.137)	(0.135)	(0.134)
Region Dummy	-0.329**	-0.329**	-0.329***	-0.332***
	(0.148)	(0.124)	(0.0952)	(0.0952)
Openness	-0.275**	-0.275**	-0.275**	-0.276**
	(0.124)	(0.123)	(0.120)	(0.122)
Constant	-1.681***	-1.681***	-1.681***	-1.873***
	(0.434)	(0.470)	(0.354)	(0.403)
Observations	519	519	519	519
Adj. R-squared	0.203	0.203	0.203	0.202

Note: Test for heterogeneity: (chi2(518)= 550.16; P>chi2=0.317)Robust, Heteroskedastcity and Autocorrelation Consistent and Clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Source: Authors' estimates.

Our results from MST-MRA also stand in contrast to the conclusions of DP08. They found a negative and insignificant coefficient on ln(df) and suggested that there is a lack of evidence to support the idea that development aid has an effect on economic growth. Once again, this is based on a simple bivariate MST, which fails to take into account other explanatory variables. This negative conclusion on aid effectiveness does not survive when the bivariate model is expanded to the multivariate context.

IV. Conclusions

Our main aim was to contribute to the aid-growth literature and associated policy debates using the meta-analysis of 68 studies employed by DP08. We also use the same measure of effect size (partial correlation) and the same moderator variables for the multivariate analysis. There are four major differences: (i) model selection, i.e. the choice between the random and fixed effects model; (ii) the way the effect size is treated for papers that include non-linear terms; (iii) choice of statistical weighting for the effect estimates both in calculating the weighted average effect and in the funnel plots; (iv) in the process of data entry and coding. Having fully replicated the analysis by DP08 and identified our differences, we expand their meta-analysis in various ways that better reflect the econometric, statistical and data challenges faced in this type of research. What did we find?

On the data issues, some recoding and filling in missing values resulted in an increase in the number of observations for our meta-analysis.

In relation to model selection, our results show that the fixed effects model assumption of a single true effect common to all studies is unrealistic in the aid-growth literature. Specifically, statistical tests reveal that there is heterogeneity in the estimate of the true effect of aid on growth across the 68 studies. Furthermore, the effect homogeneity assumption can be rejected, from the outset, on theoretical grounds. We thus emphasize that the random effects model is to be preferred as it allows for between study heterogeneity.

Accordingly, we calculated the weighted average effect of aid on growth (using the inverse of the variance as weight) relying on the random effects model. Our results show that the weighted average effect of aid on growth from the 68 studies is positive and statistically significant with a magnitude of 0.098. This finding stands in contrast to DP08, but allowing for heterogeneity across the studies the random effects results can reject DP08's estimate of 0.08 at the 5 per cent level of significance.

The partial effect of aid on growth for regressions that include interaction terms is not measured correctly as the coefficient of the interaction term(s) is not taken into account. This can potentially bias the weighted average effect of aid. By calculating the weighted average effect of aid separately for regressions with and without non-linear terms, the weighted average effect estimate of aid in the random effects model emerges as positive and significant with a magnitude of 0.14. This shows how disregarding the coefficient of the interaction terms in the calculation of the partial effect matters for the results. We thus suggest that future meta-analysis of aid and growth needs to find a way to properly incorporate the partial effect of aid from studies that include a non-linear term.

Having calculated the weighted average effect of aid on growth from the 68 studies, we moved on to check whether the observed effect is genuine or an artefact of publication bias using FAT-MRA and the General-to-Specific (GETS) modelling approach in choosing the important study characteristics (moderator variables) that help to explain the heterogeneity in research design across studies. The multivariate FAT-MRA results clearly suggest that publication bias is not a problem in the aid-growth literature once the heterogeneity is controlled for. The measure of publication bias obtained from the multivariate FAT-MRA model appears to be statistically indistinguishable from zero, which is in line with the reasonably symmetrical funnel plot depicted in Figure 1.1.A. In the same vein, the FAT-MRA results reported in Table 4 also confirm the positive and significant effect of aid on growth as depicted by the positive and statistically significant coefficient of precision.

The genuineness of the observed effect and hence the absence of publication bias in the aid-growth literature is further underpinned by the results of our MST-MRA regressions. As shown in Table 6, there is evidence of a clear empirical effect that goes beyond publication bias. Though the coefficient that verifies the authenticity of the impact of aid on growth is not significant in the bivariate MST, the authenticity of the observed positive and significant aid-growth impact becomes evident once we move to a multivariate setting. As shown in Mekasha and Tarp (2011), these findings are robust in different samples.

We also highlight the importance of heterogeneity in the true effect of aid on growth across the studies under review. As is evident from the Q-test for heterogeneity reported under Table 4, there still exists excess (unexplained) variation despite the inclusion of the relevant moderator variables. This confirms the presence of real heterogeneity in the true effect of aid on growth that goes beyond heterogeneity in research design. This is again consistent with the assumptions inherent in the random effects model and shows that the effect homogeneity assumption of the fixed effects model is not tenable.

To sum up, we have shown that the conclusions in DP08 do not hold when one applies meta-analysis rigorously to the aid-growth literature. We found a positive and significant effect of aid on growth and importantly found no evidence to suggest presence of publication bias. That said, and as pointed out from the outset, even if meta-analysis can potentially address the subjectivity associated with narrative literature reviews, it is far from flawless. For instance, subjectivity remains a threat in meta-analysis unless researchers carefully handle the judgment calls they encounter in various stages of the meta-research process. Moreover, differences in the quality of the primary studies and the lack of a reasonably objective tool to measure quality appear to be a major caveat, and especially for observational data-based studies. Differences in quality can lead to heterogeneity in effect estimates and unless properly captured, the heterogeneity can wrongly be perceived as publication bias. On top of this, and as our results show, there is real heterogeneity in the true effect of aid on growth that goes beyond methodological heterogeneity; that is, heterogeneity persists even after controlling for study characteristics. Given also the fact that meta-analysis is a method, which is more appropriate for data generated through Random Control Trials, the application to the aggregate aid-growth literature should only be undertaken with great caution.

One should not overstate the implications of the results from macro-level aid-growth meta-analysis. Nonetheless, such analysis, if applied rigorously according to best practice, can help in giving useful insight into the qualitative aspects of the research process; for example, to identify the presence or absence of publication bias in the literature under consideration. Besides, identification of the most relevant study characteristics that explain heterogeneity in the effect estimate can be relied on to improve research design of future primary studies. Regarding the quantitative results, although we do find evidence that is in line with Arndt et al. (2010), we remain vigilant in drawing strong implications. This is not only because of the limitations of applying meta-analysis to the macro-level aid-growth literature but also due to the fact that the estimate here is obtained by combining inherently heterogonous effect estimates. Moreover, the conclusions that emerge from the present review are obviously not the whole story about aid effectiveness. Economic growth, though important, is only one of the multifaceted development objectives of foreign aid. It should be noted that poverty reduction is now the main aim and target of most foreign aid programmes.[29] Finally, we fully agree with calls to improve the design and implementation of aid to the benefit of the poorest people in the poorest countries. Aid processes are complex and few would (and certainly not the present authors) dispute that they can be improved.

29 See Feeny and Ouattara (2009), Feeny (2003) and Gomanee et al. (2005).

References

Abreu, M., de Groot, H.L., and Florax, R.J. (2005) A meta-analysis of beta-convergence: the legendary 2%. *Journal of Economic Surveys*, 19(3): 389-420.

Arndt, C., Jones, S., and Tarp, F. (2010) Aid, growth, and development: Have we come full circle? *Journal of Globalization and Development*, 1(2): 1-27.

Borenstein, M., Hedges, L., and Rothstein, H. (2007) Meta analysis: Fixed effect vs. random effects. Available at: " http://www.meta-analysis.com/downloads/Meta-analysis_fixed_effect_vs_random_effects_sv.pdf ".

Bullock, R.J. and Svyantek, D.J. (1985) Analyzing meta-analysis: Potential problems, an unsuccessful replication, and evaluation criteria. *Journal of Applied Psychology*, 70(1): 108-115.

Burnside, C. and Dollar, D. (2000) Aid, policies, and growth. *American Economic Review*, 90(4): 847-868.

Card, N.A. (2011) Applied Meta-Analysis for Social Science Research. Methodology in the Social Sciences. [New York] Guilford Publications.

Chauvet, L. and Guillaumont, P. (2004) Aid and growth revisited: Policy, economic vulnerability and political instability. In B. Tungodden, N. Stern and I. Kolstad, editors, *Toward Pro-Poor Policies – Aid, Institutions and Globalization*, pages 95-109. World Bank: Oxford University Press.

Dalgaard, C.-J., Hansen, H., and Tarp, F. (2004) On the empirics of foreign aid and growth. *Economic Journal*, 114(496): F191-F216.

Doucouliagos, H. and Paldam, M. (2008) Aid effectiveness on growth: A meta study. *European Journal of Political Economy*, 24(1):1 - 24.

Doucouliagos, H. and Paldam, M. (2010) Conditional aid effectiveness: A meta-study. *Journal of International Development*, 22(4): 391-410.

Doucouliagos, H. and Paldam, M. (2011) The ineffectiveness of development aid on growth: An update. *European Journal of Political Economy*, 27(2): 399-404.

Egger, M., Smith, G.D., Schneider, M., and Minder, C. (1997) Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, 315(7109): 629-634.

Feeny, S. (2003) The impact of foreign aid on poverty and human wellbeing in Papua New Guinea. *Asia Pacific Development Journal*, 10(2): 73-93.

Feeny, S. and Ouattara, B. (2009) What type of economic growth does foreign aid support? *Applied Economics Letters*, 16(7): 727-730.

Gomanee, K., Girma, S., and Morrissey, O. (2005) Aid, public spending and human welfare: Evidence from quantile regressions. *Journal of International Development*, 17(3): 299-309.

Greenland, S. (1994) Invited commentary: A critical look at some popular meta-analytic methods. [American Journal of Epidemiology] 140(3): 290-296.

Hansen, H. and Tarp, F. (2000) Aid effectiveness disputed. *Journal of International Development*, 12(3): 375-398.

Hansen, H. and Tarp, F. (2001) Aid and growth regressions. *Journal of Development Economics*, 64(2): 547-570.

Harbord, R.M., Harris, R.J., and Sterne, J.A. (2009) An updated test for small-study effects in meta-analyses. In Sterne, J. A., editor, *Meta-Analysis in Stata: An Updated Collection from the Stata Journal*, pages 138-148. Stata Press.

Havranek T. and Irsova, Z. (2011), Estimating vertical spillovers from FDI: Why results vary and what the true effect is. *Journal of International Economics*, 85(2): 234-244.

Havranek, T. and Irsova, Z. (2012) Publication bias in the literature on FDI spillovers. *Journal of Development Studies*, forthcoming.

Higgins, J.P. and Green, S. (2011) *Cochrane handbook for systematic reviews of interventions*. Wiley Cochrane Series. Chichester: John Wiley & Sons.

Higgins, J.P. and Thompson, S.G. (2002) Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21(11):1539-1558.

Ioannidis, J.P. and Trikalinos, T.A. (2007) The appropriateness of asymmetry tests for publication bias in meta-analyses: A large survey. *Canadian Medical Association Journal*, 176(8): 1091-1096.

Jüni, P., Witschi, A., Bloch, R., and Egger, M. (1999) The hazards of scoring the quality of clinical trials for meta-analysis. *The Journal of the American Medical Association*, 282(11): 1054-1060.

Krolzig, H.-M. and Hendry, D.F. (2001) Computer automation of general to specific model selection procedures. *Journal of Economic Dynamics and Control*, 25(6-7): 831-866.

McGillivray, M., Feeny, S., Hermes, N., and Lensink, R. (2006) Controversies over the impact of development aid: it works; it doesn't; it can, but that depends. *Journal of International Development*, 18(7): 1031-1050.

Mekasha, T.J. and Tarp, F. (2011) Aid and growth: What meta-analysis reveals. Helsinki: World Institute for Development Economic Research (UNU-WIDER), *WIDER Working Paper WP/22*.

Morrissey, O. (2001) Does aid increase growth? *Progress in Development Studies*, 1(1): 37-50.

Mosley, P. (1986) Aid-effectiveness: The micro-macro paradox. *IDS Bulletin*, 17(2): 22-27.

Rajan, R.G. and Subramanian, A. (2008) Aid and growth: What does the cross-country evidence really show? *The Review of Economics and Statistics*, 90(4): 643-665.

Rose, A.K. and Stanley, T.D. (2005) A meta-analysis of the effect of common currencies on international trade. *Journal of Economic Surveys*, 19(3): 347-365.

Schell, C.L. and Rathe, R.J. (1992) Meta-analysis: a tool for medical and scientific discoveries. *Bulletin of the Medical Library Association*, 80(3): 219-222.

Stanley, T.D. (2001) Wheat from chaff: Meta-analysis as quantitative literature review. *Journal of Economic Perspectives*, 15(3): 131-150.

Stanley, T.D. (2005) Beyond publication bias. *Journal of Economic Surveys*, 19(3): 309-345.

Stanley, T.D. (2008) Meta-regression methods for detecting and estimating empirical effects in the presence of publication selection. *Oxford Bulletin of Economics and Statistics*, 70(1): 103-127.

Sterne, J.A. and Egger, M. (2001) Funnel plots for detecting bias in meta analysis: Guidelines on choice of axis. *Journal of Clinical Epidemiology*, 54(10): 1046-1055.

Sterne, J.A. and Harbord, R.M. (2009) Funnel plots in meta-analysis. In Sterne, J.A., editor, *Meta-Analysis in Stata: An Updated Collection from the Stata Journal*, pages 109-122. Stata Press.

Tarp, F. (2006) Aid and development. *Swedish Economic Policy Review*, 13(3): 9-61.

Temple, J. (2010) Aid and Conditionality. [in D. Rodrik and M.Rosenzweig]*Handbook of Development Economics*, Volume 5, [Amsterdam: Elsevier] pp. 4416-4522.

Tsikata, T. (1998) Aid effectiveness - a survey of the recent empirical literature. Washington, DC: International Monetary Fund, *IMF Papers on Policy Analysis and Assessments 98/1*.

Vevea, J. and Hedges, L. (1995) A general linear model for estimating effect size in the presence of publication bias. *Psychometrika*, 60(3): 419-435.

Wanous, J.P., Sullivan, S.E., and Malinak, J. (1989) The role of judgment calls in meta-analysis. *Journal of Applied Psychology*, 74(2): 259-264.

White, H. (1992) What do we know about aid's macroeconomic impact? An overview of the aid effectiveness debate. *Journal of International Development*, 4(2): 121-137.

Chapter 5

Aid and Income:
Another Time Series Perspective

Aid and Income: Another Time Series Perspective*

Matthijs Lof[†], Tseday Jemaneh Mekasha[‡] and Finn Tarp[§]

Abstract

In a recent article, Nowak-Lehmann, Dreher, Herzer, Klasen, and Martínez-Zarzoso (2012) (henceforth NDHKM) conclude that foreign aid has not had a significant effect on income, based on evidence from panel data potentially covering 131 countries over the period 1960-2006. The present study provides a replication of the empirical results reported by NDHKM and uncovers a number of critical issues regarding data handling and model specification. We also show that NDHKM's use of a single-equation regression is not a suitable empirical strategy for estimating the causal effect of aid on income. Given the nature of the variables and the question under investigation, a Vector Autoregressive (VAR) model can better address the inherent endogeneity problem in the aid-growth relationship. Evidence from a Panel VAR model estimated on the dataset of NDHKM, suggests a positive and statistically significant long run effect of aid on income.

* This study was prepared under the ReCom programme (see http://recom.wider.unu.edu/). The authors would like to acknowledge insightful comments by Phil Abbott, Tony Addison and Katarina Juselius. The paper is published in World Development.

† University of Helsinki and

‡ University of Copenhagen

§ UNU-WIDER and University of Copenhagen

1 Introduction

Researchers interested in foreign aid have, for several decades, done their best to empirically estimate the impact of aid on economic growth. This has not been easy, and both methodologies and results have varied over time. The aid effectiveness literature has passed through at least four different generations with each generation having its own distinguishing analytical features (see Hansen and Tarp (2000), Arndt et al. (2010)). A positive aid-growth association has been reported as characteristic across the first three generations of aid-growth empirical work surveyed by Hansen and Tarp (2000); but the fourth generation work discussed in Arndt et al. (2010) has suggested that aid may be impotent in spurring growth. [1] The balance of evidence in the last 3-4 years does appear, however, to be shifting again towards noting a positive and significant impact of aid on growth at the macro level.[2]

The use of time series techniques like co-integration analysis and vector autoregressive (VAR) models has so far been quite limited in aid-growth empirical research. Yet, studies are starting to emerge following better availability of data. One recent contribution is Juselius et al. (2012) who carry out a comprehensive study of the long-run effect of aid on a set of key macroeconomic variables including economic growth for a group of 36 SSA countries. Their findings provide clear support for a positive long-run impact of aid on the macro economy of recipient countries. Another recent time series contribution in relation to aid and growth is the paper by Nowak-Lehmann et al. (2012), henceforth NDHKM, who conclude that aid has an "insignificant or minute significant negative impact on per capita income" of recipient countries.

Overall, as noted in Juselius et al. (2012), the divergent evidence on aid effectiveness is perplexing in light of the fact that the data on aid and other macro variables used in most papers comes from the same publicly available databases. In explaining this, Juselius et al. (2012) argue that the choices researchers make regarding data transformations, econometric models, estimation methods and assumptions related to endogeniety or exogeniety are the main underlying reasons behind the observed discrepancies.

The primary objective of this study is therefore to illustrate the aforementioned points with reference to the aid-growth literature. Particularly, we show how inappropriate use of time series techniques including data transformation and econometric model choices can easily misrepresent key elements about how aid is allocated and lead to incorrect conclusions about aid effectiveness. We illustrate these points focusing on the recent contribution by NDHKM. Although we welcome their effort as a step towards increasing

[1] See Rajan and Subramanian (2008) and Moyo (2009).
[2] See, for example, Clemens et al. (2012) and Arndt et al. (2010).

the application of time series techniques in empirical aid effectiveness research, the NDHKM paper suffers from serious limitations as demonstrated in detail below. The second objective of our study is to present alternative empirical evidence on the effects of aid on income, by applying VAR models instead of the single-equation model considered by NDHKM, while using the same datatset. We argue that this methodology accurately addresses the endogeneity problem at hand in the aid-growth relationship, and is better suited to estimate the dynamic long-term effects of aid on income.

To achieve the objectives of this study, we begin by replicating the regression results reported by NDHKM. For this exercise, we make use of the replication files provided by NDHKM in the data archive of the *Canadian Journal of Economics*. The regressions are for the most part based on a panel of 50 countries, which is claimed to be "virtually balanced" with only 3% of the observations missing (NDHKM p. 298). Our replication reveals, however, that this is not the case. In most of the regressions only 30-40% of the available observations are actually used for estimation. The main reason for this omission is that NDHKM estimate a regression model that includes logarithmic transformations of variables that are not strictly positive.

Although the unbalancedness of the panel affects the asymptotic and finite-sample properties of the employed estimators (see Wooldridge (2001)), this is not our main point. We acknowledge that imperfect datasets are part of the reality in which empirical economists live. Macroeconomic panels are often unbalanced due to the fact that the starting period from which economic variables are available typically varies across countries. Researchers thus face a choice between optimizing the amount of observations, which then constitute an unbalanced panel, or to balance the panel, by cutting early observations from countries with long time series data.

The problem we address here goes much deeper and has serious implications for the results and conclusions reached. To begin, the observations in NDHKM are not simply *missing*, they are actually *omitted* by the authors. NDHKM compile an impressive dataset including relatively long time series on aid, income and other macroeconomic variables for a large group of countries. However, by trying to take logarithms of variables with negative values, a substantial fraction of this dataset is simply disregarded. While typically an unbalanced panel consists of time series of different length, in this case the logarithmic transformation creates huge gaps within the time series, which makes analyzing the dynamic properties of the data very difficult, if not impossible. The regression model, which is a log-linearization of a multiplicative Solow-type growth model, is mis-specified since not all the variables in the model are strictly positive.

Apart from these issues with data and model specification, the estimation results in NDHKM cannot be interpreted as a causal effect of aid on income. Although the applied methodology enables the analyst to consistently estimate the co-integrating coefficient, even when the regressor (aid) is endogenous, interpreting this estimate as a causal relationship between aid and income requires strict exogeneity of aid. In spite of this, NDHKM interpret their statistically insignificant estimate of the co-integrating coefficient as evidence of lack of a causal relationship between aid and income. A serious attempt to isolate potential causal (negative or positive) effects of aid on income is missing. Although tempting, interpreting a statistically insignificant co-integrating coefficient as lack of a causal relationship between aid and growth is in the present context misleading. This is not only because of the difficulty of interpreting insignificant coefficient estimates, but also due to the widely recognised fact that causality in the aid-growth relationship runs in both directions with potentially opposing effects.[3] Thus, without a clear identification strategy to isolate these two effects, finding a negative or insignificant parameter for aid, does not necessarily reveal anything about the impact of aid on growth.

Arguably, a system approach such as the VAR model applied in this study provides illuminating insights when estimating the intertemporal effects of aid on income. Since the seminal work by Sims (1980), VAR models have become the benchmark in empirical macroeconomics. In contrast, in the aid literature VAR models have not yet gained the same popularity, although there have been some recent applications of VAR models, such as Osei et al. (2005), Hansen and Headey (2010), Gillanders (2011), Juselius et al. (2012) and Kang et al. (2012). In the present study we apply a Panel VAR model to the dataset of NDHKM to investigate the effect of aid on income. By allowing explicitly for an effect of aid on income as well as an effect of income on aid, we find that the former effect is both positive and significant.

The study is structured as follows. In Section 2, after presenting the replication results, we discuss the data handling concerns uncovered by the replication exercise. In Section 3, we review the problems with the empirical strategy of NDHKM, and introduce our own strategy. Section 4 presents results from estimating VAR models on the NDHKM dataset. Section 5 concludes that when a Panel VAR model is applied to the same dataset as in NDHKM, a positive and statistically significant long run effect of aid on growth emerges.

2 Replication results

We begin this replication study by noting that we are able to exactly replicate virtually all the empirical results reported by NDHKM. Tables 1-7 show the replications of the corresponding Tables 1-7 in NDHKM. Except for the sixth column of Table 6, these tables match the results reported by NDHKM. After outlining the empirical model, our concerns

[3] Since donors give more aid to poor countries and lower their assistance as recipient countries get richer, this bi-directional relation creates a problem of endogeniety, which is a widely accepted challenge in the aid-growth empirical research.

raised by this replication exercise are discussed below. NDHKM estimate the following model, relating income per capita (*LY*) to population growth (*LPOPGPLUS*), domestic savings (*LSDOMY*), net external savings (*LSEXTNY*) and net aid transfers (*LSNATY*), with all variables measured in logs.:

$$LY_{i,t} = b_{0,i} + b_1 LSDOMY_{i,t} + b_2 LSEXTY_{i,t} + b_3 LSNATY_{i,t} + b_4 LPOPGPLUS_{i,t} + u_{i,t} \quad (1)$$

Domestic savings, external savings and net aid transfers are expressed as (log) ratios to GDP. Tables 1-5 provide estimates of this model, using different subsamples of the smaller "balanced" panel of 50 countries, while Table 6 presents estimates based on a larger panel of 131 countries. Table 7 presents the estimated effects of aid on investment, domestic savings and the real exchange rate. For the estimates reported in Tables 1-5 and 7, NDHKM apply the Dynamic GLS (DGLS) estimator, by Stock and Watson (1993).[4] This method involves adding *l* lagging and leading differences of all regressors to equation (1). Throughout their paper, NDKMH set *l=2*, without further elaboration on this choice.

Since all variables are subjected to a logarithmic transformation, it is required that the original series (in levels) are strictly positive. As it turns out, this is not the case and results in the omission of a large fraction of the available data. To illustrate this problem, we focus on the fourth column of Table 1 where the estimates of equation (1) are reported, using all the covariates, for a panel of 50 countries over the period 1960-2006. After adjusting the endpoints to the dynamic specification of the model, there are 41 observations available per country. A balanced panel should therefore include 41×50=2,050 observations. Our replication shows that the full model in the fourth column of Table 1 in NDKMH is based on only 755 observations, implying a loss of 63% of the available observations.

Consider, for example, the top-left plot in Figure 1, which depicts domestic savings, net external savings and net aid transfers, in levels, for Algeria. For all the three variables, full time series data over the period 1960-2006 is available. However, since net external savings are negative during several periods, these observations are lost after the logarithmic transformation (bottom-left plot).

Because equation (1) is supplemented with two leading and lagging differences of all regressors, at least six subsequent observations are required within a country, to include one observation for estimating the model. Making matters even more challenging, the DGLS estimator requires one additional observation for estimating ρ, the autocorrelation parameter of the residual term *u*. Therefore, in order to include observation *t* for country *c* in the

[4] The authors indicate that in estimating the impact of aid on growth their preferred approach is DGLS (the results reported from Table 1-5).

estimation, all variables need to be observed for seven periods, from period $t-4$ to $t+2$. As Figure 1 shows, this happens only once for Algeria, during 1984-90. As a result, t=1988 is the only observation from Algeria used for estimating equation (1). The right-hand side plots in Figure 1 tell a similar story for Swaziland. In levels, there are two short gaps in the observed data. After the logarithmic transformation, there is only one interval left during which all variables are observed for at least seven subsequent periods: 1986-92. The only observation from Swaziland used for estimating equation (1) is t=1990.

In addition to resulting in omission of observations, the logarithmic transformation of domestic savings and net external savings is questionable. In levels, Figure 1 shows a very clear co-movement between these variables. After taking logs, this information is completely lost. Given that the observations are not randomly omitted, but are systematically dropped for country-year pairs with non-positive savings values, the coefficient estimate of aid may potentially be underestimated. For a given level of aid, country-year pairs with non-positive saving values are cases where the returns to aid may be higher.

Although Algeria and Swaziland are the worst cases, the problem is widespread. Figure 2 shows the distribution of observations per country included in the estimation of equation (1). The full potential of 41 observations is realized in only one country (Egypt). In 37 out of the 50 countries, less than half the observations are actually used. All our tables show the number of observations used for estimation relative to the potential number of observations.

A similar critique also applies to the estimates based on the larger panel of 131 countries, which are reported in Table 6. We have been able to replicate all the results in this table except the sixth column. The issue of missing observations as the result of a failed logarithmic transformation applies here as well, as is evident from Figure 3, which shows the distribution of observations per country included in the estimation reported in the third column of Table 6. This estimation is based, on average, on only 1.5 observations per country, which is clearly insufficient to offer a time series perspective.

The regression model (1) is derived in NDHKM (p. 293) by log-linearizing a Cobb-Douglas production function in which income is the product of the inputs of domestic savings, net external savings and net aid transfers. Given that income is strictly positive, the inputs are required to be strictly positive as well, for this multiplicative relation to hold. Our finding of negative inputs therefore clearly reveals, in addition to the empirical problem of missing data, mis-specification of the theoretical model.

Overall, in our assessment, the reported results in NDHKM do not provide any evidence in favour or against the effectiveness of aid. Apart from the panels being highly unbalanced, the remaining observations are often not clustered, but scattered over the full potential sample, which makes time series analysis nearly impossible.

3 Empirical Strategy

Regardless of the issues related to data, estimating equation (1) by DGLS is not a suitable empirical strategy when the aim is to estimate the effect of aid on income. NDHKM acknowledge that aid is unlikely to be exogenous with respect to GDP. A proper application of instrumental variables can achieve the identification of the causal effect of aid on income in a single-equation framework. However, finding valid instruments, which are uncorrelated with the residuals but sufficiently correlated with the endogenous regressors to ensure strong identification, is often a daunting task (see Clemens and Bazzi (2009)). Clemens et al. (2012) also discuss the challenge of finding a reliable instrument for aid as a major problem in aid-growth empirical research. While lags of the endogenous regressors are often legitimate instruments, their use in this application would be invalid because the residuals from the co-integrating relationship are strongly autocorrelated. NDHKM list the difficulties associated with finding valid instruments as one of the reasons for using the DGLS estimator instead.

NDHKM cite the result of Stock and Watson (1993), who note that the DGLS estimator is unbiased even when the regressors are endogenous. If the variables are co-integrated, the DGLS estimator does indeed give unbiased estimates of the co-integrating vector. However, it is a serious misunderstanding that the parameters of a co-integrating vector can be interpreted as a causal effect. As indicated in Stock and Watson (2011, p.697), strict exogeneity of the regressors is required for such a causal interpretation. A co-integrating vector by itself does not reveal any direction of cause and effect and hence no conclusions on the "impact of aid on income" can be drawn based on the DGLS estimates reported by NDHKM.

Alternatively, a Vector Autoregression (VAR) can be applied, in which the dynamics of income, aid and possibly more variables are modelled jointly. Rather than identifying an instantaneous (static) causal effect, VAR models are able to show the dynamic intertemporal impact of a shock to one variable on the future path of another variable. Since aid is not necessarily supposed to improve income per capita immediately, but rather to improve conditions for growth in the longer run, a dynamic model such as a VAR provides a starting point for assessing the long-term impact of aid on income.

In the next section, we therefore present a Panel-VAR (PVAR) model for aid and income, estimated on the same dataset as NDHKM. By computing impulse response functions based on the estimated VAR, we analyze the dynamics of income over a period of 10 years following a shock to aid. With both aid and income treated as endogenous a priori, the VAR allows us in addition to explicitly consider a shock to income and its effects on aid.

VAR models have become the benchmark tool in macroeconometrics, for example for estimating the effects of monetary and fiscal policies.[5] For such applications a multi-equation model is attractive since fiscal and monetary policies do not only affect the performance of the economy, but the state of the economy itself is likely to have an impact on the policies.[6] The same argument applies to the relationship between aid and income. Surely, when donor countries make decisions regarding development aid, they take into consideration the economic conditions in the recipient countries (this is built into the aid allocation formula of the International Development Association (IDA), for example). When estimating the effect of aid on income, it is therefore essential to disentangle it from the effect of income on aid.

In a recent study, Juselius et al. (2012) estimate separate co-integrated VAR models for 36 African countries, which are supplemented with country-specific dummies to indicate periods of economic and political turmoil. In our approach the observations of all countries are instead pooled to estimate a PVAR with fixed parameters. Our PVAR is therefore a dynamic multiple-equation extension of the fixed-effects model considered by NDHKM. The advantage of pooling the data is that it dramatically increases the size of the dataset. Rather than estimating the country-specific VAR using T observations, we estimate a PVAR with $T \times N$ observations. In our case, T=37 and N=55. By assuming fixed parameters, there are many more observations available to estimate the parameters but this, of course, comes at a cost. Assuming constant parameters across countries can be highly restrictive, while country-specific dummies to account for extreme events are not included. We acknowledge these restrictions and emphasize that we estimate the average effects of aid for the reasons outlined in the introduction. In specific countries these effects may differ from the ones presented in Section 4.

Against this background, we aim to keep our model parsimonious. In order to provide an answer to the question raised by NDHKM ("Does foreign aid raise per capita income?"), we fit a bivariate VAR model to aid and GDP, which provides an alternative to the NDHKM single-equation regression model with aid as the only regressor (Table 1, 1st column), for which NDHKM report a negative correlation between aid and income. Following the convention in the aid literature, we transform both variables into logs. Juselius et al. (2012) report evidence in favour of a multiplicative rather than additive relationship between aid and income, which makes the logarithmic transformation required. Nevertheless, we avoid the problems with taking logs of negative numbers (discussed in Section 2), by excluding domestic and external savings from the model and by considering countries that are net aid-receivers only. As a robustness check, however, we also estimate

[5] See Caldara and Kamps (2008), Chung and Leeper (2007), Mountford and Uhlig (2009), Stock and Watson (2001) and the papers cited therein.
[6] For applications to fiscal policy, see Blanchard and Perotti (2002).

the VAR supplemented with domestic and external savings, while keeping both these variables in levels rather than logs.

NDHKM and Juselius et al. (2012) find aid and income to be co-integrated. It would therefore be appropriate to estimate a co-integrated VAR (CVAR), as Juselius et al. (2012) do on a country-by-country basis. We consider instead a VAR in levels, because the estimation of a CVAR on panel data with homogeneity of the co-integrating restrictions imposed across countries would be fairly complicated and, more importantly, unnecessary for the analysis we conduct in this paper. When the variables are co-integrated, the VAR in levels and CVAR are asymptotically equivalent. Clements and Hendry (1995) find that imposing co-integrating restrictions in an otherwise unrestricted VAR brings little gains when the objective is to forecast, which is essentially what we do by computing impulse-responses. Also the non-standard asymptotic distribution of the parameter-estimators in an unrestricted VAR (due to non-stationarity), is not an issue for our analysis, since our inference does not rely on the asymptotic distribution of the estimators. Instead, we compute confidence bounds for the impulse-response functions using a bootstrap simulation.

Unlike NDHKM, we do not use the aid-to-GDP ratio. Although considering aid as a ratio of GDP is not uncommon in the literature on aid effectiveness, it implies a certain restriction on the long-term relation between GDP, aid and population, which Juselius et al. (2012) test and reject for all the 36 African countries in their dataset. Moreover, transforming aid into a ratio of GDP makes it harder to identify the effect of aid on GDP. For example, consider a negative shock to GDP, which by construction reduces GDP per capita and raises the aid-to-GDP ratio. In the model by NDHKM, this negative co-movement between the regressor and explanatory variable would be interpreted as a negative effect of aid on GDP, even if the original shock to GDP could be entirely unrelated to aid. Our VAR model therefore includes aid and GDP in levels per capita, while we also examine the robustness of our results by considering the aggregate levels of both aid and GDP.

After estimating the model, we compute impulse response functions. In order to identify the shocks, we impose a recursive structure, which makes the order of the variables relevant. Since Sims (1980) it is in this literature generally considered sufficient (and even preferable) to provide a loose/intuitive justification for the ordering of variables, rather than to formulate an exact structural economic theory. For example, Caldara and Kamps (2008) argue that there is in general a considerable delay between political decisions on government spending and the actual spending. Macroeconomic conditions therefore have an impact on government spending only after a lag, while the reverse effect may occur immediately. The same argument can be applied to spending on aid.[7] In our VAR, aid is therefore placed before GDP. We acknowledge that relying on this recursive identification approach has limitations. Nevertheless, our methodology is arguably superior to the approach applied by NDHKM in terms of addressing the endogeneity issue which is inherent in aid-growth empirical work. As we are mainly interested in the long-term impact of aid, the relevance of the recursive order is actually rather limited. Moreover, given that we consider a VAR of only two variables, it is relatively straightforward to check the sensitivity of the results with respect to our assumptions by simply reversing the order of the variables, which we do in the next section as one of our robustness checks.

A potential shortcoming of VAR models is that, unlike a correctly specified structural model, it does not always reveal the mechanism through which the effects occur even if it provides an empirical description of the dynamic interaction between the variables. However, by supplementing the VAR with the appropriate variables, one can gain insight into how these mechanisms may work. For example, Osei et al. (2005) investigate the effects of aid on fiscal variables, Hansen and Headey (2010) consider trade-balances, while Kang et al. (2012) add exchange rates to their VAR. Since our main purpose is to illustrate how VAR–based results differ from the single-equation framework applied by NDHKM, we consider a bivariate VAR including only aid and GDP, although we do include domestic and external savings as a robustness check.

4 VAR results

We apply the exact same dataset as NDHKM, from which we obtain net aid receipts per capita (*aid*) and income per capita (*gdp*). The dataset features observations on 131 countries, for a maximum of 47 periods (1960-2006). We balance the panel, by including only countries for which the entire time-series of observations are available. With a starting date of 1960, we can include 44 countries in a balanced panel. When we postpone the starting date to 1970, the number of available countries increases to 55. With 1980 as the starting date, there is sufficient data for 67 countries. We choose the middle ground here, with T=37 (1970-2006) and N=55. Table 8 provides a list of countries. In addition to increasing the number of available countries, excluding the early years of aid data can be justified for data quality reasons (e.g. Juselius et al., 2012). At the end of this section, we verify the robustness of our results by varying the starting date to 1960 and 1980.

We estimate the following Panel VAR:

[7] That is, since donors have to observe the GDP shock in the recipient country before making a political decision and allocate aid, it is reasonable to assume that aid allocation occurs with some lags after the GDP shock takes place. On the other hand, the potential effect of aid on GDP can be expected to start improving the conditions for growth right away.

$$y_{it} = \mu_i + \sum_{k=1}^{p} A_k y_{it-k} + \varepsilon_{it}, \qquad i = 1 \dots N, \qquad t = 1 \dots T, \qquad (2)$$

in which $y_{it} = (aid_{it}, gdp_{it})$ gives the state of the variables in country *i* during period *t*, μ_i is a 2×1 country-specific intercept term, A_k are 2×2 matrices of coefficients, ε_{it} is a 2×1 residual term and *p* denotes the number of lags.

In estimating the PVAR we follow common practice by combining first-differencing and GMM estimation, applying lagged values as instruments (Arellano and Bond, 1991). First-differencing the model (2) eliminates the country-specific intercept, thereby avoiding the problem of inconsistency of the fixed-effects estimator for dynamic panel data regressions (Nickell, 1981). The application of lagged values as instruments is valid when the residuals ε_{it} are not autocorrelated. Table 9 shows Akaike's information criteria (AIC), Bayesian information criteria (BIC) and Durbin-Watson statistics (DW) for the bivariate VAR with one up to six lags. The BIC is minimized for a one-lag VAR, but in this case the DW statistics suggest the residuals are autocorrelated. Relying on the AIC, the selected number of lags is two, in which case there is no indication of autocorrelation, such that the use of lagged values as instruments is appropriate. We therefore choose *p=2*.

Moving to the results, Figure 4 shows the orthogonalized impulse response functions, with the dynamic effects over 10 years on aid (left) and income (right) following a positive shock to either aid (top row) or GDP (bottom row), with 90% bootstrap confidence bounds based on 100,000 replications.

Figure 4.b illustrates the effect of a one unit positive shock to log-aid (i.e. the effect of a 1% increase in aid receipts) on log-income. In the short run (0-2 years), the impact of aid is insignificant. In the long run, however, the impulse response in Figure 4.b demonstrates a clearly positive and significant response. The observed difference between short- and long-term impacts demonstrates why a dynamic model structure, like in a VAR, is crucial for evaluating these effects. A static model, like Eq. (1) only considers the instantaneous impact, but is unable to capture the long-term effect, i.e. the impact on income, multiple years after receiving aid.

Moreover, Figures 4.a-b show that a shock to aid is in itself transitory, while its effect on income seems rather persistent. These results suggest that a temporary increase in aid spending can push income to a permanently higher level, which is certainly a more positive assessment of aid than reported by NDHKM. Our results are in line with other recent VAR-based analyses of aid-effectiveness. Gillanders (2011) fits a fixed-effects PVAR to aid-per-capita and GDP growth and finds that aid has a significant positive, although small, effect on GDP. Juselius et al. (2012) estimate country-specific co-integrated VARs for aid, income and other macroeconomic variables and find a positive effect of aid for most of the countries.

Figure 4.c illustrates how the use of a single-equation framework may lead to confusion about the impact of aid. A shock to income has an estimated negative and persistent effect on aid, which is (although not statistically significant) of larger magnitude than the positive effect of aid on income. Both the intertemporal effects in Figure 4.b-c play a role in the long-run. Given the opposite signs of these effects, and the larger size of the negative effect, it should come as no surprise that a negative and/or insignificant long-run relation between aid and income is found using a single equation framework, even if the impact of aid on income is positive and significant.

Figure 4.b, showing the intertemporal effects of aid on income, is reproduced in Figure 5 based on several alternative VAR specifications, which are listed in Table 10. In Figure 5.a-b, the starting date is set to 1960 and 1980, respectively. This leads not only to a different length of the time-series, but also to the inclusion or exclusion of certain countries. In Figure 5.c, we consider aggregate aid levels instead of aid per capita, while in Figure 5.d both aid and income are measured in aggregate levels. In Figure 5.e, we reverse the order of the variables, placing GDP before aid. The main result, a positive effect of aid on income in the long run, is robust to all these alternatives. Although the long-term impact of aid is not in all cases significant at the 10% level, the confidence intervals do lie almost entirely in the positive domain. Comparing Figure 4b and Figure 5.a-b further shows that confidence intervals have clearly narrowed over time, suggesting that the evidence has become more decisive towards a positive impact of aid.

In Figure 5.f, we take first-differences of aid and GDP (both in levels per-capita). A shock to aid seems to have a positive short-run effect on differenced income, which converges to zero after some periods. The gradual decline of the impact is consistent with the decreasing slope of the impulse response function for the VARs in levels. In the final two columns of Figure 5, domestic and external savings per capita (sav_{it} and ext_{it}) are added to the VAR. Since we cannot take logs of domestic and external savings (which would lead to the problems raised in Section 2), we consider the model first with aid and GDP measured in logs and domestic and external saving in levels (Figure 5.g), and second with all four variables measured in levels (Figure 5.h). In both these cases, the estimated VARs become unstable, presumably due to the inclusion of variables in levels instead of logs, while the underlying relation is multiplicative. As a result of this instability, confidence bounds cannot be computed. We therefore report these impulse response functions without confidence bounds, and interpret this as a strengthening of our argument in Section 2 that the Solow-type model including domestic and external savings (Eq. 1) is not correctly specified. Finally, Figure 5.i is based on the VAR with all four variables measured in differences. The resulting impulse response function looks similar to Figure 5.f.

Although providing a definitive answer to the aid effectiveness question is not the primary objective this study, we believe that the results from this exercise can, with some caution, be considered as indicative time series evidence on aid effectiveness. Overall, the results presented in Figures 4 and 5 indicate consistently a positive long-term impact of aid on income, which is in stark contrast to the results reported by NDHKM, even though the results are based on the same dataset.

5 Conclusion

The main purpose of this study was to illustrate how inappropriate applications of time series techniques including data mishandling, model mis-specification and choice of inapproriate models and/or misreading of results can lead to misguided conclusions and inferences regarding the effectiveness of foreign aid. More specifically, we have demonstrated how a system of equations approach based on VAR models addresses the well-recognized issue of endogeniety in aid-growth analysis. In the process, we have also shown how the single equation approach applied in NDHKM is not well suited to handle the aid-growth relationship.

In light of the serious problems related to data handling and usage of time series techniques we uncovered in the replication exercise, we argue that the evidence in NDHKM does not provide a sound time series perspective on aid and growth. Even when appropriate methodology and data is applied, insignificant results, in the terminology of Temple (2010), only amounts to ‘absence of evidence’ and should not be confused for ‘evidence of absence’ of the effect of aid on income.

When a Panel VAR model is applied to the same dataset as in NDHKM, which better addresses the fact that causality in the aid-growth relationship runs in both directions with potentially opposing effects, a positive and statistically significant long run effect of aid on growth emerges. This result is consistent with other emerging time-series evidence.

References

Arndt, C., S. Jones, and F. Tarp (2010) 'Aid, Growth, and Development: Have We Come Full Circle? ,' *Journal of Globalization and Development*, *1*(2), 5.

Arellano, M., and S. Bond (1991) 'Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations,' *The Review of Economic Studies* 58, 277-297.

Bhargava, A., Franzini, L., & Narendranathan, W. (1982). Serial correlation and the fixed effects model. The Review of Economic Studies, 49(4), 533-549.

Blanchard, O., and R. Perotti (2002) 'An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output,' *The Quarterly Journal of Economics 117* (4), 1329–1368.

Caldara, D., and C. Kamps (2008) 'What are the effects of fiscal shocks? A VAR-based comparative analysis,' Working Paper Series 877, European Central Bank.

Chung, H., and E. Leeper (2007) 'What Has Financed Government Debt? ,' CAEPR Working Papers 2007-015, Center for Applied Economics and Policy Research, Economics Department, Indiana University Bloomington.

Clemens, M. A., S. Radelet, R. R. Bhavnani, and S. Bazzi (2012) 'Counting Chickens when they Hatch: Timing and the Effects of Aid on Growth,' *Economic Journal 122*(561), 590–617.

Clemens, M. and S. Bazzi (2009) 'Blunt Instruments: On Establishing the Causes of Economic Growth,' Working Papers 171, Center for Global Development.

Clements, M. P., and Hendry, D. F. (1995). Forecasting in cointegrated systems. Journal of Applied Econometrics, 10(2), 127-146.

Gillanders, R. (2011) 'The Effects of Foreign Aid in Sub-Saharan Africa, *School Of Economics, University College Dublin Working Papers*.

Hansen, H., and D. Headey (2009) 'The Short-Run Macroeconomic Impact of Foreign Aid to Small States: An Agnostic Time Series Analysis,' *The Journal of Development Studies* 46 (5), 877–896.

Hansen, H., and F. Tarp (2000) 'Aid effectiveness disputed,' *Journal of International Development 12* (3), 375–398.

Juselius, K., N. Framroze-Møller, and F. Tarp (2012), "The Long-Run Impact of Foreign Aid in 36 African Countries: Insights from Multivariate Time Series Analysis," *Oxford Bulletin of Economics and Statistics*, forthcoming.

Kang, J., A. Prati, and A. Rebucci (2012) 'Aid, Exports, and Growth: a Time-Series Perspective on the Dutch Disease Hypothesis,' *Review of Economics and Institutions 3*(2).

Mountford, A., and H. Uhlig (2009) 'What are the effects of fiscal policy shocks?,' *Journal of Applied Econometrics 24* (6), 960–992.

Moyo, D. (2009) *Dead aid: why aid is not working and how there is another way for Africa*, Allen Lane.

Nickell, S**.** (1981) 'Biases in dynamic models with fixed effects, *Econometrica* 49(6), 1417-1426.

Nowak-Lehmann, F., A. Dreher, D. Herzer, S. Klasen, and I. Martínez-Zarzoso (2012) 'Does foreign aid really raise per capita income? A time series perspective,' *Canadian Journal of Economics* 45(1), 288–313.

Osei, R., O. Morrissey, and T. Lloyd (2005) 'The fiscal effects of aid in Ghana,' *Journal of International Development* 17(8), 1037–1053.

Rajan, R. G., and A. Subramanian (2008) 'Aid and Growth: What Does the Cross-Country Evidence Really Show?,' *The Review of Economics and Statistics* 90(4), 643–665.

Sims, C.A. (1980) 'Macroeconomics and reality,' *Econometrica* 48(1), 1–48.

Stock, J. H., and M. Watson (1993) 'A simple estimator of cointegrating vectors in higher order integrated systems," *Econometrica* 61(4) , 783–820.

Stock, J. H., and M. W. Watson (2001) 'Vector Autoregressions,' *Journal of Economic Perspectives* 15 (4), 101–115.

Stock, J. H., and M. Watson (2011) *Introduction to Econometrics* Addison-Wesley series in economics, Pearson Education.

Temple, Jonathan R.W. (2010) 'Chapter 67 - Aid and Conditionality,' in D. Rodrik and M. Rosenzweig, eds., *Handbooks in Economics*, Vol. 5 of *Handbook of Development Economics*, Elsevier, 4415 – 4523.

Wooldridge, J. M. (2001) *Econometric analysis of cross section and panel data*, MIT press.

Tables and Figures

TABLE 1: Impact of aid on income

Dependent variable:	*LY*	*LY*	*LY*	*LY*
LPOPGPLUS	.	.	.	0.00
LSDOMY	.	0.08	0.07	0.07
LSEXTNY	.	.	0.04	0.05
LSNATY	-0.02	-0.01	-0.01	-0.02
ρ	0.97	0.97	0.98	0.99
N	57	56	50	50
T	41	41	41	41
K	2120	1693	794	755
*K/(N*T)*	0.91	0.74	0.39	0.37

Notes: Estimates of equation (1). t-values are identical to NDHKM and therefore not reported. *N* refers to the cross sectional dimension (amount of countries), *T* to periods and *K* to amount of observations used for estimation.

TABLE 2: Differing impact depending on aid-to-GDP ratio

Aid-to-GDP ratio	Above average	Below average
Dependent variable:	*LY*	*LY*
LPOPGPLUS	0.04	0.37
LSDOMY	0.05	0.16
LSEXTNY	0.04	0.06
LSNATY	-0.03	-0.01
ρ	0.98	0.99
N	23	27
T	41	41
K	343	412
*K/(N*T)*	0.36	0.37

Notes: See Table 1

TABLE 3: Differing impact depending on HDI

HDI	<0.5	0.5-0.799	>0.8
Dependent variable:	*LY*	*LY*	*LY*
LPOPGPLUS	-0.53	0.43	0.68
LSDOMY	0.06	0.09	1.91
LSEXTNY	0.02	0.05	-1.01
LSNATY	-0.03	-0.01	-0.17
ρ	0.97	1.00	0.35
N	20	25	4
T	41	41	41
K	303	413	30
*K/(N*T)*	0.37	0.40	0.18

Notes: See Table 1

TABLE 4: Differing impact depending on income level

Income level	LDC	GNI<735	736<GNI<9075
Dependent variable:	*LY*	*LY*	*LY*
LPOPGPLUS	-0.23	-0.30	0.23
LSDOMY	0.05	0.06	0.18
LSEXTNY	0.08	0.05	0.06
LSNATY	-0.01	-0.02	-0.01
ρ	0.97	0.98	0.99
N	18	24	24
T	41	41	41
K	295	397	321
*K/(N*T)*	0.40	0.40	0.33

Notes: See Table 1

TABLE 5: Differing impact depending on region

Region	Caribbean	Latin America	Africa	Asia
Dependent variable:	*LY*	*LY*	*LY*	*LY*
LPOPGPLUS	2.87	0.58	-0.10	-0.51
LSDOMY	0.16	0.12	0.06	0.02
LSEXTNY	0.06	0.06	0.04	0.02
LSNATY	-0.04	-0.02	-0.01	-0.03
ρ	0.98	0.95	0.96	1.01
N	5	11	25	6
T	41	41	41	41
K	69	117	356	136
*K/(N*T)*	0.34	0.26	0.35	0.55

Notes: See Table 1

TABLE 6: Impact of aid on income (sample of 131 countries)

Method	FE	FE	FE+GLS	GMM	GMM	SUR
Data	Annual	5y-average	5y-average	5y-average	5y-average	5y-average
Dependent variable	*LY*	*LY*	*LY*	*LY*	*LY*	*LY*
LPOPGPLUS	-0.12	-0.08	0.17	0.37	0.28	.
LSDOMY	0.09	0.10	0.02	0.04	0.01	.
LSEXTNY	.	0.01	0.01	0.01	0.01	.
LSNATY	-0.06	-0.05	-0.02	-0.02	-0.02	.
N	131	131	131	131	131	.
T	41	8	8	8	8	.
K	1728	346	198	198	115	.
*K/(N*T)*	0.32	0.33	0.19	0.19	0.11	.

Notes: See Table 1

TABLE 7: Indirect impact of aid

Dependent variable:	Investment	Domestic Savings	Real exchange rate
LPOPGPLUS	0.42	.	.
LSDOMY	0.29	.	-0.15
LSEXTNY	0.04	-0.12	-0.51
LSNATY	0.54	0.58	0.72
ρ	50	56	20
N	41	41	41
T	795	1915	327
K	0.39	0.83	0.40
*K/(N*T)*	0.42		

Notes: See Table 1

TABLE 8: Countries

Algeria[60]	Cote d'Ivoire[60]	Kenya[60]	Philippines[60]
Argentina[70]	Dominica[80]	Lesotho[60]	Rwanda[60]
Bangladesh[80]	Ecuador[60]	Liberia[60]	Saudi Arabia[70]
Belize[60]	Egypt[60]	Madagascar[60]	Senegal[60]
Benin[60]	El Salvador[60]	Malawi[60]	Seychelles[60]
Bhutan[80]	Gambia[70]	Mali[70]	Sierra Leone[60]
Bolivia[60]	Ghana[60]	Mauritania[60]	Sri Lanka[60]
Botswana[60]	Grenada[80]	Morocco[60]	Sudan[60]
Burundi[60]	Guatemala[60]	Mozambique[80]	Suriname[80]
Cameroon[60]	Guinea[80]	Nepal[60]	Swaziland[70]
Central African Rep.[60]	Guinea-Bissau[80]	Nicaragua[60]	Syria[70]
Chad[60]	Guyana[60]	Niger[60]	Togo[60]

China[80]	Haiti[60]	Nigeria[60]	Tunisia[70]
Colombia[70]	Honduras[60]	Pakistan[60]	Turkey[70]
Comoros[80]	India[60]	Panama[60]	Uruguay[70]
Congo, D.R. [60]	Indonesia[60]	Paraguay[60]	Venezuela[80]
Congo, R. [60]	Jordan[80]	Peru[70]	

Notes: Countries included in VAR analysis.

(60): Countries included in datasets 1960-2006, 1970-2006 and 1980-2006

(70): Countries included in datasets 1970-2006 and 1980-2006

(80): Countries included in dataset 1980-2006

TABLE 9: VAR model specification

p	1	2	3	4	5	6
AIC	0.80	0.72	0.83	0.98	1.08	1.18
BIC	1.12	1.36	1.79	2.26	2.68	3.11
d_{AID}	2.36	1.98	1.99	1.91	1.95	1.98
d_{GDP}	1.49	1.81	1.81	1.98	1.97	1.96

Notes: Akaike's information criteria, Bayesian information criteria and Durbin-Watson statistics for both residual series. Panel-VAR(p) process for Aid and GDP per capita (both in logs), with T=37 (1970-2006) and N=55. See Bhargava et al. (1982) for details on Durbin-Watson statistics for panel data.

TABLE 10: Robustness checks

	y	T	N	Per capita	Logs
(*)	(aid, gdp)	37 (1970-2006)	55	yes	yes
(a)	(aid, gdp)	47 (1960-2006)	44	yes	yes
(b)	(aid, gdp)	27 (1980-2006)	67	yes	yes
(c)	(aid, gdp)	37 (1970-2006)	55	GDP only	yes
(d)	(aid, gdp)	37 (1970-2006)	55	No	yes

(e)	(gdp, aid)	37 (1970-2006)	55	yes	yes
(f)	(Δaid, Δgdp)	37 (1970-2006)	55	yes	yes
(g)	(aid, gdp, ext, sav)	37 (1970-2006)	41	yes	Aid and GDP only
(h)	(aid, gdp, ext, sav)	37 (1970-2006)	41	yes	no
(i)	(Δaid, Δgdp, Δext, Δsav)	37 (1970-2006)	41	yes	Aid and GDP only

Notes: VAR specifications. (*): Benchmark model with impulse response functions presented in Figure 4. (a)-(i):

Alternative VAR specifications presented in Figure 5 (a)-(i). T: Time-series dimension. N: Cross-sectional dimension.

Final two columns indicate whether variables are measured in aggregates or per capita terms, and in logs or levels.

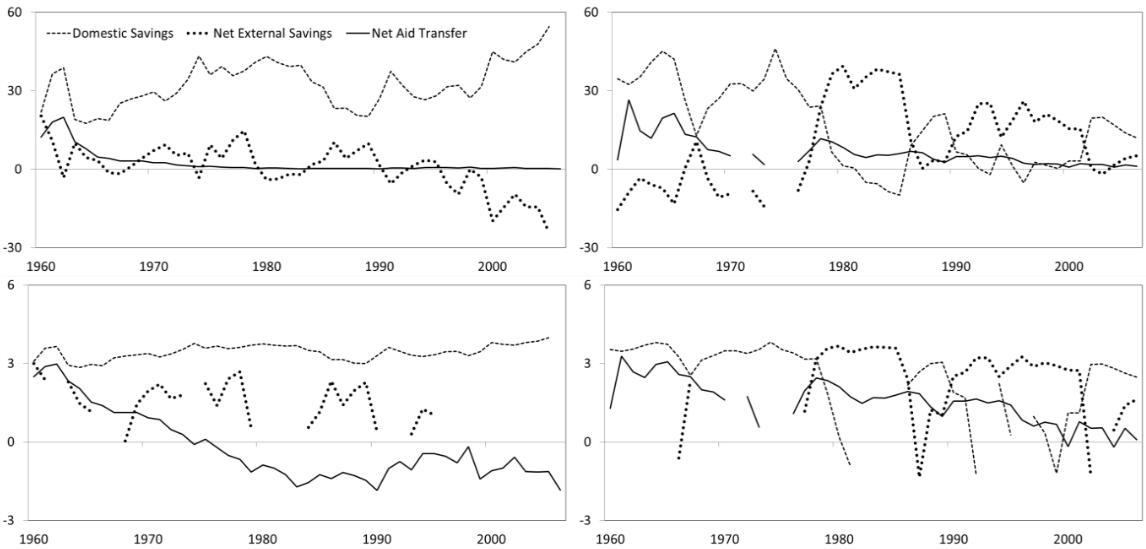


Figure 1: Domestic Savings, Net External Savings and Net Aid Transfer for Algeria (left) and Swaziland (right), in levels (top) and logs (bottom), for the period 1960-2006

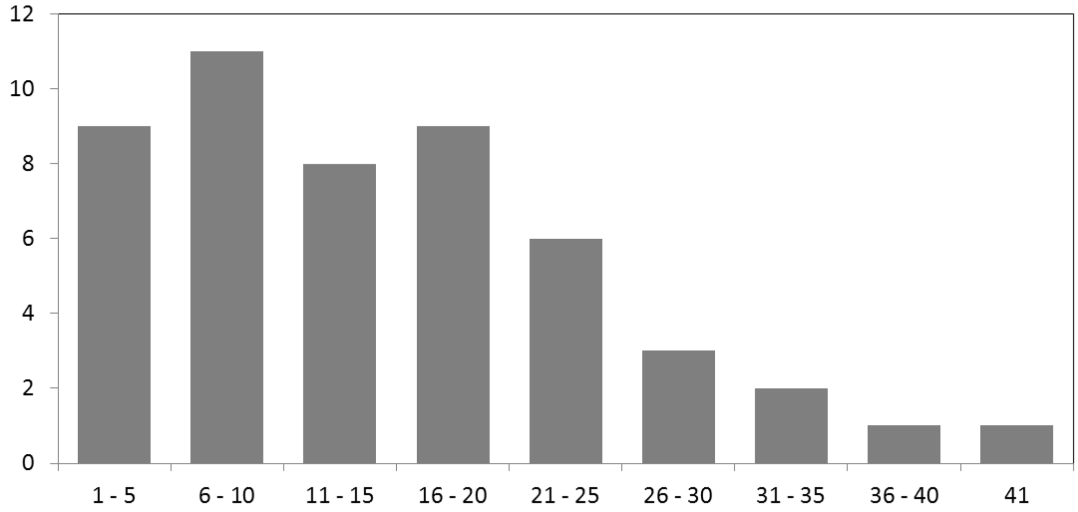


Figure 2: Distribution of included observations for the estimation of equation (1): Table 1, 4th column. Histogram depicts amount of countries (y-axis) with amount of included observations (x-axis).

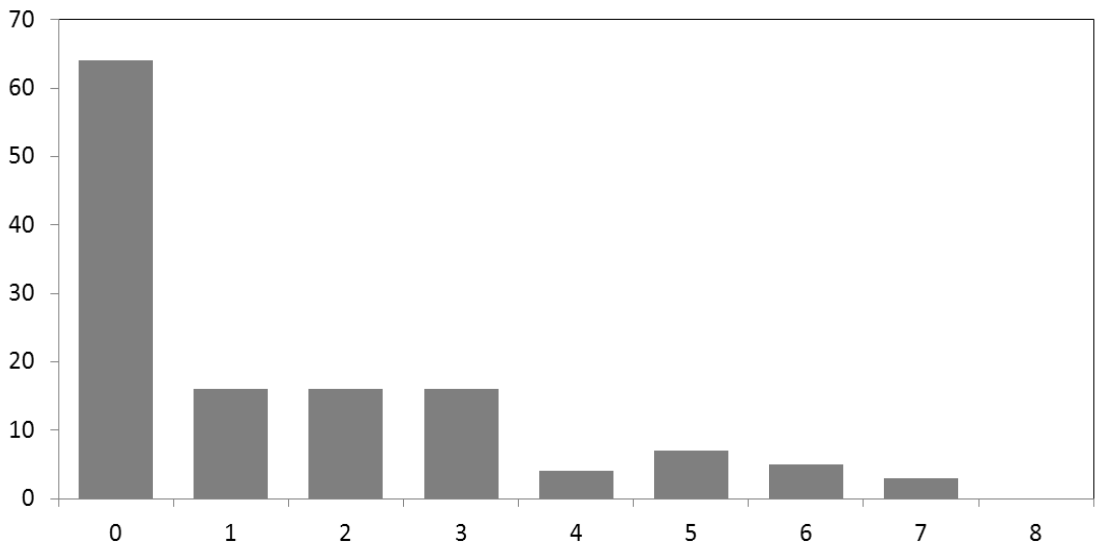


Figure 3: Distribution of included observations for the estimation of equation (1): Table 6, 3rd column. Histogram depicts amount of countries (y-axis) with amount of included observations (x-axis).

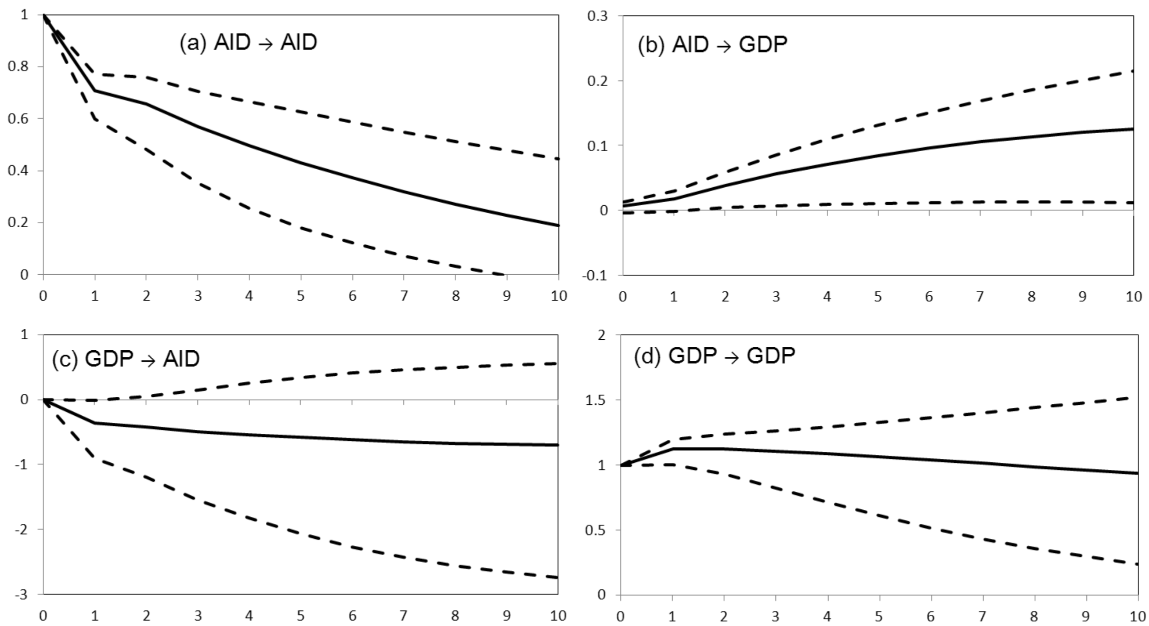


Figure 4: Impulse response functions based on PVAR (2) with *p=2* for log-Aid per-capita and log-GDP per-capita. *T*=37 (1970-2006) and *N*=55. 90% bootstrap confidence bounds based on 100,000 replications.

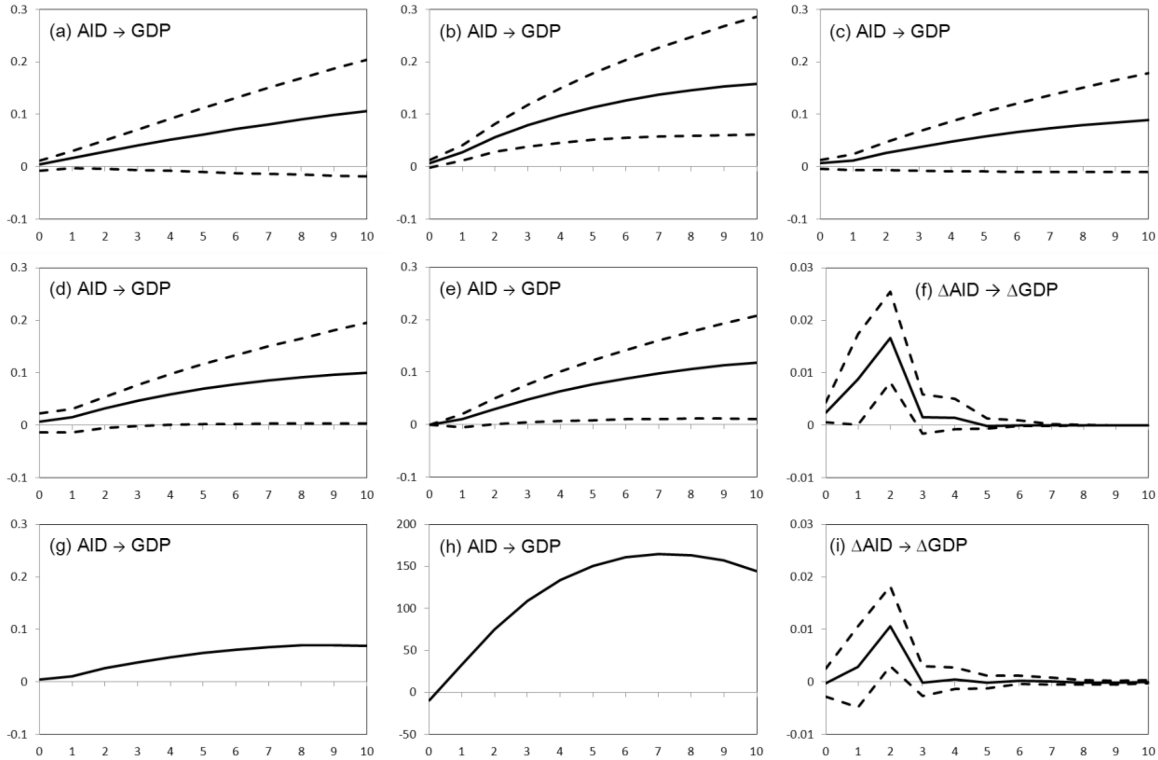


Figure 5: Robustness checks for Figure 4.b See Table 10 for details on the alternative VAR specifications. 90% bootstrap confidence bounds based on 100,000 replications.