PhD Thesis No. 159

Dynamic Aspects of Entrepreneurial Behavior

by

Bertel Schjerning

# DYNAMIC ASPECTS OF ENTREPRENEURIAL BEHAVIOR

By

Bertel Schjerning

PhD Thesis
Submitted to
Department of Economics
University of Copenhagen
October 2007

# TABLE OF CONTENTS

Acknowledgement

Introduction and Summary

*A special thanks to my wife Kjelfred and my son Viggo. When I go off to conferences for several days, work late nights, or don't answer your questions because my thoughts are elsewhere, you have been more patient with than I could ever ask for. This is more appreciated than you think. I know this thesis surely makes little sense to you. However, with out you being around me every day – it would not make sense to me either. I therefore dedicate this work to you.*

*Bertel Schjerning*
*Copenhagen, August 2007*

Introduction and Summary

The thesis consists of four chapters. They are self-contained; and can be read independently. Each of the four chapters is connected to dynamic aspects of entrepreneurial behavior and the effects of uncertainty related to future states of the world. Current decisions faced by entrepreneurs are usually made under uncertainty and can have long-run implications. Therefore, intertemporal incentives underlie much of the behavior by entrepreneurs. In this thesis, I aim contributing to an improved understanding of the importance of intertemporal incentives and the role of uncertainty for the decision to become and entrepreneur and the subsequent behavior. The thesis is structured as follows:

Chapter 1 presents an empirical analysis based on micro data, considering how existing earnings differentials and differences in income uncertainty can explain observed occupational choices, i.e. the decision to become an entrepreneur or a wage worker.

Chapter 2 develops a theory of occupational choice in the presence of credit constraints. The model developed explicitly incorporates intertemporal incentives with respect to both the occupational choice and wealth accumulation and takes into account that future income is uncertain.

Chapter 3 investigates a source of human capital accumulation that takes place via employment in high-productivity firms, in particular, foreign-owned firms. This is hypothesized to give rise to subsequent productivity transfers when workers later move to self-employment (entrepreneurship) or alternative wage employment.

Chapter 4 develops useful models and techniques to solve and estimate dynamic structural models of discrete choices. As such, this chapter is not specific to entrepreneurial decision making, but it was inspired by the requirement to take a more structural empirical approach than in Chapter 1, where the link between a behavioral model and the econometric specification admittedly could be more explicit. Specifically, the original objective of Chapter 4, was to develop a procedure to solve and estimate a dynamic structural model for the decision to become an entrepreneur, allowing for lots of observed and unobserved heterogeneity in, e.g., individual specific skill endowments (entrepreneurial ability), and serial correlation in the income processes faced by workers and entrepreneurs. This will be a good stepping stone for doing a full structural estimation of the (more complicated) model in Chapter 2. This is an interesting project that is still subject to ongoing research.

A more detailed summary of the content of the four chapters is given below.

CHAPTER 1

Chapter 1 on earnings, uncertainty, and the self-employment choice, is co-authored with a fellow PhD student Daniel le Maire. The main objective of this Chapter is to investigate the relationship between the occupational choice and the distributions of associated monetary gains in different occupations. Specifically, we analyze how existing earnings differentials and differences in income uncertainty can explain observed occupational choices. A particular focus is on explaining why fewer women choose to become self-employed.

We use a large longitudinal data set based on Danish register data from the Integrated Database for Labor Marker Research (IDA) covering the period 1980 to 1996. These data provide detailed annual individual information about income, wealth, education, labor market status (occupation), region of residence, and immigration status. Since the panel covers more than 15 years, we can track long sequences of individual occupational choices.

To evaluate whether the self-employed are compensated for their risk-taking, individual level information about the expected income (and the expected distribution of income) in self-employment and wage-employment is required. To obtain this information, we estimate earnings functions for both self-employment and wage-employment. These functions are then in turn used to predict an individual's income (and the uncertainty of this income) in different occupations. The random components of the model are partitioned into transitory and permanent chocks, which in turn are used to create occupational and education specific measures of income variance (uncertainty) and skewness (the risk/chance of very low/high incomes).

This information is used to explain the observed occupational choices of the individuals. Specifically, individuals are assumed to choose between three labor market states: self-employment, wage-employment and unemployment. The underlying assumption is that the occupational choice reflects an optimal trade-off between the different (earnings) attributes of the three occupations. Since we focus in particular on the trade-off between expected returns and uncertainty, we include not only predicted earnings in the occupational choice model, but also the corresponding measures of income variance and skewness in the different occupations.

Several interesting results emerge from our analysis on Danish longitudinal register data. First, the (taxable) of self-emplyed displays a concentration around the kink points in the tax system since self-employed can retain earnings and thereby transfer income across years. Second, when comparing earnings distributions based on different income measures, we find that the dispersion of incomes is in general much larger for the self-employed. Third, our results suggest that expected income and income uncertainly are important determinants in the choice of occupation. As expected, people are in general risk averse, preferring high mean earnings with low variances. However, this is much more pronounced among women and may therefore explain their limited participation in self-employment, as this occupation displays much more income uncertainty. Finally, our results suggest that nonwestern immigrants are marginalized into self-employment.

CHAPTER 2

Other dynamic considerations are likely to play an important role for the individual choice of becoming an entrepreneur. That is, individuals do not only compare earnings profiles in wage work and entrepreneurship when deciding on their occupations. For instance, future expected earnings in the different occupations, and − at least in entrepreneurship − are likely to depend on both accumulated human capital and savings. In the absence of external financing, savings may thus determine how much the entrepreneur can invest in the firm. This creates a complicated decision problem, in part because savings are affected by current and future income.

Chapter 2 formalize these aspects to provide a better understanding of observed occupational choices by individuals, in particular, the observed transitions between wage work and entrepreneurship. The model in Chapter 2 is an intertemporal model of saving, consumption, human capital accumulation and occupational choice in the presence of liquidity constraints, income uncertainty, and entry costs. Self-employment yields an uncertain return which depends on both the human capital (labor market experience) of the individual and the capital invested in the company. Wage employment, on the other hand, also yields an uncertain return, but this does not depend on capital investments and human capital. Human capital (experience) is accumulated in both wage work and

entrepreneurship, but does only affect returns in entrepreneurship. Entry into entrepreneurship, however, is associated with a fixed, sunk start-up cost, which is lost completely if the business must close down and the agent returns to wage work.

The individual is credit constrained which implies that the fixed start-up cost must be financed out of her savings. Furthermore, the subsequent investments in physical capital in the firm are also limited by the amount of individual savings – but these investments can at least be recouped by selling the acquired capital upon closing the company and returning to wage work. Together, these assumptions reflect that some investments are irreversible (the entry costs) while others can be reversed (the physical capital) – at least at some cost.

Thus in each period, the agent not only decides on her occupation – i.e. whether to be wage employed or entrepreneur – but she also decides on the division of her current resources (her wealth) between saving (which includes investments in physical capital), consumption and possibly entry costs in each period.

A prominent feature of the model is that it generates a well-defined transition pattern between entrepreneurship and ordinary wage work. This feature is consistent with empirical findings of sequential entries and exits. The model predicts that workers may transit back and forth between wage employment and entrepreneurship. The latter occupation becomes more attractive as the worker accumulates sufficient wealth (and human capital) as this gives her a higher (expected) return in entrepreneurship because she can then acquire more physical capital. Wage work, on the other hand, becomes relatively more attractive when individual savings are depleted, e.g., following a series of negative shocks to entrepreneurial income which forces the individual to use her savings (i.e., sell her capital investments) to maintain consumption.

Second, the credit constraints induce entrepreneurs to accumulate savings and may thus explain why we empirically observe a concentration of assets among the entrepreneurs. When entrepreneurs are credit constrained, the accumulated savings determine how much physical capital they can acquire and thus the expected pay-off from entrepreneurship. Furthermore, entrepreneurs save to maintain their position as entrepreneurs to avoid potential costs associated with later re-entry. A result related to policy initiatives is that the entrepreneurial saving motive is affected by the tightness of credit constraints in a

nonmonotone way: Initially, as credit constraints become less binding, potential high-productivity entrepreneurs with relatively low asset holdings will find it optimal to save more. The reason is that the perspectives for (future) entrepreneurship become better for this particular group when it is possible to borrow for investments in larger projects. As credit constraints are further loosened, and eventually become irrelevant, the incentive to save for entrepreneurial reasons disappears.

Third, in the absence of entry costs, the probability of exiting entrepreneurship will generally be declining in the age of the firm. The reason is that entrepreneurs accumulate assets (savings) over time, making them more resistant to negative chocks. In the presence of entry costs, however, the probability of exit from entrepreneurship is initially increasing in the elapsed duration and then decreasing. Entry costs alter the transition patterns between entrepreneurship and ordinary wage work fundamentally: Due to entry costs, individuals will never enter entrepreneurship if there is a significant risk that they will not be able to maintain their business in at least the following period or two. Therefore, the exit probability is very low for short durations. For longer durations, however, the entrepreneurs become more resistant to negative chocks due to a larger amount of accumulated capital. When this effect dominates, the probability of exit starts to decline again.

In sum, the model in Chapter 2 has several predictions which are consistent with observed behavior. Furthermore, the model adds to our understanding of the effects of credit constraints on observed entrepreneurial behavior as well as the potential consequences of government intervention in this area.

CHAPTER 3
Chapter 3 on foreign owned firms, productivity transfers, and entrepreneurship is co-authored Associate Professor Nikolaj Malchow-Møller and Professor James R. Markusen.

This Chapter follows the idea outlined in Chapter 2 that entrepreneurial skills are – at least partly – derived through labor market experience. More precisely, it investigates – both theoretically and empirically – a specific source of such human capital accumulation; namely the accumulation that takes place via employment in high-productivity firms, in particular, foreign-owned firms. This is hypothesized to give rise to subsequent

productivity transfers when workers later move to self-employment/entrepreneurship or alternative wage employment.

Why are foreign-owned firms hypothesized to be important channels of productivity transfers? Because these firms are often believed to possess superior knowledge, production technologies, or management techniques compared to the average domestic firm. Empirically, it has also been documented in a number of studies that different firms pay different wages. In particular, the literature has found higher wages in foreign-owned firms and larger firms.

However, it remains an open question how these higher wages arise. Do they reflect actual labor productivity differences between the firms, or are they just a premium paid by these firms to their workers? Yet another possibility is that they merely reflect a better selection of workers by these firms. And if they reflect productivity differences are these then transferable to subsequent employments in other firms and/or to self-employment? In other words, do they reflect a higher degree of learning/human capital accumulation in these firms? This chapter offers a theoretical explanation for these observations which builds on Melitz's (2003) model of an industry structure with heterogeneous firms extended with a learning-on-the-job model from Ethier and Markusen (1996). The paper thus provides a theory of an important channel for productivity transfers to domestic firms and self-employment which is based on human capital accumulation and worker mobility instead of externalities. Furthermore, the paper investigates empirically the implications of the model.

Many of the hypotheses advanced by the simple theoretical model are verified in the estimations and point to an important role for learning and worker mobility as a channel for productivity transfers from high-productivity firms to entrepreneurship.

CHAPTER 4

In Chapter 4, I focus on the development and implementation of a useful technique to solve and estimate more rich and realistic dynamic structural models of discrete choices. In turn, these models can be used to improve our understanding of dynamic aspects in entrepreneurial decisions and the role of risk and uncertainty in the behavior of entrepreneurs. As mentioned above, this chapter is not specific to the analysis of entrepreneurial choice, but it was initially motivated by the requirement to take a more structural empirical approach than in Chapter 1 and eventually to perform a full structural estimation of the model in Chapter 2. However, tractable methods dealing with sufficieny flexible models were not readily available.

The methods developed in Chapter 4 allow for lots of flexibility in the modeling choice that obviates many of the limitations of approachable models. Usually, unobservables are assumed additively separable in utility, conditionally independent, and extreme value distributed. In contrast, the suggested approach can potentially allow for random taste variation, unrestricted substitution patterns, and correlation in unobserved state variables over time. The strategy is to combine polynomial approximation methods and ideas from the literature on discrete choice models with simulation.

When solving and estimating the model, I use Chebyshev polynomials to approximate expected value functions over (observed and unobserved) continuous state variables and simulation techniques to evaluate integrals. This strategy has several important spin-offs. First, it helps ameliorating the 'curse of dimensionality'; the well known exponential increase in computer time and storage requirements. Second, the approach permits very fast and accurate simulation of likelihood functions, once expected value functions are approximated. Third, we can easily allow for lots of heterogeneity, greatly expanding the range of models that can be considered. Fourth, data on continuous state variables do not have to be discretized, mitigating serious problems with approximation errors.

When evaluating the approach, I find that a fifth-order polynomial provides sufficiently accurate approximations. The approximation errors, transmitted to conditional choice probabilities, likelihood functions and structural parameter estimates, are practically eliminated at this level of approximation. I also compare the approach to conventional discretization methods and find that discretizing data is clearly inefficient and results in

substantial approximation error and variation loss in state variables. This is transmitted to significant bias in parameter estimates. When comparing it to the suggested approximation approach, I find that at least 100 grid points are needed to match the precision we obtain when solving the model with a fifth order approximating Chebyshev polynomial, i.e. with 6 Chebyshev nodes. Hence, Chebyshev polynomials are a powerful tool in ameliorating the curse of dimensionality and mitigating the problem with approximation error associated with discretization of the data.

I also provide a Monte Carlo experiment that highlights the importance of heterogeneity bias in dynamic programming discrete choice models and illustrate a model with unobserved heterogeneity in utility.

Although different in topic and methodology, each of the four chapters in the thesis, was meant to contribute to the overall topic: Dynamic Aspects of Entrepreneurial Behavior.

### References

ETHIER, W. J., AND J. R. MARKUSEN (1996): "Multinational Firms, Technology Diffusion and Trade," *Journal of International Economics*, 41(1-2), 1–28.

MELITZ, M. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71(6), 1695–1725.

# EARNINGS, UNCERTAINTY, AND THE SELF-EMPLOYMENT CHOICE

DANIEL LE MAIRE AND BERTEL SCHJERNING

ABSTRACT. This paper investigates the relationship between self-employment choice, expected earnings, and uncertainty. Several interesting results emerge from our analysis on Danish longitudinal register data: Firstly, self-employed (taxable) personal income bunch at kink points in the tax system since self-employed can retain earnings and thereby transfer income across tax-years. Secondly, expected income level and income variance are important determinants in choice of occupation. Thirdly, men put more emphasis on expected earnings level, while women appears more risk averse, which contribute to explain why fewer women are self-employed. Finally, our results suggest that non-western immigrants are marginalized into self-employment.

Keywords: Occupational choice, self-employment, wage-differentials, income uncertainty, risk aversion, overconfidence, self-selection, gender differences.

JEL codes: J16, J24, J31, C33, C35.

1. INTRODUCTION

Compared to wage work, self-employment is a fundamentally different occupation with respect to the type and source of income. While wage workers receive a wage which is subject to a relatively small level of uncertainty, self-employed individuals often face considerably more variation in their income. Moreover, since self-employed typically use own wealth to finance their business, they bear the risk associated with starting up the firm. Therefore, the expected income and the uncertainty of this income are likely to be important determinants of an individual's occupational choice.

The main objective of this paper is to investigate the relationship between the occupational choice and the distributions of associated monetary gains in different occupations. Specifically, we analyze how existing earnings differentials and differences in income uncertainty can explain observed occupational choices. A particular focus is on explaining why fewer women choose to become self-employed.

If individuals are risk averse, we would expect that the self-employed should be compensated for facing higher income uncertainty. However, earnings-differentials may arise for other reasons than risk compensation: Hamilton (2000) argues that cross-sectional earnings differentials may arise due to i) different earnings-experience profiles, ii) self-selection, and iii) non-pecuniary benefits. Hamilton finds that mean and median incomes are lower in self-employment than in wage-employment in the US, although those in the higher income brackets earn more in self-employment than in wage-employment. Hamilton concludes that individuals choose self-employment primarily because of non-pecuniary benefits.

An alternative (or complementary) explanation is that those who choose to become self-employed may be less risk-averse than the typical wage employed. They may even be risk-lovers. In a recent paper by Elston, Harrison, and Rutström (2006), experiments are used to characterize the attitudes to risk among entrepreneurs. Their main finding is that full-time entrepreneurs are less risk-averse and exhibit a significant joy of winning compared to non-entrepreneurs and part-time entrepreneurs.

Yet another explanation relates to the individual's subjective assessment of the probability of success. While Coelho and de Meza (2006) provide experimental evidence that

entrepreneurs tend to overestimate their chance of success, Elston, Harrison, and Rutström (2006) do not find systematic judgmental error of profitability. However, it is found that part-time entrepreneurs are reluctant to enter markets where profitability is based on their perception of their relative skill ability.

Evidence from existing Danish questionnaire surveys shows that men focus more on the expected income level than women when choosing occupation, whereas women emphasize non-pecuniary benefits (Statistics Denmark, 1999; and Kjeldsen and Nielsen, 2000). Thus, 90 per cent of the women who had a child in the age of 0-2 years at the time of the business start-up state that an important reason for becoming self-employed was to make it easier to combine family life and work.

With respect to risk aversion, Byrnes, Miller, and Schafer (1999) analyze 150 psychological studies of risk-taking behavior, and find that in 14 out of 16 tasks, women are more risk-averse. However, according to Croson and Gneezy (2004) the evidence of women being more risk-averse is less clear in the economics literature which has typically focused on financial risk.

Several studies have suggested that overconfidence is part of human nature, e.g. Svenson (1980) reports that 90 per cent of Swedish drivers rate themselves above average. Recently and in relation to occupational choice, Niederle and Vesterlund (2006) find from the conduction of experiments that more women than men prefer to work under a non-competitive piece-rate compensation system rather than under a competitive tournament compensation scheme even though women are found to be as productive as men. Niederle and Vesterlund (2006) conclude that the reason for this difference is that men are too overconfident and enjoy competition more. In other words, too many low productivity men enter the competitive tournament, while productive women do not enter enough.

To evaluate whether the self-employed actually are compensated for their risk-taking, individual level information about the expected income (and the expected distribution of income) in both self-employment and wage-employment is required. To obtain this information, we estimate earnings functions for self-employed and wage-employed separately. However, individuals would be expected to select themselves into the type of occupation where they are most productive. Therefore, we estimate earnings functions for each occupational choice, using the dynamic panel data sample selection model of

Vella and Verbeek (1998, 1999). This also allows us to disentangle the role of unobserved heterogeneity and state dependence in the occupational choices. We find evidence of state dependence in the occupational choices.

The estimated earnings functions are then in turn used to predict an individual's income (and the uncertainty of this income) in different occupations. The random components of the model are partitioned into transitory and permanent shocks, which in turn are used to create occupational and education specific measures of income variance (uncertainty) and skewness (the risk/chance of very low/high incomes).

Rather that rather than characterizing the entrepreneur, we directly evaluate the impact of earnings on the choice of becoming self-employed, wage-employed or unemployed by examining the roles of expected earnings, risk aversion and over-confidence. This is done for each gender separately. Our results complement existing evidence from experimental economics, providing an potential explanations for the substantial gender gap in the probability of choosing to become self-employed.

We use a large longitudinal data set based on Danish register data from 1980 to 1996, providing us with detailed individual information about income, wealth, education, labor market status (occupation), region of residence, and immigration status. Since the panel covers more than 15 years, we can track long sequences of individual occupational choices and, thereby, appropriately investigate the dynamics of the self-employment choice.

Our results point to a large role for monetary aspects when choosing occupation. As expected, people prefer the sector with the highest expected income and lowest expected variance and, thus, on average appear risk-averse. We find that men put more emphasis on the earnings level, while women appear more risk-averse, which could be one of the crucial reasons why fewer women are self-employed. We do not find evidence of overconfidence. If anything, women instead seem to under-estimate their chance of success compared to men.

The explanatory power of the occupational choice model is quite impressive considering that we only include the predicted income level, variance and skewness. However, we explain much less of the variation in the realized occupational choices for the group of non-western immigrants. Immigrants are interesting with respect to occupational choice since they are more likely to start up their own business than natives. We find that immigrants

put much less emphasis on the earnings level. These findings provide additional evidence for immigrants being marginalized into self-employment as Blume, Ejrnæs, Nielsen, and Würtz (2005) suggest. From their analysis on Danish transition data it is found that most non-western immigrants entering self-employment come from unemployment and that they do not use self-employment as a stepping stone for becoming wage-employed.

The rest of the paper is organized as follows: In section 2 we describe the data used in the analysis. In section 3 we formulate the econometric specification. In section 4, we present the results. Section 5 concludes.

## 2. Data

The data we use in this paper is an unbalanced panel data set for 1980-1996. The data is a representative 10 per cent sample extract drawn from the Integrated Database for labor market Research (IDA) and the Danish Income Registry (IKR) both maintained by Statistics Denmark. IDA and IKR are both longitudinal data based on register data for all individuals in Denmark. Since data originates from administrative records covering the entire Danish population there is only natural attrition in the data, i.e. birth, death and migration of individuals. The occupational status is observed once a year (the last week of November). We divide the labor market status into three states; self-employed, wage-employed, and unemployed. Since the panel covers more than 15 years, we have the possibility to track individuals over long time periods (before, during and after self-employment) and, thereby, appropriately control for the dynamics of the occupational choice. These high-quality Danish data contains very detailed individual information concerning, e.g., income, wealth, education, labor market status, region of residence, and immigration status. Moreover, the data also includes the same information for cohabitants allowing us to aggregate variables to the household level.

In order to avoid distortions in the results due to retirement patterns and educational attainment we restrict the sample to include persons aged 30-55 years only. This leaves us with $2,424,694$ observations in total of which $1,130,635$ are women.

**FIGURE 1: DISTRIBUTION OF DISPOSABLE INCOME IN 1996**
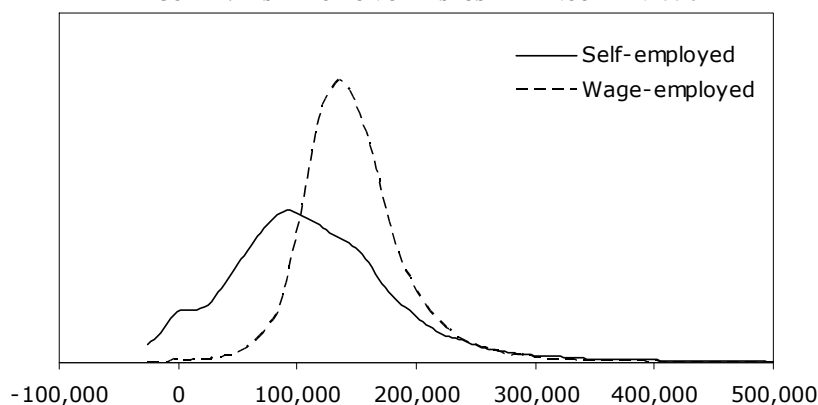


**TABLE 1: INCOME DISTRIBUTIONS IN 1996, SELECTED PERCENTILES**

| Percentile | Personal Income | | Disposable Income | |
|---|---|---|---|---|
| | Self-empl. | Wage-empl. | Self-empl. | Wage-empl. |
| 5 | 0 | 111,055 | -12,529 | 85,848 |
| 10 | 7,745 | 135,380 | 12,687 | 99,698 |
| 25 | 75,980 | 167,222 | 61,119 | 119,634 |
| 40 | 120,678 | 190,022 | 88,076 | 133,592 |
| 50 | 148,590 | 205,274 | 104,021 | 142,430 |
| 60 | 183,967 | 221,686 | 121,745 | 151,914 |
| 75 | 242,834 | 253,018 | 152,229 | 169,066 |
| 90 | 345,187 | 325,042 | 202,619 | 202,231 |
| 95 | 458,649 | 387,028 | 248,060 | 229,819 |

For the analysis of occupational choice we need to decide on an income measure to use. One obvious candidate is disposable income since this measure is closely related to current consumption possibilities and, hence, utility.[1]

Figure 1 shows kernel densities for the disposable income for self-employed and for wage-employed in 1996. Both distributions are right-skewed with the distribution of incomes from self-employed being most right skewed. From both Figure 1 and Table 1 it can be seen that the mean disposable income for self-employed is considerably below the mean income for wage-employed. However, due to the skewness the 90th percentile earns more in self-employment than the equivalent in wage-employment.

---

[1]We compute the gross income including wage-income, capital income, labor market contributions (since 1994), taxable and non-taxable benefits. In order to obtain the disposable income we subtract the tax payments.

**FIGURE 2: DISTRIBUTION OF PERSONAL INCOME IN 1996**
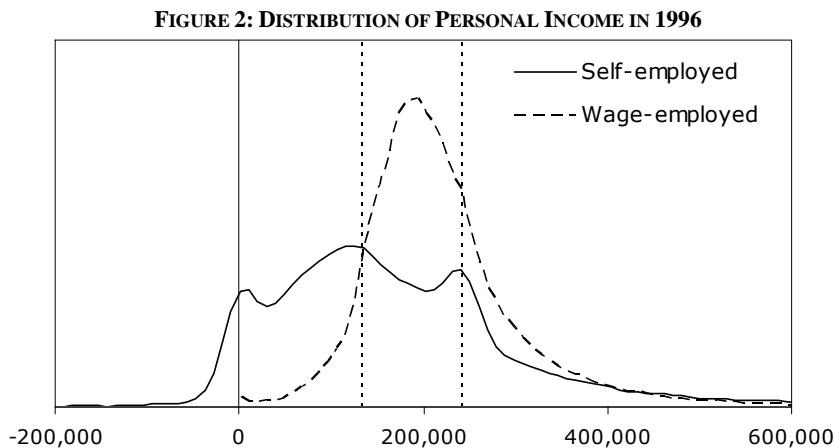


Figure 1 and Table 1 confirm the US evidence presented in Hamilton (2000), who also finds that mean and median incomes are lower in self-employment, but that those in the higher percentiles earn more in self-employment relative to wage-employment.

In Figure 2 we have depicted the (taxable) personal income for respectively wage-employed and self-employed together with two dotted vertical lines indicating where the medium and upper tax brackets set in. In contrast to wage-employed, self-employed tend to bunch just below where the tax brackets set in. This can be due to self-employed being in charge of their own working time, but it may also reflect that self-employed are building up inventories and capital stocks or have other means of extracting income from their firm (possibly also in the grey area between firm economics and personal economics). Finally, an institutional feature ("Virksomhedsordningen") allows self-employed to retain earnings in the firm.

The bunching at the tax brackets suggests that adding retained earnings (less of taxes) to the disposable income constitutes a better income measure for self-employed and we only use this income measure in the rest of the paper. As shown in Figure 3 we find that the unconditional mean and median incomes are larger in self-employment than in wage-employment in contrast to the US evidence in Hamilton (2000) and in contrast to when applying the narrow income measure.

## 3. ECONOMETRIC SPECIFICATION

The organization of this section, can be summarized as follows: First, we consider the estimation of conditional earnings functions using Vella and Verbeek (1998, 1999) sample

FIGURE 3: DISTRIBUTION OF DISPOSABLE INCOME PLUS RETAINED EARNINGS
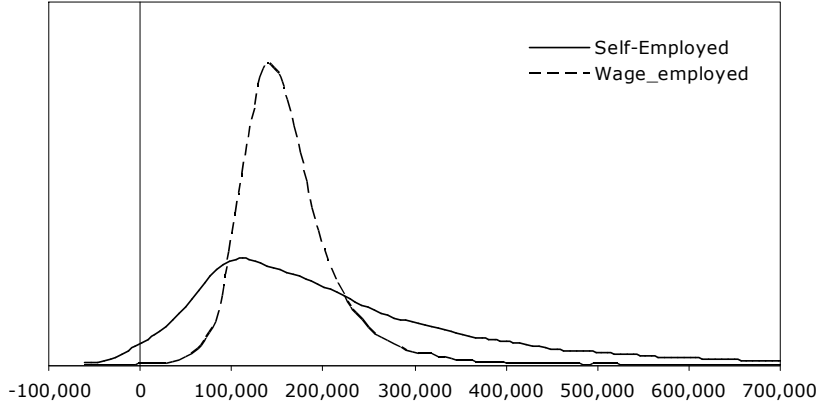


selection model for panel data. Hereafter, we construct income uncertainty and skewness measures. Finally, using measures for expected income, uncertainty and skewness, we model the occupational choice in a conditional logit model.

3.1. **Earnings Conditional on Occupational Choice.** For each person we separately predict the disposable income including retained earnings from being self-employed and being wage employed. The chosen income measure is disposable income including retained earnings. We use unemployment benefits for the group of unemployed. For each occupation we model earnings as a simple log-linear mincer earnings equation

$$(3.1) \qquad\qquad \ln y_{nt}^* = x_{nt}\beta + \alpha_n + \varepsilon_{nt}$$

where $n$ indexes individuals $(n = 1, .., N)$ and $t$ indexes time $(t = 1, ...., T)$; $y_{nt}^*$ is annual disposable income plus retained earnings, $\beta$ is a vector of unknown coefficients to be estimated, $x_{nt}$ is a vector covariates, $\alpha_n$ represents unobserved heterogeneity and $\varepsilon_{nt}$ is a normally distributed disturbance.

Since we observe earnings for the chosen occupational status only, the conditional earnings functions will in general be estimated on a non-random selected sample. There are several arguments, why self-selection may be an issue in the present context. In the Roy (1951) model the individual ex-ante knows her sector-specific productivity, and will select herself into the sector, where she is most productive. Furthermore, if the incomes in the two sectors are highly correlated, the most productive persons will select the sector with

the largest dispersion of sector specific abilities, while the least productive will select the sector with the smallest dispersion.

The Danish labor market is characterized by a compressed wage structure as a consequence of the generous unemployment benefit level and a high degree of organization on both employer and worker sides. As argued by Malchow-Møller, Markusen, and Skaksen (2006) such institutional arrangement may well imply that the most productive are not paid according to their marginal product and, therefore, the most able may select themselves into self-employment. On the other hand, the least productive may not have a sufficiently high productivity to earn the minimum wage in paid employment. Consequently, marginalization may also push the least productive into self-employment. Blume, Ejrnæs, Nielsen, and Würtz (2005) argue that this indeed is the case for non-western immigrants in the Danish labor market.

Yet another type of selection, ex-post self-selection, arises in leaning models such as Jovanovic (1979) and Jovanovic (1984), where persons have no ex-ante knowledge of their productivity, but consecutively observe output realizations. Persons experiencing poor output realizations will quit and search for a new match.

To control for the selection problem we use the Vella and Verbeek (1998, 1999) dynamic panel data application of Heckman's two-step sample selection model. The selection is modelled as a dynamic random effects probit, which allows us to separate two sources of persistence in the occupational choice: Persistence as a result of unobserved heterogeneity and (true) state dependence. Since we do not observe the first occupational choice, we cannot assume that the initial observation of the occupational is truly exogenous. We use the Wooldridge (2005) way of handling the initial conditions problem and, thus, allow the unobserved heterogeneity to be correlated with the initial dependent variable.

We will now briefly explain the model.[2] We consider a model consisting of two equations, where the parameters of equation (3.1) are of primary interest, while the selection equation below is a reduced form equation for the occupational choice. The selection part of the

---

[2]For a detailed treatment of the model see Vella and Verbeek (1999).

model can be summarized as

$$(3.2) \qquad d_{nt}^* = x_{nt}\gamma + \phi d_{nt-1} + \mu_n + \eta_{nt}$$

$$(3.3) \qquad d_{nt} = \mathbf{1}\left(d_{nt}^* > 0\right)$$

$$(3.4) \qquad \ln y_{nt} = \ln y_{nt}^* \text{ if } d_{nt} = \mathbf{1}$$

$$= 0 \text{ (unobserved) otherwise}$$

where $y_{nt}^*$ and $d_{nt}^*$ are latent endogenous variables with observed counterparts $y_{nt}$ and $d_{nt}$.

The equation of interest is assumed to have the usual error component structure, where $\alpha_n \sim iN\left(0, \sigma_\alpha^2\right)$ and $\varepsilon_{nt} \sim iN\left(0, \sigma_\varepsilon^2\right)$. For the selection equation we allow for unobserved heterogeneity through random individual effects, such that the selection equation has the following two-component error structure $\mu_n \sim iN\left(0, \sigma_\mu^2\right)$ and $\eta_{nt} \sim iN\left(0, \sigma_\eta^2\right)$. We allow for correlation between the individual effects as well as correlation between the idiosyncratic disturbances, that is $\text{cov}(\alpha_n, \mu_n) = \sigma_{\alpha\mu} \neq 0$ and $\text{cov}(\varepsilon_{nt}, \eta_{nt}) = \sigma_{\varepsilon\eta} \neq 0$. Finally, denote $\xi_{nt} = \alpha_n + \varepsilon_{nt}$ , $v_{nt} = \mu_n + \eta_{nt}$ , $\mathbf{x}_n = [x_{n1}, ..., x_{nT}]'$ and let $\boldsymbol{v}_n$ be a $T$ vector of $v_{nt}$.

Assume now

$$(3.5) \qquad \boldsymbol{v}_n | x_n \sim iN\left(0, \sigma_\mu^2 \mathbf{ii}' + \sigma_\eta^2 \mathbf{I}\right)$$

$$(3.6) \qquad E\left[\xi_{nt} | x_n, v_n\right] = \tau_1 v_{nt} + \tau_2 \bar{v}_n$$

where $\bar{v}_n = T^{-1} \sum_{t=1}^{T} v_{nt}$ and where $\tau_1 = \sigma_{\varepsilon\eta}/\sigma_\varepsilon^2$ and $\tau_2 = T\left(\sigma_{\alpha\mu} - \sigma_{\varepsilon\eta}\sigma_\mu^2/\sigma_\varepsilon^2\right) / \left(\sigma_\eta^2 + T\sigma_\mu^2\right)$ are constants to be estimated and $\mathbf{i}$ is a column of ones. Note that equation (3.6) imposes strict exogeneity of $x_{nt}$, such that errors are assumed to be independent of future and lagged values of $x_{nt}$. To estimate the conditional mean for the dependent variable in the equation of interest, we condition on the chosen occupation

$$E\left[\ln y_{nt} | x_n, d_{n0}, d_n\right] = x_{nt}\beta + E\left[\xi_{nt} | x_n, d_{n0}, d_n\right]$$

where $E\left[\xi_{nt} | x_n, d_{n0}, d_n\right]$ is the selection bias induced by correlation between the errors in the two equations.

Under these assumptions, it can be shown that the conditional mean of the error-term from the selection equation, $E\left[v_{nt}|x_n, d_{n0}, d_n\right]$ can be estimated by the following expression

$$(3.7) \quad \tilde{v}_{nt} = \frac{1}{\int f\left(d_n|x_n, \mu_n\right) f\left(\mu_n\right) d\mu_n} \int \left(\mu_n + E\left[\eta_{nt}|x_n, \mu_n\right]\right) f\left(d_n|x_n, \mu_n\right) f\left(\mu_n\right) d\mu_n$$

This expression can be approximated by quadrature methods or simulation. Once we have estimated the reduced form parameters for the selection equation, we can easily simulate the conditional error $\tilde{v}_{nt}$.[3]

After computing $\tilde{v}_{nt}$ and the individual specific means $\bar{v}_n = \frac{1}{T_n}\sum_t^{T_n} \tilde{v}_{nt}$ we can estimate the following equation by the simple linear random effects model

$$\ln y_{nt} = x_{nt}\beta + \tilde{v}_{nt}\theta_1 + \bar{v}_n\theta_2 + \alpha_n + \varepsilon_{nt}$$

### 3.2. Uncertainty and Skewness Measures.

For each category in our disaggregated education breakdown shown in Table A.1 we estimate the occupational-specific measures of variance and skewness of the income processes. This is done separately for men and women.

We divide the uncertainty into a permanent part relating to the variance of the individual time-constant $\alpha_n$ and into a transitory uncertainty relating to the time-varying error-term.[4] Among the covariates in $x_{nt}$ we have included 28 educational dummies.[5] We define $a_n = \exp\left(\alpha_n\right)$ and $e_{nt} = \exp\left(\varepsilon_{nt}\right)$ and compute the variance $R$ and the skewness $K$ for each education type $l$ by

$$R_l^\alpha = \frac{1}{N_l}\sum_{n=1}^{N_l}\left(a_{nl} - \bar{a}_l\right)^2 \qquad K_l^\alpha = \frac{1}{N_l}\sum_{n=1}^{N_l}\left(a_{nl} - \bar{a}_l\right)^3$$

$$R_l^\varepsilon = \frac{1}{T}\frac{1}{N_l}\sum_{n=1}^{N_j}\left(e_{nlt} - \bar{e}_l\right)^2 \qquad K_l^\varepsilon = \frac{1}{T}\frac{1}{N_l}\sum_{n=1}^{N_l}\left(e_{nlt} - \bar{e}_l\right)^3$$

By averaging the residuals only on education groups, we effectively assume that the income uncertainty does not depend on for example experience, which is obviously an

---

[3]The procedure is summarized in algorithm 1 in the appendix.

[4]Recently, Diaz-Serrano, Hartog, and Nielsen (2003) have used a similar approach in the context of educational choice.

[5]In the IDA database there are 1,750 different educations, but in order to secure representativity we operate with 28 education groups only (see Table A.1 in the Appendix). We have aimed at securing representativity by not making a too disaggregated educational break-down, but on the other hand aimed at selecting as homogeneous groups as possible.

approximation. Averaging the incomes on other variables as well is not feasible with the detailed education break-down used.

For an unemployed there is no or very little uncertainty regarding income. Consequently, we set the variance and skewness equal to zero.

3.3. **A Model of Occupational Choice.** The behavioral framework underlying the occupational choice model is simple: We assume that individuals each period associate each occupation with a continuous random utility function, $U_{nit}$ , where each occupation is indexed by $i \in [se, we, ue]$. Each period individuals choose between self-employment (se), wage-employment (we) and unemployment (ue) to maximize the $U_{nit}$.[6] Random utility is assumed to be a linear function of occupational specific earnings, and the variance and skewness of permanent and transitory income shocks. Hence, the random utility function can be written as

$$U_{nit} = \mathbf{x}_{nit}\boldsymbol{\beta} + \delta_i + \epsilon_{nit} \;\; \text{with} \;\; n = 1, ..., N \quad \text{and } t = 1, .., T$$

where $\delta_i$ is a choice-specific constant, $\mathbf{x}_{nit} = \left[ \hat{Y}_{nit}, R_l^\alpha, K_l^\alpha, R_l^\varepsilon, K_l^\varepsilon \right]$ denotes the set of attributes associated with each occupation, $\boldsymbol{\beta}$ is a vector of coefficients related to the the choice specific attributes $\mathbf{x}_{int}$. The error component $\epsilon_{nit}$ is assumed to be individual-, choice,- and time specific and distributed according to a Type I extreme value distribution. With this distributional assumption, we end up with McFadden's well known Conditional Logit model for discrete choices.

## 4. RESULTS

4.1. **Self-selection and Earnings Differentials.** In this section, we investigate the extent to which earnings differentials can be explained by individuals self-selecting themselves into the different occupations. To account for the potentially important selection problems, we estimate the model sample selection model of Vella and Verbeek (1998,1999). First, we estimate the selection equation given by equation (3.2) and equation (3.3) by a

---

[6]Even though each individual maximizes utility each period by choosing occupations this need not be equivalent to maximization of life-time utility given by a discounted sum of period utility. However, this simplification is needed to make the model operational.

dynamic random effects probit. Hereafter, we estimate the parameters in the conditional earnings function in equation (3.1).

Since the choice of labor market state differs considerably between the genders, the sample correction and the prediction of incomes are done separately for men and women. Additionally, the existence of wage differentials between the genders suggests that it would be appropriate to run the wage equations separately.

The results from the selection equations given in table A.2 suggest that the impact of the lagged dependent variable is positive and highly significant, indicating the presence of substantial state dependence. State dependence can be a result of cost of and uncertainty of labor market transitions and is likely to be amplified for transitions into self-employment in the presence of start-up costs. In an intertemporal model of occupational choice Schjerning (2005) shows that the combination of irreversible start-up costs and income uncertainty introduces an option value of being self-employed. To avoid potential start-up costs associated with later re-entry, the self-employed is willing to wait until good times occur rather than temporarily leaving self-employment. This introduces a value of waiting and consequently we will see later entry and later exit.

The magnitude of state dependence for the self-employed is substansial: Being self-employed in the previous period increases the probability of being self-employed in the current period from 1.2 to 41.5 per cent for females and from 4.0 per cent to 48.7 per cent for men. As a comparison, the marginal effect from previous wage-employment is 19.9 per cent for females and 24.0 per cent for men.

The results from the selection equation suggest that, in general, the probability of being self-employed varies much between the educational categories both with respect to length and type of education. Although the picture is quite mixed, it seems to be the case that unskilled and some groups of highest education are the most likely to become self-employed. The latter is due to the fact that the self-employed include professionals such as practitioner doctors, dentists, lawyers and accountants.

The estimated earnings equations are given in table A.3. As dependent variable we use the disposable income including retained earnings. We allow for unobserved heterogeneity in the form of random effects, and we control for the usual socio-demographic variables. We find positive coefficients on marriage for men, while they are negative for women. The

origin variables have the expected signs and magnitudes, i.e., non-western immigrants earn considerably less than western immigrants, second generation immigrants and natives. It is striking that non-western immigrants are more likely to become self-employed even though they should expect a much lower income in self-employment compared to wage-employment.

We find the usual hump-shaped effect of age, which obviously captures labor market experience. We have included dummies for each education from our detailed educational break-down shown in the appendix. The general picture is as expected that the longer education, the higher disposable income. As one would presume, the returns to education differ remarkably between the educations. For example, the returns to humanities are lower compared to social sciences at each length of education reflecting the relatively higher unemployment rate that may lead to accepting jobs below the educational level.

If education is a signal, so that employers use education to screen potential workers, we would expect lower returns to education in self-employment. There does not seem to be much evidence for the signalling hypothesis.

Since we do not wish to rely on the non-linearity of the selection equation to identify the selection effects in the income equations we need to exclude at least one variable. We use the lagged dependent variable, household wealth, dummies for children in the household and a dummy for the spouse being self-employed.

The inclusion of the correction terms account for the selection bias induced by the correlation between unobservables in the selection model and earnings equations. The coefficients to the correction terms $\upsilon_{nt}$ and $\bar{\upsilon}_n$ are statistically significant in all four regressions. In the case of men, the coefficient on both correction terms are negative, implying that the marginalization on average dominates. Taken literally, we have that those in wage-employment will tend to earn more in self-employment than those already self-employed.

In contrast to this, the coefficient on the individual specific correction term, $\bar{\upsilon}_n$ is positive in the self-employment earnings equation for women implying that those in wage-employment have a lower self-employment potential than the currently self-employed. Since the income is measured on a yearly basis a possible explanation for the positive

selection into self-employment relates to differences in the hours of work between wage-employed and self-employed. About 20 per cent of female wage-employed work part-time and if this fraction is larger than the corresponding for women in self-selection a positive selection into self-employment will, on average, emerge. In recent work by Carrasco and Ejrnæs (2003) it is in fact argued that the relative low share of female self-employed in Denmark can be explained by the relative high level of flexibility in the Danish labor market providing the possibility to work part time in paid employment. Similar arguments apply to women planning to have children, as the opportunities for paid maternity leave are better in wage-employment. Another explanation might be glass-ceiling effects in wage-employment, see e.g. Albrecht, Björklund, and Vroman (2003) for Swedish evidence.

4.2. **The Occupational Choice.** The occupational choice model is estimated for a several different subsamples. The results from these estimations are shown in Table 2. Each row in the Table corresponds to the results for a different subsample. The figures in the Table show the effects of the mean, variance (uncertainty) and skewness (i.e., in this case the chance of very high incomes) of predicted earnings conditional on the occupational choice. Note that in the estimations, variance and skewness of both transitory and permanent shocks are included in the model. For expositional purposes, however, we only report variance and skewness of the permanent income component in the Table.[7]

The coefficients to mean earnings gives the marginal utility of expected income, while the (negative of the) coefficient to the variance can be interpreted as the marginal (dis-) utility of income uncertainty. To give an example, a positive coefficient to expected income is associated with individuals consistently choosing occupations with higher levels of expected earnings, while a negative coefficient to variance emerges when individuals choose occupations with little income uncertainty.

---

[7]Since the earnings equations were estimated with age variables and time-dummies there is no aggregate time variation left in the error-terms, but still individual specific variation occurs. Alvarez, Browning, and Ejrnæs (2002) find that Danish income processes are particularly heterogenous. Steep income-experience profiles imply a large variance, but when controlling for the income level, we should due to income smoothing expect that a flat income-tenure profile is preferred. However, a steeper wage-experience profile may indicate greater possibilities such as promotion for wage-employed and business expansion for self-employed, which may explain the positice coefficient to the variance of the temporary income shocks. The skewness of the time-varying part does not seem to play any role.

**TABLE 2: CHOICE OF LABOR MARKET STATUS**
**(CONDITIONAL LOGIT MODEL)**

| Subsample | Mean Earnings | | Distribution of Permanent Chock's | | | | MRS Variance/ Mean | Skewnes/ Mean | Sample size | R$^2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Variance | | Skewness | | | | | |
| All | 1.58 | (0.007)** | -0.15 | (0.004)** | -0.02 | (0.001)** | -0.10 | -0.010 | 7,274,082 | 51 |
| Non-western immigrants | 1.08 | (0.044)** | 0.05 | (0.025) | -0.02 | (0.005)** | 0.04 | -0.016 | 119,412 | 23 |
| | | | | | | | | | | |
| Women | 0.40 | (0.017)** | -0.22 | (0.011)** | -0.04 | (0.001)** | -0.56 | -0.091 | 3,391,905 | 59 |
| Non-western immigrants | 0.81 | (0.095)** | -0.34 | (0.070)** | 0.02 | (0.009)* | -0.42 | 0.022 | 45,564 | 29 |
| Married | 1.00 | (0.023)** | -0.27 | (0.014)** | -0.03 | (0.002)** | -0.27 | -0.032 | 2,514,129 | 61 |
| HH. Wealth(t-1)>500.000 | 0.85 | (0.048)** | -0.14 | (0.022)** | -0.01 | (0.003)** | -0.16 | -0.013 | 493,101 | 64 |
| Father self-employed | 1.29 | (0.148)** | 0.20 | (0.030)** | -0.06 | (0.009)** | 0.16 | -0.043 | 64,728 | 64 |
| age<40 | 0.40 | (0.026)** | -0.11 | (0.013)** | -0.05 | (0.002)** | -0.27 | -0.118 | 1,353,195 | 60 |
| age>45 | 0.58 | (0.030)** | -0.37 | (0.025)** | -0.02 | (0.002)** | -0.63 | -0.038 | 1,148,580 | 57 |
| | | | | | | | | | | |
| Men | 2.81 | (0.016)** | -0.10 | (0.005)** | 0.00 | (0.001) | -0.04 | -0.001 | 3,882,177 | 44 |
| Non-western immigrants | 1.81 | (0.088)** | 0.14 | (0.027)** | -0.03 | (0.006)** | 0.08 | -0.018 | 73,848 | 20 |
| Married | 2.39 | (0.020)** | -0.15 | (0.006)** | 0.01 | (0.001)** | -0.06 | 0.003 | 2,761,398 | 48 |
| HH. Wealth(t-1)>500.000 | 1.42 | (0.044)** | -0.18 | (0.012)** | 0.06 | (0.002)** | -0.13 | 0.045 | 474,390 | 43 |
| Father self-employed | 3.77 | (0.110)** | -0.10 | (0.023)** | -0.02 | (0.007)** | -0.03 | -0.006 | 91,098 | 40 |
| age<40 | 3.33 | (0.027)** | -0.07 | (0.007)** | -0.02 | (0.002)** | -0.02 | -0.006 | 1,538,208 | 50 |
| age>45 | 2.53 | (0.027)** | -0.10 | (0.009)** | 0.01 | (0.002)** | -0.04 | 0.004 | 1,348,191 | 39 |

Notes: Standard errors are in parentheses. * significant at 5%; ** significant at 1%. Other controls: Occupational specific constants and measures of the temporary components of estimated chocks (skewness and variance)

To make results comparable across different subsamples, we compute the marginal rate of substitution (MRS) between the variance and mean earnings and between the skewness and mean earnings. The MRS can be interpreted as the rate at which you are willing to trade off more uncertainty for higher income.[8] These results are shown in the right part of the Table.

Considering the full sample (the first row in the Table), the results point to a large role for monetary aspects when choosing occupation. As expected, people's choice of occupation is positively affected by expected (mean) earnings and negatively by a higher variance of the income. Thus, on average, people appear risk averse. These findings are found to be robust to various sample decompositions.

Turning to the differences between the genders, we find that men put more emphasis on the earnings level, while women appear more risk averse. This is reflected in the much lower value of the MRS estimate for women. This could be one of the main reasons why fewer women choose to become self-employed.

---

[8]This normalization is important since estimates from two different subsamples are not directly comparable due to differences in the variances of the unobserved factors.

Women seem to be behaving in a less risk averse manner when household wealth exceeds DKK 500,000. This is perfectly consistent with standard models of intertemporal behavior that find that the degree of effective risk aversion is decreasing in wealth; see, e.g., Deaton (1991), Carroll (1997), and Schjerning (2005).

The finding that married women appear less risk averse than other women is also fully consistent with models from the literature on family economics that point to risk sharing as being a potentially important economic gain from marriage, see e.g. Hess (2004).

A similar variation in men's attitudes towards risk is not found. An interesting finding, however, is that the MRS between income uncertainty and expected earnings is virtually zero compared to females. This confirms the evidence from Danish questionnaires, referred to above, which pointed to men putting much more emphasis on monetary gains (expected income) than women.

Finally, a positive coefficient to skewness is interpreted as being consistent with evidence of overconfidence. If people systematically prefer occupations with a high degree of skewness (a chance of very high incomes) it may be due unrealistic, strong beliefs in their own ability.

For the full sample, a negative coefficient to skewness is found. Hence, on average, there is no evidence of overconfidence. The more detailed results with respect to this behavioral hypothesis are mixed and inconclusive. If anything, men behave somewhat more overconfidently than females. This result match those found from experimental studies.

It is striking that we in the model for immigrants only can explain 17 per cent compared to 50 per cent in the other models. Moreover, the coefficient to income is much lower than in the other conditional logit models. Hence, other important (unobserved) factors, such as lack of opportunities in the ordinary labor market and non-pecuniary benefits may be much more relevant in explaining their occupational choice. Hence, the low explanatory power, and the lower coefficient to income points to self-employment being the last resort due to marginalization in wage-employment. We also find that non-western immigrants appear less risk averse. This may be due to marginalization forcing immigrants to accept insecure and low paid occupations, but it can also be a consequence of cultural differences in the attitudes towards self-employment.

## 5. Conclusions

This paper uses high quality Danish longitudinal register data, to investigate the relationship between self-employment choice, expected earnings and income uncertainty. We proceed in the following steps: Firstly, we estimate of conditional earnings functions using the sample selection model of Vella and Verbeek (1998, 1999). Secondly, using measures for expected income, uncertainty and skewness, we model the occupational choice in a conditional logit model.

Comparing earnings distributions based on different income measures, we find that i) the dispersion of incomes is in general much larger for the self-employed and ii) Danish self-employed earn more than wage-employed when retained earnings are included in the income measure. Contrary to wage-workers, self-employed (taxable) personal income bunch at kink points in the tax system since self-employed (unlike wage workers) has the possibility to retain earnings and thereby transfer income across years. The progressive Danish income tax system provides strong incentives to make such transfers.

Several experimental studies have found that while men are more competitive, women are more risk averse. In the context of occupational choice, we find that men put more emphasis on the income level, while women seem to be more risk-averse. This result is found to be robust to various sample decompositions.

Linking the behavioral results from the experimental literature with income distributions in self-employment and wage-employment may explain why fewer women become self-employed. We find that part of the gender gap can be explained by gender differences in the trade-offs between income level and the variance of incomes. However, we find no effect of skewness of incomes.

Non-western immigrants are overrepresented in self-employment. The occupational choice model performs considerably worse for this group and we find smaller effects of income level and variance. Furthermore, the sample selection model shows that non-western are more likely to become self-employed even though they should expect a much lower income in self-employment than native Danes. This suggests that non-western immigrants are marginalized into self-employment

# REFERENCES

ALBRECHT, J., A. BJÖRKLUND, AND S. VROMAN (2003): "Is There a Glass Celing in Sweden?," *Journal of Labor Economics*, 21(1), 145–177.

ALVAREZ, J., M. BROWNING, AND M. EJRNÆS (2002): "Modelling Income Processes with Lots of Heterogeneity," CAM Working Paper 2002-01.

BHAT, C. (2001): "Quasi-Random Maximum Simulated Likelihood Estimation of the Multinomial Mixed Logit Model," *Transportation Research*, B 35, 677–693.

BLUME, K., M. EJRNÆS, H. S. NIELSEN, AND A. WÜRTZ (2005): "Labor Market Transitions of Immigrants with Emphasis on Marginalization and Self-Employment," working paper.

BYRNES, J. P., D. C. MILLER, AND W. D. SCHAFER (1999): "Gender Differences in Risk-Taking: A Meta-Analysis," *Psychological Bulletin*, 125, 367–383.

CARRASCO, R., AND M. EJRNÆS (2003): "Self-Employment in Denmark and Spain: Institution, Economic Conditions and Gender Differences," CAM Working Paper 2003-06.

CARROLL, C. D. (1997): "Buffer Stock Saving and the Life Cycle/Permanent Income Hypothesis," *Quaterly Journal of Economics*, 112(1), 1–56.

COELHO, M., AND D. DE MEZA (2006): "Self Deception, Self Selection, Self Destruction: An Experimental Investigation of Adverse Selection," working paper.

CROSON, R., AND U. GNEEZY (2004): "Gender Differences in Preferences," working paper.

DEATON, A. (1991): "Savings and Liquidity Constraints," *Econometrica*, 59, 1121–1248.

DIAZ-SERRANO, L., J. HARTOG, AND H. S. NIELSEN (2003): "Compensating Wage Differentials for Schooling Risk in Denmark," IZA Discussion Paper no. 963.

ELSTON, J. A., G. W. HARRISON, AND E. E. RUTSTRÖM (2006): "Characterizing the Entrepreneur Using Field Experiments," working paper.

HAMILTON, B. H. (2000): "Does Entrepreneurhip Pay? An Empirical Analysis of the Returns to Self-Employment," *Journal of Political Economy*, 108, 604–631.

HESS, G. D. (2004): "Marriage and Consumption Insurance: What's Love Got to Do with It?," *Journal of Political Economy*, 112, 290–318.

JOVANOVIC, B. (1979): "Job Matching and theTheory of Turnover," *Journal of Political Economy*, 87, 972–990.

——— (1984): "Matching, Turnover and Unemployment, Journal of Political Economy," *Journal of Political Economy*, 92, 108–122.

MALCHOW-MØLLER, N., J. R. MARKUSEN, AND J. R. SKAKSEN (2006): "Labour Market Institutions, Learning, and Self-Employment," CEBR discussion paper 2006-24.

NIEDERLE, M., AND L. VESTERLUND (2006): "Do Women Shy Away from Competition? Do Men Compete Too Much?," working paper.

ROY, A. D. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3, 135–146.

SCHJERNING, B. (2005): "A Dynamic Model for Occupational Choice, Savings and Formation of Entrepreneurial Human Capital under Liquidity Constraints and Income Uncertainty," working paper.

STATISTICSDENMARK (1999): "Generel Erhvervsstatistk," Statistiske Efterretninger 1999:19.

SVENSON, O. (1980): "Are We All Less Risky and More Skillful Than our Fellow Drivers?," *Acta Psychologica*, 47, 143–148.

TRAIN, K. (2003): *Discrete Choice Models with Simulation.* Cambridge University Press, Cambridge.

VELLA, F., AND M. VERBEEK (1998): "Whose Wages Do Unions Raise? A Dynamic Model of Unionism and Wage Rate Determination for Young Men," *Journal of Applied Econometrics*, 13(2), 163–183.

VELLA, F., AND M. VERBEEK (1999): "Two-Step Estimation of Panel Data Models with Censored Endogenous Variables and Selection Bias," *Journal of Econometrics*, 90, 239–263.

WOOLDRIDGE, J. M. (2005): "Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Effects," *Journal of Applied Econometrics*, 20(1), 39–54.

.

## 6. Appendix

**Algorithm 1.** *Estimation of conditional error-term from the selection equation, $E\left[v_{nt}|x_{nt}, d_{n0}, d_{nt}\right]$*

(1) *For a given set of parameter values $\theta_1 = (\gamma, \phi, \sigma_\mu)$ take a draw from $\mu_n^r$ from $f(\mu_n|\sigma_\mu) = N(0, \sigma_\mu)$ and calculate the likelihood for individual $i$ conditional on the draw*

$$f(d_n, d_{i0}|x_n, \mu_n^r) = \prod_{t=1}^{T_n} f(d_{nt}|x_{nt}, \mu_n^r) f(d_{n0}|x_{nt}, \mu_n^r)$$

*where $f(d_{nt}|x_{nt}, \mu_n^r) = \Phi_{nt} d_{nt} + (1 - \Phi_{nt})(1 - d_{nt})$ and where $\Phi_{nt} \equiv \Phi(x_{nt}\gamma + \phi d_{nt-1} + \mu_n^r)$*

(2) *Repeat many times and average the results to obtain the Simulated Log Likelihood function (SLL)*

$$SLL = \ln \frac{1}{R} \sum_r^R f(d_n, d_{i0}|x_n, \mu_n^r)$$

(3) *Choose $\theta_1^{MSL}$ so that SLL is maximized*

(4) *Given the MSL estimates from the fist stage regression $\theta_1^{MSL}$, we can easily simulate $\hat{v}_{nt}$. Take $R$ draws from $f(\mu_n|\sigma_\mu^{MSL})$ and calculate the simulated counterpart of $\hat{v}_{nt}$*

$$\tilde{v}_{nt} = \frac{1}{\frac{1}{R}\sum_r^R f(d_n, d_{i0}|x_n, \mu_n^r)} \frac{1}{R} \sum_r^R (\mu_n^r + E[\eta_{nt}|x_n, \mu_n^r]) f(d_n, d_{i0}|x_n, \mu_n^r)$$

*where $E[\eta_{nt}|x_n, \mu_n^r] = \dfrac{d_{nt}\phi_{nt}}{\Phi_{nt}} - \dfrac{(1 - d_{nt})\phi_{nt}}{1 - \Phi_{nt}}$ is the cross-sectional generalized residual for the probit model and where $\phi_{nt} \equiv \phi(x_{nt}\gamma + \phi d_{nt-1} + \mu_n^r)$*

To improve coverage of the integrals and reduce simulation noise, we use Halton Draws.[9]

.

---

[9]Halton draws provides a superior coverage as it induces negative correlation across individuals. In the context of discrete choice models, Bhat (2001) found in a Mixed Logit Model, that 100 Halton draws provided more precise results than 1000 standard pseudo random draw. Train (2003) provide a comprehensive and excellent treatment of several variance reduction techniques.

**TABLE A.1: MEAN, VARIANCE AND SKEWNESS
(EDUCATIONAL BREAKDOWN)**

| | | # observations | | Mean disposable income | | Variance | | | | Skewness | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Transitory effect | | Permanent effect | | Transitory effect | | Permanent effect | |
| | | se | we | se | we | se | we | se | we | se | we | se | we |
| Missing education | | 3,867 | 29,167 | 209,691 | 163,169 | 35.633 | 0.049 | 2.196 | 0.224 | 47.9 | 11.8 | 6.1 | 3.9 |
| Primary School | *Basic school* | 80,596 | 657,151 | 217,842 | 133,278 | 35.476 | 0.037 | 2.296 | 0.099 | 236.9 | 36.9 | 25.8 | 7.5 |
| Secondary school | *General* | 4,509 | 43,962 | 267,294 | 166,541 | 0.660 | 0.039 | 3.216 | 0.223 | 25.5 | 2.1 | 6.9 | 3.6 |
| | *Commercial and technical* | 1,362 | 12,906 | 353,176 | 218,601 | 1.062 | 0.100 | 6.970 | 0.806 | 11.9 | 20.8 | 6.9 | 13.1 |
| Vocational training | *Shop assistents* | 24,343 | 345,545 | 232,634 | 151,299 | 2.342 | 0.034 | 0.974 | 0.122 | 104.9 | 9.8 | 5.3 | 6.2 |
| | *Building and construction* | 18,769 | 113,541 | 225,676 | 170,977 | 0.571 | 0.024 | 0.554 | 0.072 | 28.6 | 5.1 | 7.5 | 4.4 |
| | *Metal* | 17,887 | 147,347 | 240,837 | 177,130 | 0.219 | 0.028 | 0.784 | 0.075 | 11.2 | 19.9 | 5.7 | 4.5 |
| | *Graphic* | 2,061 | 18,304 | 277,758 | 202,616 | 3.609 | 0.029 | 2.278 | 0.161 | 22.1 | 4.5 | 7.2 | 8.3 |
| | *Technical* | 2,606 | 33,878 | 158,173 | 129,978 | 1.342 | 0.035 | 1.147 | 0.097 | 18.8 | 14.5 | 3.0 | 2.7 |
| | *Service and transport* | 10,725 | 23,617 | 124,194 | 131,480 | 1.225 | 0.051 | 0.478 | 0.095 | 42.1 | 22.1 | 1.7 | 1.5 |
| | *Food* | 20,383 | 49,809 | 314,211 | 170,244 | 1.155 | 0.032 | 0.764 | 0.091 | 83.8 | 5.1 | 5.3 | 4.7 |
| | *Health care* | 2,254 | 67,273 | 138,455 | 117,762 | 0.659 | 0.026 | 0.896 | 0.052 | 14.2 | 5.2 | 3.3 | 1.2 |
| Post secondary | *Humanities and social sciences* | 1,576 | 18,953 | 185,609 | 153,600 | 0.224 | 0.039 | 2.013 | 0.141 | 2.1 | 2.5 | 3.6 | 4.5 |
| | *Technical* | 3,847 | 32,146 | 246,378 | 181,377 | 0.202 | 0.040 | 0.594 | 0.112 | 3.7 | 15.2 | 3.0 | 6.7 |
| | *Agriculture* | 917 | 8,504 | 306,245 | 157,736 | 0.218 | 0.039 | 0.697 | 0.066 | 3.2 | 5.1 | 1.1 | 1.1 |
| | *Health care* | 206 | 10,378 | 114,543 | 129,392 | 1.163 | 0.024 | 0.812 | 0.051 | 4.5 | 2.0 | 0.3 | 0.9 |
| | *Police and defence* | 327 | 15,524 | 234,734 | 202,600 | 0.194 | 0.020 | 0.621 | 0.079 | 3.7 | 6.2 | 3.2 | 3.8 |
| Higher education short cycle | *Humanities* | 3,033 | 157,352 | 174,274 | 156,319 | 7.537 | 0.021 | 0.893 | 0.054 | 48.4 | 2.6 | 3.1 | 1.5 |
| | *Social sciences* | 1,780 | 22,332 | 491,580 | 251,898 | 0.251 | 0.073 | 1.292 | 0.270 | 3.2 | 15.7 | 2.5 | 3.8 |
| | *Technical* | 3,338 | 40,915 | 308,715 | 246,404 | 0.483 | 0.041 | 1.327 | 0.126 | 17.3 | 8.7 | 8.4 | 10.5 |
| | *Health care* | 1,962 | 65,014 | 168,011 | 137,621 | 0.095 | 0.026 | 0.659 | 0.057 | 2.0 | 2.7 | 4.8 | 4.7 |
| | *Food, agriculture and transport* | 730 | 15,661 | 276,414 | 216,992 | 0.883 | 0.032 | 2.391 | 0.073 | 19.7 | 2.4 | 5.1 | 1.7 |
| | *BA* | 469 | 5,205 | 336,057 | 251,710 | 0.828 | 0.087 | 1.234 | 0.320 | 8.4 | 11.9 | 1.5 | 2.4 |
| Higher education MA level | *Humanities* | 578 | 22,592 | 209,513 | 190,457 | 0.162 | 0.038 | 0.868 | 0.068 | 2.3 | 26.6 | 1.7 | 1.5 |
| | *Natural sciences* | 159 | 9,727 | 226,141 | 217,852 | 0.179 | 0.023 | 1.104 | 0.070 | 2.8 | 1.2 | 4.0 | 3.8 |
| | *Social sciences* | 3,589 | 25,269 | 476,302 | 251,683 | 0.321 | 0.047 | 0.778 | 0.190 | 19.3 | 6.9 | 3.5 | 5.4 |
| | *Technical* | 1,912 | 18,252 | 334,441 | 255,821 | 0.553 | 0.048 | 2.193 | 0.161 | 7.1 | 17.9 | 10.3 | 7.0 |
| | *Food* | 1,431 | 5,250 | 358,845 | 228,455 | 0.122 | 0.024 | 0.446 | 0.073 | 5.8 | 2.1 | 7.6 | 2.4 |
| | *Health care* | 7,069 | 18,412 | 441,567 | 260,069 | 0.074 | 0.038 | 0.199 | 0.085 | 2.4 | 13.1 | 3.0 | 1.3 |

## TABLE A.2: SELECTION EQUATIONS
### (RESULTS FROM A BINARY PROBIT WITH RANDOM EFFECTS)

| | | Males | | | | Females | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Self-employment | | Wage-employment | | Self-employment | | Wage-employment | |
| | | Coefficient | Std. | Coefficient | Std. | Coefficient | Std. | Coefficient | Std. |
| Lagged dependent, y(t-1) | | 2.590 | 0.007 | 1.415 | 0.004 | 2.651 | 0.010 | 1.226 | 0.005 |
| Initial dependent, y(0) | | 1.962 | 0.018 | 1.918 | 0.010 | 1.641 | 0.023 | 1.540 | 0.011 |
| Age | | 0.705 | 0.071 | -0.157 | 0.044 | 0.731 | 0.096 | 0.974 | 0.047 |
| Age squared | | -0.827 | 0.082 | 0.067 | 0.050 | -0.822 | 0.110 | -1.231 | 0.053 |
| Wealth (in mio dkr, 1996 prices) | | 0.014 | 0.002 | 0.001 | 0.001 | 0.005 | 0.002 | 0.008 | 0.001 |
| No. of children aged 0-6 | | 0.049 | 0.006 | -0.019 | 0.004 | 0.049 | 0.010 | -0.098 | 0.006 |
| No. of children aged 7-17 | | 0.030 | 0.004 | -0.011 | 0.003 | 0.012 | 0.006 | 0.001 | 0.003 |
| Married | | -0.006 | 0.009 | 0.112 | 0.007 | -0.047 | 0.011 | 0.064 | 0.007 |
| Immigrant (western) | | 0.043 | 0.030 | -0.155 | 0.023 | 0.132 | 0.032 | -0.159 | 0.021 |
| Immigrant (non-western) | | 0.230 | 0.028 | -0.481 | 0.020 | 0.124 | 0.037 | -0.529 | 0.024 |
| Second generation immigrants | | 0.092 | 0.084 | -0.111 | 0.070 | 0.132 | 0.109 | 0.037 | 0.079 |
| Spouse self-employed | | 0.453 | 0.013 | 0.137 | 0.005 | 0.399 | 0.011 | 0.152 | 0.005 |
| Regional | *Copenhagen* | 0.009 | 0.017 | -0.131 | 0.012 | 0.033 | 0.022 | -0.097 | 0.013 |
| dummies | *Large city* | 0.012 | 0.016 | -0.096 | 0.012 | 0.022 | 0.020 | -0.161 | 0.012 |
| | *Rural* | 0.097 | 0.013 | -0.123 | 0.009 | 0.120 | 0.015 | -0.197 | 0.009 |
| Missing education | | 0.066 | 0.029 | -0.030 | 0.023 | 0.152 | 0.037 | -0.109 | 0.025 |
| Secondary | *General* | 0.061 | 0.026 | 0.073 | 0.020 | 0.133 | 0.034 | 0.093 | 0.022 |
| school | *Commercial and technical* | 0.070 | 0.043 | 0.153 | 0.038 | 0.002 | 0.070 | 0.190 | 0.042 |
| Vocational | *Shop assistents* | -0.038 | 0.014 | 0.176 | 0.012 | -0.023 | 0.013 | 0.155 | 0.008 |
| training | *Building and construction* | -0.025 | 0.014 | 0.067 | 0.011 | 0.180 | 0.084 | -0.011 | 0.067 |
| | *Metal* | -0.072 | 0.014 | 0.152 | 0.011 | 0.403 | 0.139 | 0.063 | 0.103 |
| | *Graphic* | -0.025 | 0.036 | 0.041 | 0.028 | 0.252 | 0.082 | -0.182 | 0.055 |
| | *Technical* | -0.010 | 0.046 | 0.087 | 0.033 | 0.071 | 0.029 | 0.000 | 0.019 |
| | *Service and transport* | 0.188 | 0.041 | -0.145 | 0.032 | 0.500 | 0.025 | -0.370 | 0.020 |
| | *Food* | 0.175 | 0.018 | -0.156 | 0.015 | 0.110 | 0.057 | 0.051 | 0.037 |
| | *Health care* | -0.143 | 0.086 | 0.418 | 0.066 | -0.099 | 0.022 | 0.351 | 0.013 |
| Post | *Humanities and social sciences* | 0.171 | 0.069 | 0.007 | 0.059 | 0.176 | 0.036 | 0.063 | 0.023 |
| secondary | *Technical* | -0.056 | 0.026 | 0.191 | 0.021 | 0.057 | 0.054 | 0.223 | 0.035 |
| | *Agriculture* | -0.050 | 0.060 | 0.126 | 0.047 | -0.172 | 0.084 | 0.439 | 0.054 |
| | *Health care* | -0.751 | 0.361 | 0.478 | 0.153 | -0.204 | 0.064 | 0.443 | 0.038 |
| | *Police and defence* | -0.425 | 0.053 | 0.763 | 0.039 | -0.422 | 0.322 | 0.532 | 0.122 |
| Higher | *Humanities* | -0.365 | 0.026 | 0.587 | 0.018 | -0.221 | 0.019 | 0.552 | 0.012 |
| education | *Social sciences* | 0.003 | 0.035 | 0.256 | 0.029 | -0.097 | 0.057 | 0.418 | 0.037 |
| short cycle | *Technical* | -0.048 | 0.024 | 0.302 | 0.019 | 0.139 | 0.118 | 0.073 | 0.078 |
| | *Health care* | 0.153 | 0.077 | 0.285 | 0.065 | -0.177 | 0.024 | 0.785 | 0.016 |
| | *Food, agriculture and transportat* | -0.258 | 0.043 | 0.478 | 0.033 | -0.095 | 0.114 | 0.277 | 0.073 |
| | *BA* | 0.043 | 0.067 | 0.170 | 0.054 | -0.238 | 0.189 | 0.324 | 0.094 |
| Higher | *Humanities* | -0.358 | 0.047 | 0.510 | 0.032 | -0.044 | 0.055 | 0.297 | 0.032 |
| education | *Natural sciences* | -0.390 | 0.069 | 0.616 | 0.051 | -0.221 | 0.122 | 0.352 | 0.065 |
| MA level | *Social sciences* | 0.198 | 0.029 | 0.073 | 0.023 | 0.195 | 0.051 | 0.218 | 0.034 |
| | *Technical* | 0.021 | 0.033 | 0.264 | 0.027 | 0.112 | 0.101 | 0.110 | 0.072 |
| | *Food* | 0.223 | 0.053 | 0.034 | 0.042 | 0.395 | 0.094 | -0.008 | 0.062 |
| | *Health care* | 0.587 | 0.031 | -0.333 | 0.027 | 0.563 | 0.039 | 0.048 | 0.029 |
| | Year dummies | Yes | | Yes | | Yes | | Yes | |
| | Constant | -4.327 | 0.153 | -0.661 | 0.098 | -4.627 | 0.208 | -2.555 | 0.104 |
| | $s_\mu$ | 0.760 | 0.009 | 0.838 | 0.005 | 0.626 | 0.011 | 0.730 | 0.005 |
| Number of observations | | 1288888 | | 1288888 | | 1126960 | | 1126960 | |
| Number of individuals | | 136990 | | 136990 | | 122749 | | 122749 | |
| Log-likelihood | | -116985.808 | | -287641.544 | | -59297.355 | | -243222.453 | |

### TABLE A.3: EARNINGS EQUATIONS
### (CORRECTED FOR SAMPLE SELECTION BIAS AND UNOBSERVED HETEROGENEITY)

| | Males | | | | Females | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | Self-employment | | Wage-employment | | Self-employment | | Wage-employment | |
| | Coefficient | Std. | Coefficient | Std. | Coefficient | Std. | Coefficient | Std. |
| Age (divided by 10) | 0.707 | 0.030 | 0.323 | 0.004 | 0.282 | 0.081 | 0.069 | 0.005 |
| Age squared (divided by 1000) | -0.834 | 0.034 | -0.350 | 0.005 | -0.410 | 0.092 | -0.111 | 0.005 |
| Married | 0.110 | 0.006 | 0.063 | 0.001 | -0.285 | 0.015 | -0.101 | 0.001 |
| Immigrant (western) | -0.278 | 0.031 | -0.081 | 0.006 | -0.264 | 0.062 | -0.046 | 0.006 |
| Immigrant (non-western) | -0.608 | 0.029 | -0.187 | 0.006 | -0.543 | 0.071 | -0.065 | 0.008 |
| Second generation immigrants | 0.008 | 0.086 | -0.027 | 0.014 | -0.098 | 0.230 | 0.009 | 0.018 |
| Regional dummies *Copenhagen* | -0.239 | 0.015 | -0.085 | 0.002 | -0.141 | 0.034 | -0.044 | 0.002 |
| *Large city* | 0.019 | 0.016 | -0.039 | 0.002 | -0.033 | 0.036 | -0.062 | 0.002 |
| *Rural* | 0.041 | 0.012 | 0.014 | 0.002 | -0.072 | 0.027 | -0.042 | 0.002 |
| Missing education | -0.068 | 0.031 | 0.086 | 0.006 | -0.098 | 0.072 | 0.063 | 0.007 |
| Secondary school *General* | 0.079 | 0.029 | 0.017 | 0.005 | 0.093 | 0.062 | 0.009 | 0.005 |
| *Commercial and technical* | 0.258 | 0.047 | 0.238 | 0.009 | -0.056 | 0.131 | 0.114 | 0.010 |
| Vocational training *Shop assistents* | 0.196 | 0.016 | 0.185 | 0.003 | 0.107 | 0.029 | 0.093 | 0.002 |
| *Building and construction* | 0.025 | 0.015 | 0.066 | 0.003 | -0.152 | 0.194 | 0.052 | 0.017 |
| *Metal* | 0.071 | 0.015 | 0.087 | 0.003 | 0.175 | 0.288 | 0.101 | 0.027 |
| *Graphic* | 0.148 | 0.041 | 0.223 | 0.007 | 0.276 | 0.172 | 0.171 | 0.017 |
| *Technical* | 0.038 | 0.051 | 0.042 | 0.009 | -0.113 | 0.061 | 0.036 | 0.005 |
| *Service and transport* | 0.002 | 0.036 | 0.084 | 0.009 | -0.013 | 0.038 | -0.028 | 0.007 |
| *Food* | 0.319 | 0.016 | 0.085 | 0.004 | 0.182 | 0.099 | 0.051 | 0.008 |
| *Health care* | 0.144 | 0.112 | 0.043 | 0.012 | 0.049 | 0.051 | 0.039 | 0.003 |
| Post secondary *Humanities and social sciences* | 0.086 | 0.072 | 0.209 | 0.013 | 0.219 | 0.067 | 0.189 | 0.006 |
| *Technical* | 0.139 | 0.030 | 0.178 | 0.005 | 0.098 | 0.121 | 0.073 | 0.009 |
| *Agriculture* | 0.316 | 0.064 | 0.218 | 0.013 | 0.071 | 0.171 | 0.135 | 0.013 |
| *Health care* | 0.141 | 0.452 | -0.021 | 0.027 | -0.044 | 0.149 | 0.113 | 0.009 |
| *Police and defence* | 0.242 | 0.076 | 0.208 | 0.007 | 0.215 | 0.755 | 0.086 | 0.022 |
| Higher education short cycle *Humanities* | 0.133 | 0.040 | 0.189 | 0.004 | 0.122 | 0.052 | 0.224 | 0.003 |
| *Social sciences* | 0.660 | 0.041 | 0.426 | 0.006 | 0.443 | 0.151 | 0.276 | 0.008 |
| *Technical* | 0.300 | 0.029 | 0.401 | 0.005 | 0.324 | 0.207 | 0.385 | 0.020 |
| *Health care* | 0.279 | 0.097 | 0.180 | 0.014 | 0.468 | 0.058 | 0.219 | 0.004 |
| *Food, agriculture and transport* | 0.181 | 0.061 | 0.341 | 0.008 | 0.299 | 0.255 | 0.191 | 0.017 |
| *BA* | 0.127 | 0.068 | 0.314 | 0.011 | 0.484 | 0.313 | 0.163 | 0.015 |
| Higher education MA level *Humanities* | 0.233 | 0.080 | 0.336 | 0.007 | 0.247 | 0.109 | 0.445 | 0.007 |
| *Natural sciences* | 0.354 | 0.114 | 0.387 | 0.009 | -0.066 | 0.378 | 0.476 | 0.015 |
| *Social sciences* | 0.773 | 0.034 | 0.486 | 0.006 | 0.719 | 0.109 | 0.476 | 0.007 |
| *Technical* | 0.376 | 0.039 | 0.477 | 0.007 | 0.564 | 0.198 | 0.435 | 0.020 |
| *Food* | 0.634 | 0.054 | 0.428 | 0.013 | 0.613 | 0.200 | 0.420 | 0.023 |
| *Health care* | 0.959 | 0.028 | 0.663 | 0.008 | 1.257 | 0.069 | 0.627 | 0.010 |
| $s_{a\mu}$ | -0.062 | 0.003 | -0.112 | 0.001 | -0.024 | 0.005 | -0.076 | 0.002 |
| $s_{e?}$ | -0.012 | 0.004 | 0.010 | 0.001 | 0.092 | 0.008 | 0.013 | 0.001 |
| Constant | 15.109 | 0.068 | 15.783 | 0.010 | 15.779 | 0.180 | 16.323 | 0.011 |
| | *Summary Statistics* | | | | | | | |
| Number of observations | 168782 | | 1037089 | | 47496 | | 994058 | |
| Number of individuals | 27544 | | 122754 | | 10456 | | 117511 | |
| $s_a$ | 0.6656 | | 0.2929 | | 0.9831 | | 0.3092 | |
| $s_e$ | 0.4557 | | 0.1737 | | 0.5914 | | 0.1869 | |
| Fraction of variance due to individual specific error | 0.68 | | 0.74 | | 0.73 | | 0.73 | |
| R-squared | 0.12 | | 0.20 | | 0.11 | | 0.23 | |

# ENTREPRENEURSHIP, SAVINGS AND CREDIT CONSTRAINTS

BERTEL SCHJERNING

*October, 2006*

ABSTRACT. In this paper, I develop an intertemporal model for saving, consumption, human capital accumulation and occupational choice in the presence of liquidity constraints, income uncertainty, and entry costs. The paper extends the literature on entrepreneurship to provide an explanation of the phenomenon of frequently observed transitions. The model generates a well-defined transition pattern between entrepreneurship and ordinary wage work, where workers may transit back and forth between wage employment and entrepreneurship. This feature is consistent with empirical findings of sequential entries and exits. The model also holds quite different implications for optimal consumption and saving behavior compared to existing papers within the consumption saving literature.

## 1. Introduction

Over the past couple of decades, a substantial literature on entrepreneurship has developed. In this literature, the individual decision to become entrepreneur is probably the issue that has received most attention.

The existing theoretical models explaining the decision to become an entrepreneur have focused on the importance of: i) individual differences, including differences in risk aversion (Kihlstrom and Laffont (1979); Cramer and Praag (2001)); differences in probability assessments (de Meza and Southey (1996)) and differences in entrepreneurial ability (Brock and Evans (1986); Holmes and Schmitz (1990), and Fonseca, Lopez-Garzia, and Pissarides (2001)); and ii) institutional features such as tax schemes, set-up costs and labour market institutions (Kihlstrom and Laffont (1983); Fonseca, Lopez-Garzia, and Pissarides (2001); and Malchow-Møller, Markusen, and Skaksen (2005)).

A particular focus, however, has been on the importance of credit constraints and how these interact with wealth, risk aversion and entrepreneurial ability (Evans and Jovanovic (1989); Holtz-Eakin, Joulfaian, and Rosen (1994); Blanchflower and Oswald (1998); and Dunn and Holtz-Eakin (2000)).

In the large empirical literature the relationship between wealth and entrepreneurship has received an equal amount of attention. We can identify at least two competing explanations for the observed concentration of wealth among entrepreneurs. First, it may just be due to higher incomes earned by entrepreneurs. Second, it may be that at least part of it reflects that entrepreneurs are selected mainly from richer families because credit constraints prevent low-wealth individuals from entering entrepreneurship (see Hurst and Lusardi (2004) for a recent contribution and a critical discussion of this literature).

From a policy point of view, disentangling the relative powers of these explanations is extremely important as they have quite different policy implications. Thus, if credit constraints are to blame for the observed concentration of wealth, this could potentially justify intervention in credit markets. Evans and Jovanovic (1989) seek to quantify the

importance of credit constraints in a static structural model. In this model, liquidity constraints discourage some people from starting up a business and those who become entrepreneurs after all use less capital. Individuals with relatively high entrepreneurial ability are most likely to be credit constrained, since they are assumed to require the highest level of capital for their businesses. The key empirical finding in Evans and Jovanovic (1989) is that credit constraints are quantitatively important and have very large welfare costs.

However, successful policy intervention in this area requires a better understating of the mechanisms at work here. While previous studies have enhanced our knowledge about the decision to become entrepreneur, most of this work seems to neglect important dynamic/intertemporal aspects of entrepreneurial behavior. One example is the structural model in Evans and Jovanovic (1989). Since the model is static, it ignores the possibility of saving to overcome liquidity constraints and the accumulation of entrepreneurial human capital. In a dynamic context, high ability individuals facing borrowing constraints will make an effort to overcome these constraints by saving. This illustrates the importance of allowing for intertemporal incentives when analyzing entrepreneurial behavior.

Economic decisions faced by entrepreneurs have long-run implications and intertemporal incentives are therefore likely to underlie much of the observed behavior by entrepreneurs. As an example, decisions to accumulate wealth and to set up a business are typically founded in expectations about the future. Due consideration of such intertemporal aspects is thus crucial for an improved understanding of what makes people become entrepreneurs, their behavior and in turn the optimal design of policies regarding, e.g., bankruptcy laws, business start-up schemes, and public loan guarantees.

In the more recent literature, there has been an increasing recognition of the importance of intertemporal aspects in entrepreneurial decision making. Quadrini (2000) , Cagetti and DeNardi (2006) and Buera (2003) have thus developed intertemporal models of entrepreneurship in the presence of credit constraints. Despite these recent advances in

entrepreneurship literature, the theoretical understanding of dynamic behavior of entre-
preneurs is still very limited. In terms of analyzing the dynamics of the entrepreneurial
choice, the main limitation of Quadrini (2000) and Cagetti and DeNardi (2006) is that
they only analyze the equilibrium around a steady state - not the individual saving and
transition behavior. The model in Buera (2003) generates well defined transition pattern
of individuals moving from wage work to entrepreneurship. However, this model cannot
explain that workers may transit back and forth between wage employment and entre-
preneurship – in part because he do not appropriately account for the effects of income
and demand uncertainty.

This paper develops a theory of occupational choice, i.e. the decision to become
entrepreneur, in the presence of credit constraints and set-up costs. It extends the
existing literature to provide a better explanation of observed saving and transition
behavior. This is done by including the effects of income/demand uncertainty, which not
only is an empirically relevant extension, but also affects both wealth accumulation and
transitions, as it may cause entrepreneurs to close down their businesses and return to
wage employment. I end up with an intertemporal model for saving, consumption, human
capital accumulation and occupational choice in the presence of liquidity constraints,
income uncertainty, and set-up/entry costs. More specifically, I assume that an infinitely
lived individual maximizes a time-separable utility function by each period choosing
between entrepreneurship and wage work, where transitions between occupations are
associated with a cost, and by dividing his resources between consumption, savings
and transition costs. Hence, the model developed explicitly incorporates intertemporal
incentives with respect to both the occupational choice and wealth accumulation and
takes into account that future income is uncertain.

The paper addresses a number of unresolved questions in relation to the dynamic be-
havior of entrepreneurs: i) Can we rationalize observed transition patterns, characterized
by continuous cycling between occupations? ii) How do start-up costs influence entrepre-
neurial saving incentives and decisions to entry and exit?, iii) What are the implications
of set-up costs for the dynamics of the occupational choice? iv) How are entrepreneurial

saving incentives and transition patterns affected by the tightness of credit constrains, and how do these effects vary across individuals with different levels of entrepreneurial ability, entrepreneurial human capital, and asset holdings? v) How does human capital accumulation affect entrepreneurial behavior?

This paper contributes to the literature on entrepreneurship by delivering a plausible explanation of observed transition patterns and an improved understanding of the intertemporal incentives. The model is interesting in its own right as it may provide an explanation for observed transitions between occupations and since it holds quite different implications for optimal consumption and saving behavior than the existing papers within the consumption saving literature (see e.g., Caroll (1997) and Deaton (1991)).

The rest of the paper is organized as follows. Section 2 describes the dynamic decision problem faced by the individual. In this section, I also give an analytical characterization of the solution and specify the individual optimization problem as a stochastic dynamic programming problem. Section 3 presents the numerical solution of a simple version of the model without human capital accumulation, where I focus on the implications of entry costs and credit constrains. Section 4 presents numerical solutions of the full model. In this section, I discuss the implications of two polar cases of human capital accumulation: i) when accumulation takes place only in entrepreneurship, and ii) when accumulation takes place only in wage work. Section 5 concludes and discuss directions for future research.

## 2. A Dynamic Model of Occupational Choice

2.1. **The Model.** We begin with a basic framework that builds on the intertemporal model introduced by Deaton (1991) of saving and consumption under liquidity constraints. The new feature is that income is not exogenous, but depends crucially on wealth and occupational choice.

In each period, individuals choose a level of consumption that maximizes a time-separable infinite-horizon utility function

$$E_t \left( \sum_{\tau=t}^{\infty} (1+\delta)^{t-\tau} u(c_\tau) \right)$$

where $\delta > 0$ is the subjective discount rate, $c_t$ is consumption at time $t$, $E_t$ summarizes expectations given the information available at time $t$ and $u(\cdot)$ is an instantaneous utility function defined over current consumption. The instantaneous utility is assumed to be of the CRRA form: $u(c_\tau) = (1-\eta)^{-1} c_\tau^{1-\eta}$, with $\eta > 0$

At the end of period $t$, the individual has assets $a_t$ and receives income $y_t$. The sum $x_t = a_t + y_t$, "cash on hand", is then divided between consumption in period $t$, $c_t$, savings, $s_t = (1+r)^{-1} a_{t+1}$ and possibly costs of switching occupation, $\phi(i_t, i_{t+1})$. Savings, $s_t$ earn interest, $r$, which become assets in the following period.

Hence, the evolution of liquid assets $a_t$ is governed by

$$a_{t+1} = (1+r)(a_t + y_t - c_t - \phi(i_t, i_{t+1}))$$

It is assumed that individuals are liquidity constrained, implying that liquid wealth can never fall below zero

$$(2.1) \qquad\qquad a_t \geq 0, \quad \forall\, t$$

To make the model a model of occupational choice, individuals choose among two mutually exclusive work alternatives: Entrepreneurship, $e$, or wage-employment, $we$. Compared to wage work, entrepreneurship is a fundamentally different occupational choice with respect to the source of income. Wage-workers inelastically supply one unit of labor at an uncertain market wage, $w\varepsilon_t^{we}$, where $\varepsilon_t^{we}$ summarizes the uncertainty in wage income and is distributed according to a truncated normal with mean 1 and variance $\sigma_{we}^2$. To ensure a bounded state space, $\varepsilon_t^{we}$ is defined on the bounded support $[\underline{\varepsilon}^{we}, \bar{\varepsilon}^{we}]$. Entrepreneurs, on the other hand, derive income from production. Hence, the state dependent income is given by:

$$y_t(i_t, h_t, a_t, \varepsilon_t(i_t)) = \begin{cases} \pi(h_t, a_t, \varepsilon_t^e) & \text{if } i_t = we \\ w\varepsilon_t^{we} & \text{if } i_t = w \end{cases}$$

where $\pi\left(h_t, a_t, \varepsilon_t^e\right)$ denotes the entrepreneurs profit function given the level of entrepreneurial human capital, $h_t$, liquid assets $a_t$ and the productivity chock, $\varepsilon_t^e$. The chock, $\varepsilon_t^e$ and entrepreneurial human capital, $h_t$ are assumed to affect the productivity of the entrepreneur directly, while liquid assets operate indirectly through possibly binding capital constraints. If entrepreneurs are capital constrained, they must use own wealth to finance their investments. Therefore the wealth the of entrepreneur affects the efficiency scale of the business.

The model also allows for an explicit role for business start-up costs. This feature has previously been analyzed in Fonseca, Lopez-Garzia, and Pissarides (2001) where a standard matching model with matching between workers and managers is used to shed light on the general equilibrium effects of start-up costs on employment and entrepreneurial activity in the economy. To my knowledge, however, this paper is the first to provide an analysis of how entry costs alter the intertemporal incentives that underlie the decision to become entrepreneur. The transition costs function, $\phi\left(i_t, i_{t+1}\right)$, is specified as:

$$
\phi\left(i_t, i_{t+1}\right) = \begin{cases} \phi^{entry} & \text{if } i_t = we \text{ and } i_{t+1} = e \\ \phi^{exit} & \text{if } i_t = e \text{ and } i_{t+1} = we \\ 0 & \text{otherwise} \end{cases}
$$

Hence if an individual switches from wage work to entrepreneurship, transition costs equals $\phi^{entry}$. Conversely, $\phi^{exit}$ are transition cost associated with closing down a business.

The model also has a specific role for human capital accumulation. While previous (empirical) studies have treated entrepreneurial experience and work experience as exogenously assigned to individuals, see e.g. Hamilton (2000), in this paper, experience (or learning by doing) will be treated as a behaviorally determined investment decision. In the spirit of the human capital literature, see e.g. Ben-Porath (1967), Blinder and Weiss (1976), and Keane and Wolpin (1997), human capital accumulation is determined jointly

with occupational choice decisions.[1] In the present context, I think of human capital as being productive only in entrepreneurship, to capture the idea that while individuals can acquire managerial/entrepreneurial skills in both wage work and entrepreneurship activities, they are only useful in the latter activity. Entrepreneurial human capital, $h_t$, is assumed to evolve according to

$$h_{t+1} = \gamma h_t + \Delta\left(i_t\right)$$

where $\gamma < 1$ is the depreciation rate of experience and $\Delta\left(i_t\right)$ is the amount of entrepreneurial human capital gained in occupation $i_t$

If individuals choose to run their own business, they must devote their entire labor endowment to operate the business and have to decide how much capital to invest in the business. As soon as the occupational choice is made, the investment decision is purely static. Entrepreneurs derive income from the production of a single homogeneous good according to a Cobb-Douglas production function $f\left(h_t, k_t\right) = \theta k_t^{\alpha_k} h_t^{\alpha_h} \varepsilon_t$, defined over two production factors - entrepreneurial human capital, $h_t$ and the amount of capital invested in the business, $k_t$. Individuals are assumed to differ with respect to their initial level of assets $a_0$ and their entrepreneurial ability, $\theta$.

Once the investment decision is made, the entrepreneur receives a realization of the stochastic element of production, $\varepsilon^e$. The disturbance $\varepsilon^\varepsilon$ summarizes the uncertainty in entrepreneurial income. $\varepsilon^e$ is assumed to be independent and identically distributed with bounded support $[\underline{\varepsilon}, \overline{\varepsilon}]$, mean, $\mu_\varepsilon = 1$, and variance, $\sigma_\varepsilon^2$. As we shall see later, the assumption about boundedness is necessary to ensure a compact state space.

If $k_t > a_t$, the entrepreneur is a net borrower and must rent the remaining capital at a fixed interest rate, $r$. However, in line with Evans and Jovanovic (1989), it is assumed that entrepreneurs can only borrow up to an amount proportional to the stock

---

[1]Keane and Wolpin (1997) study the career decisions of young men in a finite horizon model, where individuals can choose between schooling, three work alternatives, and retirement. While early contributions for simplicity assumed that human capital is homogeneous, in Keane and Wolpin (1997) skills are assumed to be occupational specific and their returns vary across occupations.

of liquid assets $a_t$. Letting the factor of proportion being $\lambda - 1$, where $\lambda \geq 1$, a potential entrepreneur faces the credit constraint

(2.2)
$$k_t \leq (\lambda - 1) a_t + a_t, \; \forall \; t$$

or:

$$k_t \leq \lambda a_t, \; \forall \; t$$

If $\lambda = 1$, the entrepreneur must finance all activities in the business from the holding of liquid assets, $a_t$, while there are no liquidity constraints when $\lambda \to \infty$.

This assumption can be motivated by an underlying market friction, where loan contracts are imperfectly enforceable. Cagetti and DeNardi (2006) explicitly model this type market friction, and find that it generates endogenous entrepreneurial borrowing constraints. In this setup, own wealth act as a collateral to reduce the incentive to default: The larger the amount, the entrepreneur is able to finance from own wealth, the larger the amount the creditor is able to recover. Therefore, the amount the entrepreneur is able to borrow increases with liquid asset holdings, $a_t$

At the tome when the investment is made, the entrepreneur cannot observe or foretell the value of the idiosyncratic income chock $\varepsilon_t^e$. Thus, the investment decision is taken conditional on the level of entrepreneurial ability, $\theta$, human capital, $h_t$, and liquid assets, $a_t$. In each period, the entrepreneur therefore derives his optimal investment by solving the following maximization problem

(2.3)
$$E_\varepsilon \pi \left( h_t, a_t, \varepsilon_t^e \right) = \max_{k_t \leq \lambda a_t} \left( \theta k_t^{\alpha_k} h_t^{\alpha_h} - r k_t \right)$$

At an interior maximum, the first order condition is

$$\alpha_k \theta h_t^{\alpha_h} k_t^{\alpha_k - 1} - r = 0$$

By the concavity of (2.3), the optimal level of capital can thus be written as

$$k_t^* = \min \left\{ \lambda a_t, \left[ \frac{\alpha_k \theta}{r} h_t^{\alpha_h} \right]^{1/(1 - \alpha_k)} \right\}$$

For entrepreneurs to be unconstrained we must have

$$
\begin{aligned}
k_t^* \quad &< \quad \lambda a_t \Rightarrow \\
\theta \quad &< \quad (\lambda a_t)^{1-\alpha_k} \frac{r}{\alpha_k h_t^{\alpha_h}}
\end{aligned}
$$

(2.4)

Since marginal productivity of capital is increasing in entrepreneurial ability, $\theta$ more able individuals are more likely to be credit constrained.

In sum, the profit function for entrepreneurs can be written as

$$
\pi \left(h_t, a_t\right) = \min \left\{ \theta \left(\lambda a_t\right)^{\alpha_k} h_t^{\alpha_h} \varepsilon_t - r \lambda a_t \;,\; \theta \left[\frac{\theta \alpha_k}{r} h_t^{\alpha_h}\right]^{\frac{\alpha_k}{1-\alpha_k}} h_t^{\alpha_h} \varepsilon_t - r \left(\frac{\alpha_k \theta}{r} h_t^{\alpha_h}\right)^{1/(1-\alpha_k)} \right\}
$$

If entrepreneurs are credit constrained, entrepreneurial earnings depend on individual wealth, $\alpha_t$, while earnings is independent of the level of assets if they are not, i.e. if $k_t < \lambda a_t$.

To summarize: Given current occupation, $i_t$, cash on hand, $x_t$ and entrepreneurial human capital, $h_t$ and the state dependent income function, $y_t \left(i_t, h_t, a_t, \varepsilon_t \left(i_t\right)\right)$, individuals optimally choose i) assets to carry over for the following period, $a_{t+1}$ and ii) future occupation $i_{t+1}$ to maximize a discounted stream of utility.

## 2.2. Characterization of the Solution.

I start by characterizing the solution of the model by inspecting the first order conditions for the intertemporal allocation - the Euler equation. Even though it is not possible to derive a complete analytical solution for the model, the Euler equation provide a convenient way to characterize some of the mechanisms in the model - analytically.

Since the occupational choice is discrete, individuals face only one continuous intertemporal choice, the savings decision. Therefore we can only derive one state dependent Euler equation originating from the first order condition with respect to assets in the following period, $a_{t+1}$

(2.5)
$$
u' \left(c_t\right) = \frac{1+r}{1+\delta} E_t \left\{ u' \left(c_{t+1}\right) \left[1 + \frac{dy_{t+1} \left(i_{t+1}, h_{t+1}, a_{t+1}, \varepsilon_{t+1} \left(i_{t+1}\right)\right)}{da_{t+1}}\right]\right\} + (1+r) \mu_{a_{t+1}}
$$

where $\mu_{a_{t+1}} \geq 0$ is the Lagrange multiplier associated with the borrowing restriction in (2.1).

The first order condition in (2.5) states that in optimum, it should not be possible to increase utility through a reallocation of consumption via $a_{t+1}$. Hence, marginal utility of consumption today (the left hand side) should equal the sum of i) the discounted expected marginal utility consumption in the next period, corrected for the change in future income due to the change in assets and the difference between the subjective and the objective discount rate; and ii) the shadow price of the liquidity constraint.

To identify the different savings motives is important to distinguish between the two types of borrowing constraints.

(1) The *liquidity* constraint faced by all individuals, preventing individuals to smooth consumption perfectly if income fluctuations occur.

(2) The *credit* constraint faced by entrepreneurs with a relatively low level of initial assets.

In the absence of these borrowing constraints, individuals would smooth out consumption so that discounted expected marginal utility is equalized across time periods. This results in the well known consumption/income divergence, which can be explained with essentially the same logic Friedman used long time ago (see Friedman (1957)): consumption does not respond one-for-one to transitory shocks to income because assets are used to buffer consumption against such shocks. This is referred to as the life-cycle saving motive.

However, in the presence of borrowing constraints, individuals reduce consumption today to overcome the expected utility loss induced by either of the two borrowing constraints This leads to two additional saving motives in the model: i) a precautionary saving motive and ii) a entrepreneurial saving motive[2].

---

[2]As pointed out in Kimball (1990) a key theoretical requirement to produce precautionary savings is prudence of the value function, $V(x)$. Formally, Kimball (1990) defines prudence of the value function as $-V'''(x)/V''(x)$ or equivalently the convexity of the marginal value function at $x$. The precautionary

In the present model, $u(.)$ is CRRA, with coefficient $\eta > 1$. Thereby, $E_t u'(c_{t+1}) \geq u'(c_{t+1})$ and therefore individuals reduce consumption today to overcome the expected utility loss induced by the binding liquidity constraint. In other words, individuals save to buffer against future negative income shocks.

Note that in periods where none of the liquidity constraints are binding $\mu_{a_{t+1}} = 0$ and $k_{t+1} > \lambda a_{t+1}$ the first order condition collapses to a standard Euler equation, where discounted expected marginal utility is equalized over time. However, as pointed out by Deaton (1991), even when liquidity constraints do not bind in a given period, this does not imply that the optimal saving policy coincides with the policy function from the problem without liquidity constraints. The reason is that individuals anticipate that liquidity constraints could be binding in the future. This illustrates that the Euler equation is not a sufficient condition for optimal behavior. Rather, it puts restrictions on the allocation of resources between two successive periods.

With respect to the entrepreneurial motive, credit constraints have an additional effect. Due to the presence of credit constraints, a reallocation of current consumption into future assets, $a_{t+1}$ adds additionally to future consumption through an expected increase in future income

$$\frac{dy_{t+1t}(i_t, h_t, a_t, \varepsilon_t(i_t))}{da_{t+1}} = \begin{cases} \frac{d\pi_t(,h_t,a_t,\varepsilon_t^e)}{da_{t+1}} > 0 & \text{if } i_{t+1} = e \text{ and } k_{t+1} = \lambda a_{t+1} \\ 0 & \text{if } i_{t+1} = we \text{ or } k_{t+1} < \lambda a_{t+1} \end{cases}$$

The reason for the additional saving motive is that entrepreneurs who operate at a suboptimal level of capital due to binding credit constraints can expect an increase in profits if they save more, $\frac{d\pi(h_t, a_{t+1})}{da_{t+1}} > 0$. Just like the precautionary motive, if an individual knows that it will ever be optimal to enter entrepreneurship, this savings motive is relevant at all times (due to the recursive nature of the first order condition).

The entrepreneurial saving motive depends crucially on the factor of proportion $\lambda - 1$, at which entrepreneurs can borrow. The following proposition states how saving incentives are affected by changes in $\lambda$

---

motive is present only if individuals are prudent,. i.e. if the marginal value function at $x_t$ is convex. This is the case if the marginal instant utility is convex, i.e. $u'''(c) > 0$.

**Proposition 1.** *The entrepreneurial saving motive is affected by $\lambda$ in a non-monotone way: For small values of $\lambda$, relatively productive individuals with relatively low asset holdings will increase their savings when $\lambda$ increases. On the other hand, for $\lambda$ large enough, i.e. when credit constraints become less binding, the expected return to increased savings approaches zero.*

*Proof.* See appendix □

The intuition behind proposition 1 goes as follows: Initially, as $\lambda$ increases some highly productive individuals will find it optimal to save more as the perspectives for (future) entrepreneurship becomes better. As $\lambda$ is further increased, credit constraints eventually become irrelevant, thereby lowering the incentive to save.

Proposition 1 has important implications for the understanding the effect of changes in credit policy, e.g. government loan guaranties. According to proposition 1, relatively productive individuals will increase savings, while less productive individuals decrease savings. On the one hand, this kind of policy will increase wealth inequality in the economy and could be associated with increased probability of default. On the other hand, increasing $\lambda$ also increases the probability of entry relatively more for productive individuals - due to the increased willingness to use savings to overcome credit constraints.

2.3. **The Dynamic Programming Problem.** In specifying this as a dynamic programming problem, note that the state variables $x_t, s_t, h_t$ summarize all information about the past that bears on current and future decisions. Since $y_t$ is assumed to be *iid* distributed conditional of $h_t$, $a_t$ and $i_t$, only the sum $x_t = a_t + y_t$ "cash on hand" is relevant for current and future saving decisions and occupational choice. Note also that the problem is stationary in the sense that optimal choices do not depend on time *per se*. Hence, time subscripts can be dropped. To discriminate between the current and future periods, I therefore denote next period variables with a prime. The resulting Bellman equation can thus be formulated as

(2.6) $$V(x, h, i) = \max_{i', a' \in \Lambda(x, i, h)} u(x, i, a', i') + \beta E_t \left[ V(x', h', i') | x, h, i \right]$$

where $E_t$ summarizes expectations give the information available at the time the decision is made and the value function,$V(x, h, i)$, is the maximum expected discounted utility obtainable by the agent in the given state $(x, h, i)$. The Bellman equation express the recursive relationship between the value function in the current period, $V(x, h, i)$, current utility, $u(x, i, a', i')$, and expectations over the value function in the following period, $E_t[V(x', h', i')|x, h, i]$. Hence, individuals choose $i', a' \in \Lambda(x, i, h)$ to maximize the sum of current utility and discounted expected future utility.

Current utility is

$$(2.7) \qquad u(x, i, a', i') = u\left(x - (1+r)^{-1} a' - \phi(i, i')\right)$$

and $\Lambda(x, i, h) : \mathbb{S} \to \mathbb{D}$ is a correspondence that summarizes the feasible choice set

$$(2.8) \qquad \Lambda(x, i, h) = \left\{ (a', i') \in \mathbb{D} : \begin{array}{l} 0 \le a' \le (1+r)(x - \phi(i, i')) \\ h' = \gamma h + \Delta(i) \\ i' \in I = \{e, we\} \end{array} \right\}$$

To make notation a bit more compact, let $\mathbb{D}$ denote the set of controls and let $\mathbb{S}$ denote the state space, such that

$$(2.9) \qquad \begin{aligned} d &= (a, i) \in \mathbb{D} \\ s &= (h, x, i) \in \mathbb{S} \end{aligned}$$

Furthermore, $\Lambda(s) \subseteq \mathbb{D}$ is the non-epmty set of feasible controls that summarizes the contingent constraints on the controls $d'$ in state $s$ and $u : \mathbb{S} \times \mathbb{D} \to \mathbb{R}$ is the current pay-off function in given the current state $s \in \mathbb{S}$ and given the control $d' \in \mathbb{D}$ is applied in the following period. Finally let $f(s'|s, d')$ be the probability density that $s' \in S$, i.e. the conditional density that future state $s'$ occurs given current state and control $d'$. We can now express the infinite horizon, discounted, time separable dynamic programming problem in more compact notation

$$V(s) = \sup_{d' \in \Lambda(s)} u(s, d') + \beta \int V(s') f(s'|s, d') \, ds'$$

Solving the model is equivalent to finding a fixed point of the Bellman equation (2.6). Under certain conditions, a unique solution exists and successive iterations on the Bellman equation will guarantee global convergence to this solution. Roughly speaking, these condition will be met if the subjective discount factor is less than unity, $\beta < 1$, the state space, $\mathbb{S}$ is a compact set and the value function is bounded on this set.

I start by formulating two propositions, stating that is possible to restrict attention to a compact subset of the state space $\mathbb{S}^3$.

**Proposition 2.** *There exists $h_{high} < \infty$ such that if $h_t \leq h_{high}$ then $h'$ satisfies $h' \leq h_{high}$*

*Proof.* See appendix. $\square$

In other words, there exists an upper level of $h$, where the depreciation of experience exceeds the human capital gain in any occupation. Therefore $h_t$ is bounded above.

**Proposition 3.** *Given $h \leq h_{high}$ where $h_{high}$ satisfies the proposition 2, there exists an $a_{high} < \infty$ such that if $a \leq a_{high}$ then the optimal choice of $a'$ satisfies $a' \leq a_{high}$*

*Proof.* See appendix. $\square$

Hence, there exists some (finite) upper level of assets, $a^{high}$, where individuals who for some reason own more than this level, will stop saving.

The intuition behind Proposition 3 goes as follows: Since marginal productivity of capital is decreasing and since shocks to production has bounded support, marginal returns to savings will approach the interest rate, $r$ as $a$ increases, and earnings, $y$ will be bounded from above and below. Since individuals are impatient, in the sense that $\delta > r$, for $a$ large enough, it will be optimal to stop saving as the life-cycle motive will dominate both precautionary and entrepreneurial saving motives.

---

[3] The compactness of state space for the discrete occupational choice $i_t$ is trivial as it can only take two values.

We are now ready to formulate conditions that guarantee that the considered dynamic programming has the contraction mapping property. I formalize this in the following proposition which is stated without a formal proof

**Proposition 4.** *Let $\mathbb{S}$ be defined by (2.9) with values of $h_{high}$ and $a_{high}$ satisfying propositions (2) and (3). Let $u : \mathbb{S} \times \mathbb{D} \to \mathbb{R}$ and $\Lambda : \mathbb{S} \to \mathbb{D}$ be given by (2.7) and (2.8) respectively. Furthermore, let $f\left(s'|s, d'\right) = f\left(x'|s, d'\right)$ be a continuous density function defined on a bounded support $[x^{low}, x^{high}]$ such that $c^{low} = x^{low} - \left(1 + r\right)^{-1} a^{high} - \phi\left(i, i'\right) > 0$. Then the mapping defined by*

$$(2.10) \qquad \Gamma\left(V\right)\left(s\right) \equiv \sup_{d' \in \Lambda(s)} u\left(s, d'\right) + \beta \int V\left(s'\right) f\left(s'|s, d'\right) ds'$$

*is a contraction mapping $\Gamma : \mathbb{B} \to \mathbb{B}$ taking a complete normed vector space (i.e. a Banach space) of functions from $\mathbb{S} \to \mathbb{R}$. The nonlinear operator $\Gamma$ has a unique fixed point $V = \Gamma\left(V\right)$ and for any $V_0 \in \mathbb{B}$*

$$(2.11) \qquad \left\|\Gamma^k V_0 - V\right\| \leq \beta^k \left\|V_0 - V\right\|, \quad k = 1, 2, ...$$

Under the conditions stated in the proposition above, the dynamic programming problem has a unique fixed point and successive value function iterations will converge to the unique solution.

As mentioned, a rigorous proof will not be given here. Instead I will try to give an intuitive reasoning: First, $u$ is bounded from below as long as $c = x - \left(1 + r\right)^{-1} a' - \phi\left(i, i'\right) > 0$ and bounded above if $s = (h, x, i) \in \mathbb{S}$ is bounded. With values of $h_{high}$ and $a_{high}$ satisfying propositions (2) and (3), it is necessarily the case that $s = (x, h, i)$ stays within a compact set. Secondly, since the continuous density $f\left(x'|s, d'\right)$ has bounded support $[x_{low}, x_{high}]$ with $x_{low} = y_{low} > 0$, consumption can always be sustained above zero.

I.e. for all values of $x \in [x_{low}, x_{high}]$ it is always feasible to chose $a'$ such that $c > 0$ and consequently $u$ is bounded from above and below and the integral $\int V\left(s'\right) f\left(s'|s, d'\right) ds'$ is therefore well-defined. Third, the correspondence $\Lambda\left(s\right) \subseteq \mathbb{D}$ is non-epmty and compact valued. Finally, the effective discount factor is below one $\beta = 1/\left(1 + \delta\right) < 1$ (by

assumption). To get the intuition clear: it is necessary to bound the support of the distribution of the disturbance, $\varepsilon$ from below such that income, and thereby consumption, are bounded from below too. Otherwise the expectation of the value function may not be well-defined. Under these assumptions plus some regularity conditions, it follows that $\Gamma$ satisfies the contraction mapping theorem, see Stokey and Lucas (1989)[4]

To solve the model, we have to find a fixed point of the functional equation in (2.6). I use chebyshecv polynomials to represent the value function over the continuous state space. Since the value function has discontinuous first derivatives in the swiching point, I use *piecewise* Chebyshev polynomials to approximate the value function with an edogenously determined join point at the kink of the value function. The use of Chebyshev polynomials to approximate the value function has one important spin-off. Once the model has been solved, Chebyshev approximation of the value function can be utilized to express the policy function in any point of the state space, at almost zero computational cost. .

In line with Rust (1987) a combination of successive contraction iterations and the Newton-Kantorowich algorithm will be used. While contraction iterations guarantee convergence due to the contraction mapping property, the procedure slows down when the approximation errors $\|V_k - V\|$ become small.[5] In contrast, Newton-Kantorowich iterations are not guaranteed to converge, but converge in a quadratic rate in the neighborhood of the solution[6]. The resulting fixed point algorithm known as the *poly-algorithm,*

---

[4]It actually turns out that one of the conditions is violated in the present context. Since one of the state variables is discrete, the requirement that the state space is a convex set - is obviously not satisfied. However, the conditions stated in Stokey and Lucas (1989) are *sufficient* conditions and thus too restrictive in the present model. A more general version of the theorem is available in Denardo (1967)

[5]It follows directly from equation (2.11) in proposition 4 that the upper bound on the approximation error $\beta \|V_k - V\|$ dcreases linearly in $\|V_k - V\|$ (making convergence particular slow for $\beta$ close to 1).

[6]Kantorovich's Theorem guarantees that given a starting point $V_0$ in a domain of attraction of the fixed point $V$ of $\Gamma$ the Newton Kantorovich iterations will converge to $V$ at a quadratic rate (see Rust (1996))

TABLE 1. Baseline Parameter Values

| $r$ | $\eta$ | $\delta$ | $\theta$ | $\alpha^h$ | $\alpha^k$ | $\sigma_\varepsilon^e$ | $\lambda$ | $w$ | $\sigma_\varepsilon^{we}$ | $\gamma$ | $\Delta$ | $\phi^{exit}$ | $\phi^{entry}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.04 | 1.5 | 0.07 | 0.8 | 0 | 0.33 | 0.3 | 1 | 1 | 0.1 | $-$ | $-$ | 0 | 0 |

combines these two algorithms in order to balance robustness versus speed of convergence.

## 3. NUMERICAL RESULTS - THE CASE WITHOUT HUMAN CAPITAL ACCUMULATION

In this section, I present numerical solutions of the model. For the purpose of exposition, I will first consider a simpler version of the model where I assume that $\Delta(i) = 0$ and $h_t = 1$. The model without human capital accumulation will serve as a useful starting point, when explaining some of the key features of the model: In particular, I will discuss: i) how highly productive potential entrepreneurs can use savings to overcome binding liquidity constraints; ii) how entry and exit costs affects savings decisions and the transition between the two occupations; and iii) how individuals depending on their initial wealth and entrepreneurial productivity approach two different equilibria in the long run.

### 3.1. **Baseline Calibration.** Rather than trying to calibrate the model to observed data in order to give quantitative predictions about behavior, the baseline parameters are chosen to identify the mechanisms of the model. The baseline values used in the numerical simulations are listed in Table 3.1.

*Utility parameters:* The first two parameters will be set with little controversy. Taking the time period to be one year, I let the real interest rate of $r = 0.05$ reflect the average market return to wealth. I choose $\eta = 1.5$ as a reasonable value of the the inverse of the intertemporal elasticity of substitution (see e.g. Caroll (1997) and Deaton (1991)). Due to the functional form of the instantaneous utility function, the relative risk aversion and

the intertemporal rate of substitution are inversely related.[7] A choice of $\eta = 1.5$ therefore implies that agents are risk averse and slightly prudent - which seems empirically sensible. Harrison, Lau, and Rutstrom (2004) estimate individual risk attitudes using controlled field experiments in Denmark. Their results indicate that the average Dane is risk averse, and that risk neutrality is an inappropriate assumption to apply. They also find that risk attitudes vary significantly in the population roughly within a range of $\eta \in [0, 2]$.

I set the time preference rate $\delta = 0.07$ to be larger than the real interest rate to reflect relatively impatient agents. In the empirical savings literature, the rate of time preference has been estimated much higher. This also applies for the recent literature on experimental economics: Harrison, Lau, and Williams (2002) estimate annual individual discount rates with respect to time to be around $0.25$[8]. Deaton (1991) used $\delta = 0.1$ in the simulations of his model. Quadrini (2000) used $\delta = 0.9$ in a dynamic model with entrepreneurial savings calibrated to the US wealth distribution. However, if $\delta$ is very high relative to $r$ agents become very impatient and the incentives to accumulate assets will be almost zero. Therefore, in order to better illustrate the savings incentives in the model, I set $\delta = 0.07$.

*Income parameters:* For convenience, mean wages are normalized to one, $w = 1$ with a standard deviation of $\sigma_\varepsilon^{we} = 0.1$. I set $\alpha^k = 0.33$ approximately equal to the structural estimates of return to capital in Evans and Jovanovic (1989) and $\sigma_\varepsilon^e = 0.3$ is chosen in

---

[7]Note that this is only the case if we think of household's preferences over consumption gambles in a static context. This interpretation of $\eta$ has been subject to much criticism (see e.g. Flavin and Nakagawa (2005)) In a dynamic context, it is more relevant to think of relative risk aversion as the change in the curvature of the value function, i.e.

$$RRA = \frac{\partial^2 V(x, h)/\partial x^2}{\partial V(x, h)/\partial x} x > 0$$

Because the household's degree of risk aversion depends on the curvature of the value function, behavior towards income risks will not only depend on the curvature of the instantaneous utility function - also the state variables. In particular, very wealthy individuals will tend to be less risk averse. Therefore $\eta$ is sometimes referred to as the curvature parameter.

[8]These estimates may reflect attitudes to risk also.

accordance with their structural estimates of the dispersion in entrepreneurial earnings. Contrary to the estimates of Evans and Jovanovic (1989), I choose $\sigma_\varepsilon^{we} < \sigma_\varepsilon^e$ to reflect the very compressed wage structure in Denmark (see e.g. Malchow-Møller, Markusen, and Skaksen (2005)).

*Framework conditions:* In the baseline scenario, I set $\lambda = 1$ to reflect binding credit constraints. Hence it is not possible to borrow any funds for starting up a business. Entry and exit costs are set equal to zero, $\phi^{entry} = \phi^{exit} = 0$. In what follows, we shall see how changes in these parameters influence savings incentives, occupational choice, income etc.

Figure 1 displays the numerical solution of the value function in the baseline scenario. Two vertical lines mark two threshold levels of cash on hand: The leftmost line marks the reservation value of cash on hand, $x_r$, where individuals will choose to enter entrepreneurship. The rightmost line marks the level of cash on hands where entrepreneurs will be unconstrained, $x_u$. To the right of this line, the real interest rate would exceed the marginal product of capital - if all available funds were invested.

It should be apparent from Figure 1, that the value function, does not display the standard properties of concavity and differentiability. The value function is the upper envelope of two underlying value functions associated with each of the two occupational choices. In the crossing point, the indifferent individual switches occupation. The convexity around the kink of the value function is induced by the introduction of the investment opportunities in the model. This investment option ads an extra component to the marginal returns to savings and therefore individuals facing this option will have an entrepreneurial saving motive.

As we have discussed earlier, the value function has a kink in the crossing point. This is precisely what makes the model solution non–trivial and what causes the Chebyshev approximation method to perform poorly. Therefore I use piecewise Chebyshev polynomials with a single join point in $x_r$, which is continuously updated at each iteration. This effectively avoids numerically unstable and imprecise solutions with occilirations in the policy functions and occasionally breakdowns in algorithm.

FIGURE 1. VALUE FUNCTION
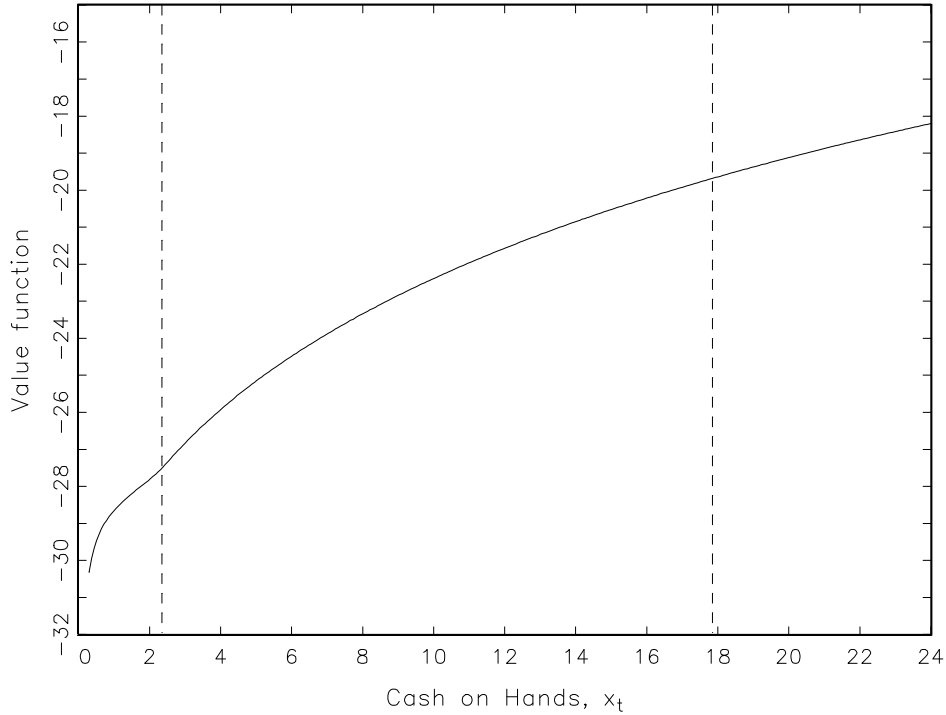


3.1.1. *Policy functions.* Figure 2 present the policy function for optimal consumption. Again, the two vertical lines mark the threshold values $x_r$ and $x_u$. Starting from the left, we see that consumption equals cash on hands as long as liquidity constraints are binding i.e. $a_{t+1} = 0$. The individual would actually like to consume more today at the expense of tomorrow's consumption. But since liquid assets can never fall below zero, this is impossible. Due to the precautionary savings motive, the consumption policy function starts to bend off in a slightly concave way around $x_t = 1$. The segment below $x = 1.5$ on the policy function coincides with the saving behavior implied by the model in Deaton (1991): Precautionary saving arises from the possibility that constraints might bind in the future. Therefore individuals use precautionary savings as an insurance against future negative shocks.

At some point, before the switching point, $x_r$, optimal consumption drops significantly: Individuals with this level of cash on hands starts to save to become entrepreneurs. In the absence of credit constraints, these individuals would have been entrepreneurs (In fact all

FIGURE 2. OPTIMAL CONSUMPTION
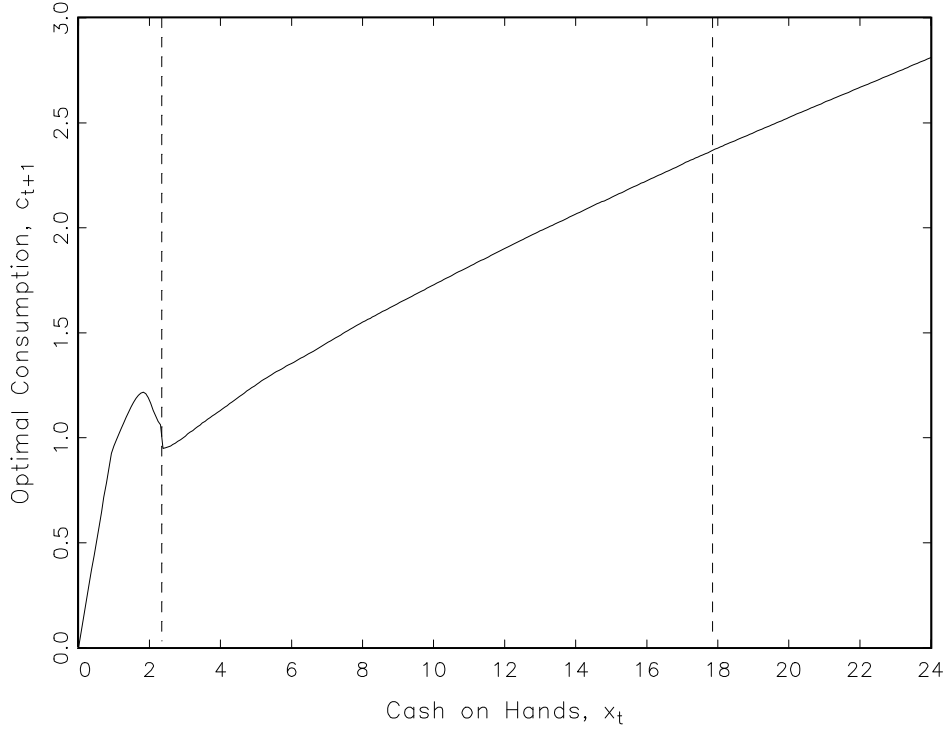


individuals would be entrepreneurs at the baseline parameter values). Instead, they use savings to overcome binding liquidity constraints thereby seeking opportunities for higher future income. At the point where the individual chooses to enter entrepreneurship, the consumption policy function drops discretely to a local minimum. Since the entrepreneur is credit constrained and therefore have to operate at a suboptimal level of capital, the effective return to savings will exceed the market interest rate. Therefore, the individual will save more and consume less.

Figure 3 displays the relationship between wealth (cash on hand) and expected future gross income, $E[y_{t+1}|x_t] + ra_{t+1}$ which is the sum of expected earnings in the next period and interest on liquid assets. The individual earns the fixed wage, $w = 1$ until he chooses to become entrepreneur at the switching point, $x_r$. For low values of $x_t$, individuals are liquidity constrained, and will therefore not save any assets for the following period, i.e. $a_{t+1} = 0$. Hence, expected future gross earnings equals $w = 1$. Hereafter, individuals first start to save for precautionary reasons (around $x_t = 1$), then for entrepreneurial

FIGURE 3. EXPECTED FUTURE GROSS INCOME
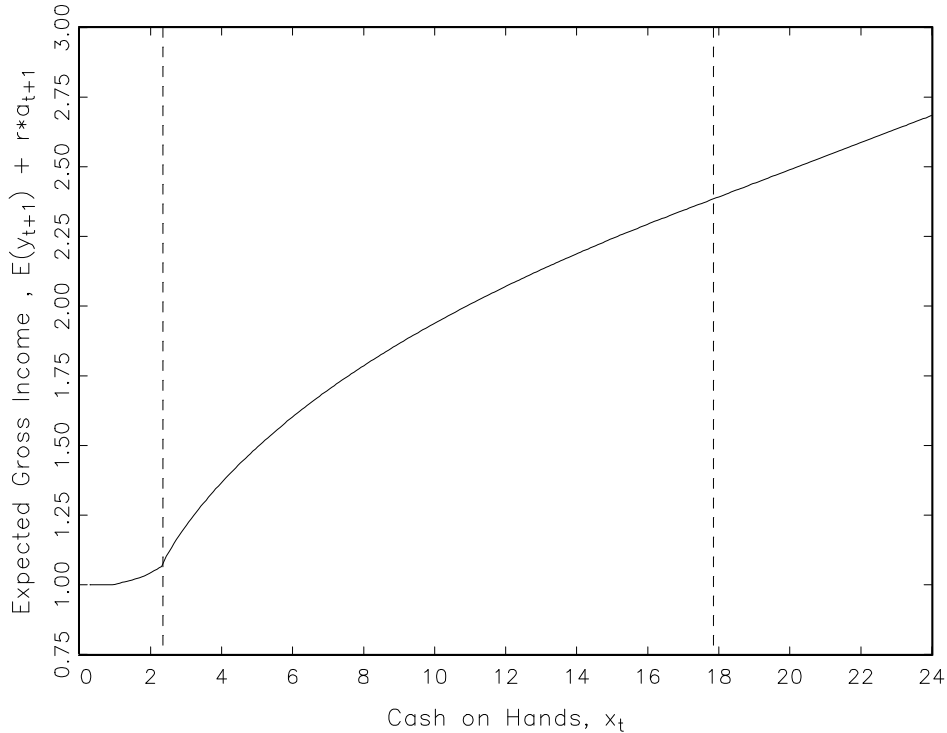


reasons (around $x_t = 1.8$). Correspondingly, asset returns, $ra_{t+1}$ starts to increase. Between the two vertical lines, i.e. when $x_t \in [x_r, x_u]$ the entrepreneur is credit constrained and operates at a suboptimal level of capital. As we move towards $x_u$, the business becomes more capital intensive and the marginal product drops until the business reaches its unconstrained scale, where marginal product equals the real interest rate. Below the point $x_u$, entrepreneurs chose to invest all available assets in their business, therefore entrepreneurial earnings increase with cash on hand until the entrepreneur is unconstrained with respect to capital, i.e. when $x_t = x_u$. At this point, the entrepreneur is able to self-finance the investments needed to operate the business at the optimal scale and entrepreneurial income (net of interest) is independent of cash on hands.

3.1.2. *Simulated Sequences.* Using the same set of baseline parameter values, stochastic model simulations are used to characterize the evolution of the state dependent variables: liquid assets, consumption and gross income. In order to characterize individual saving incentives, first note that there exists a threshold level of cash on hand, such that the

FIGURE 4. SIMULATIONS OF $c_t$, $y_t + r\alpha_t$ AND $\alpha_t$, $a_0 = 0.5$
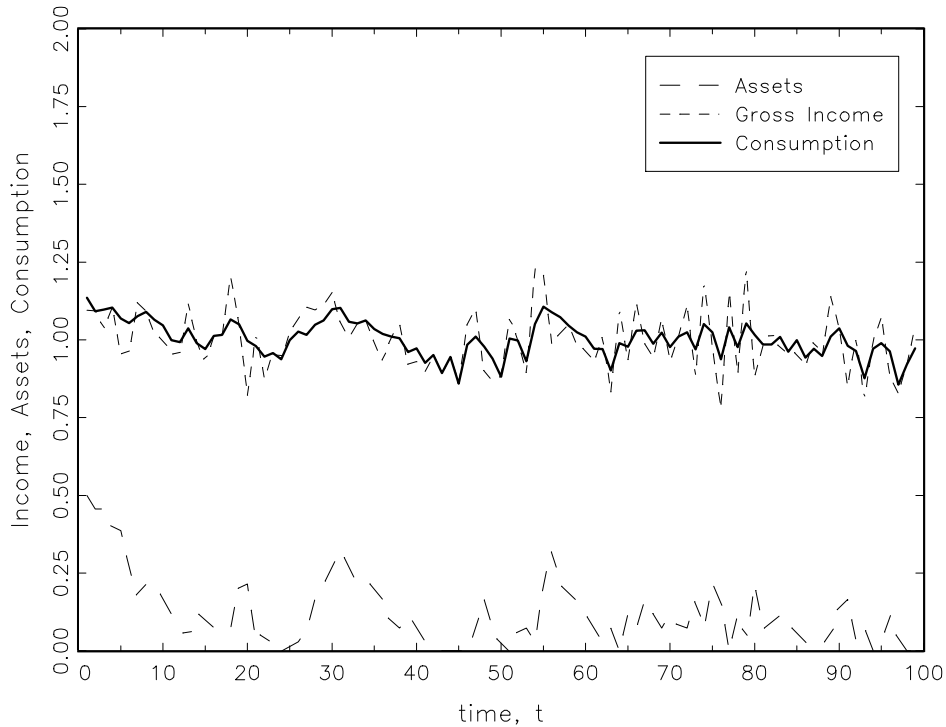


individuals with cash below this threshold, $x_t < x_{ns} \in [0, x_r]$ will not save to become entrepreneurs. Unless a sequence of unanticipated positive income chocks occurs, these individuals will instead follow a path that converges to a stationary equilibrium, where the individual remains a wage worker and keeps small levels of precautionary savings as a buffer against negative income shocks. Thus, the entrepreneurial saving motive is dominated by the incentive to smooth consumption over time and relative impatience induced by the relatively high discounting of utility, $\delta > r$.

Individuals with cash on hand above this level, i.e. $x_t \geq x_{ns}$, expect to become entrepreneurs at some point in their carrier. Depending on their current level of available funds, these individual will either save or dis-save to reach an equilibrium level of cash on hands, $x_{ss} \in [x_{ns}, \bar{x}]$ where they will stop saving.

To illustrate how the evolution in individual income, wealth and consumption are affected by the level of initial assets, I simulate sequences of these variables for two different levels of initial assets. The sequences are displayed in Figures 4 and 5.

Consider first the sequences displayed in Figure 4. The individual enters the labor market as wage worker with a relatively low level of initial assets, $a_0 = 0.5$. These initial conditions result in a realized level of cash on hands below the threshold, $x_{ns}$. Hence, rather than saving to become entrepreneur, this individual find it optimal to remain a wage worker unless an unanticipated sequence of positive income chocks is realized. After a couple of periods, the simulated sequence of $a_t$ has decreased to the stationary equilibrium, where precautionary savings are used as a buffer against negative income chocks. This individual behaves very much like the consumers in Deaton (1991). Hence, the following characteristics apply: First, consumption is notably smoother than income. Secondly, the downward spikes in consumption when liquid wealth stock-outs occur, are generally larger that the corresponding upward peaks. Consumption is therefore asymmetric, in the sense that mainly negative shocks are transmitted into consumption, whereas savings are used to smooth out positive shocks.

The displayed sequences in Figure 5 are associated with an individual entering the labor market as a wage worker with an intermediate level of initial wealth, $a_0 = 0.75$. Several things are worth noting. First, wage workers with an intermediate level of cash on hands, $x_t \in [x_{ns}, x_u]$ will have to save for several periods before entry to entrepreneurship is profitable. In fact, the wage worker associated with the simulated sequence in Figure 5 does not enter entrepreneurship until after seven periods of wage work. Hereafter, the entrepreneur keeps saving until the entrepreneurial saving motive is dominated by impatience, i.e. the incentive to smooth consumption over the life-time.

Secondly, as the business becomes more profitable, consumption increases gradually over time to reach a higher equilibrium level.

Third, despite a very fluctuating entrepreneurial income, consumption is remarkably smoother for entrepreneurs than for wage-workers due to the higher stock of assets. This is due to the simplifying assumption that investments undertaken by the entrepreneur are fully reversible. Since invested assets are perfectly liquid, savings have a dual role: As working capital and to smooth out consumption.

FIGURE 5. SIMULATIONS OF $c_t$, $y_t + r\alpha_t$ AND $\alpha_t$, $a_0 = 0.75$
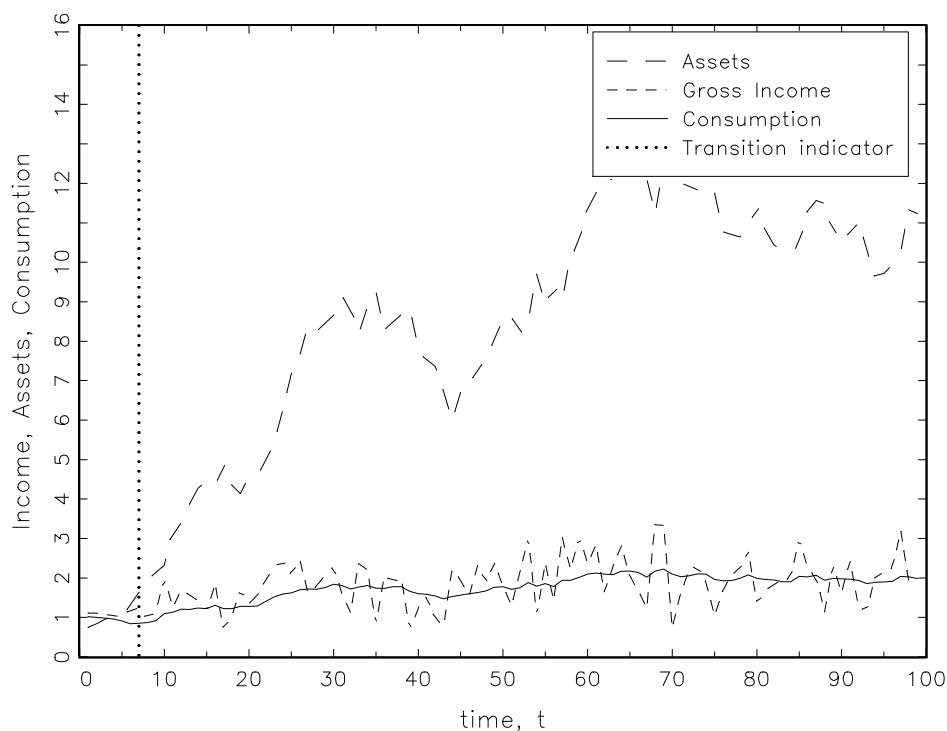


This is opposed to Fafchamps and Pender (1997), where poor households fail to undertake a profitable investment that they could, in principle, self-finance because non-divisibility and irreversibility of the investment put it out of their reach. In the literature of investment under uncertainty, (see Dixit and Pendyck (1994)) it is emphasized that uncertainty works to decrease investments when investments are irreversible. I expect similar results would be found in this paper if entrepreneurial investments were (partly) irreversible. I.e. entrepreneurs would postpone some of the investments in the business and keep a buffer of liquid assets to smooth out consumption. In that sense, the combination of irreversibility and uncertainty introduces a value of waiting. A specific analysis of this phenomenon would require an additional state variable (invested capital) and is therefore not pursued here.

3.1.3. *Transition Patterns.* The model also generates well-defined transition patterns between entrepreneurship and ordinary wage work. To illustrate this aspect of the dynamics in the occupational choice, consider a population of individuals who enter the labor market as wage workers at an intermediate level of individual wealth, $a_0 = 0.75$.

Initially, these individuals will start saving to become entrepreneurs. Depending on the realized sequence of stochastic wages, some will fall into a 'poverty trap' and remain wage workers; others will save and enter entrepreneurship within a couple of periods. As soon as they become entrepreneurs, they accumulate capital to make the business more profitable and more resistant to negative shocks. Hence, individuals 'cycle' between the two occupational alternatives until they either fall into a poverty trap or accumulate enough assets to run a profitable business.

The implied transition behavior for these initially homogeneous agents can be analyzed by inspecting the occupational specific hazard functions displayed in Figure 6. The hazard functions are calculated as follows: The conditional probability that an individual exits a given initial state after a duration of $\tau$ periods, given $\tau$ periods of survival. Hence, the hazard function by construction will sum to one over $\tau$.

Consider first the hazard out of entrepreneurship (the solid curve). Since, all individuals are initially wage workers saving to become entrepreneurs, the hazard function for entrepreneurs, is based on individuals that voluntarily entered entrepreneurship with a relatively low level of assets. Depending on the realized sequence of stochastic production shocks, some individuals will succeed in accumulating enough capital to resist negative shocks to production. In addition to the primary function as working capital, these assets serve as a buffer stock against the poverty trap. Therefore, the probability of exit to wage work will decrease with duration in entrepreneurship, i.e. the hazard function exhibits true negative duration dependence.

Now turn to the hazard function for wage workers (the dashed curve). Individuals who enter entrepreneurship will on average need 3 or 4 periods in wage employment to accumulate enough assets to start a business. If a wage worker receives a poor sequence of incomes for a longer period, available cash on hands falls below the threshold where it

FIGURE 6. SIMULATED HAZARD FUNCTIONS, $\alpha_0 = 0.75$

Nonparametric Hazard
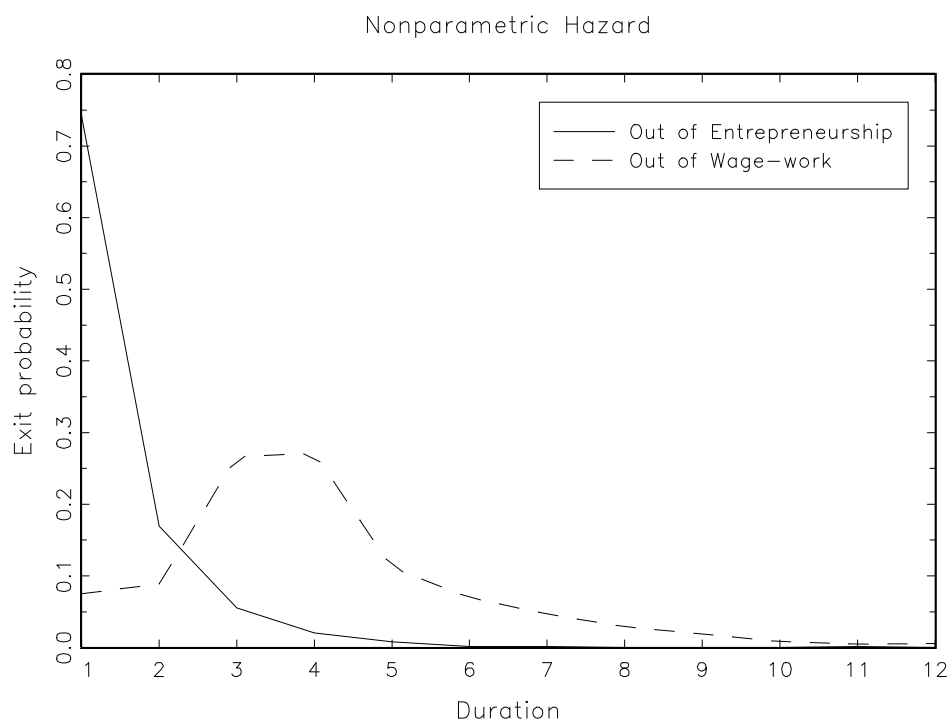


is no longer optimal to save to become an entrepreneur. Therefore, the exit probability decreases with duration. Note also that some individuals who re-entry wage work after a short duration in entrepreneurship, will relatively quickly re-exit to entrepreneurship[9]. Therefore an increased concentration around $\tau = 1$.

One important insight from the baseline scenario is that some individuals who expect to become entrepreneurs in the future, will save for several periods and accumulate a considerable amount of assets. As a back drop, note that precautionary savings models in Deaton (1991) and Caroll (1997) suggest that higher income uncertainty should lead to a higher level of precautionary savings. Previous authors has used this feature of the model to identify the level of precautionary savings from the cross sectional correlation between income risk measures and wealth holdings, in order to quantify the importance

---

[9]Of course, this is not immediately evident from the hazard function. However, as closer inspection of the simulated sequence reveals this pattern

of precautionary savings (see Browning and Lusardi (1996) for a survey of empirical applications of the intertemporal consumption model). The conclusions from these studies are very heterogeneous and consensus on the importance of precautionary savings has not yet been reached. For example, based on PSID data, Caroll and Samwick (1998) find that precautionary savings can account for as much as 40 pct. of the total wealth accumulation, while others find limited or no evidence for precautionary savings. As pointed out by Hurst, Lusardi, Kennickell, and Torrealba (2005), the observed correlation between wealth and income risk is spurious: Since entrepreneurs generally face higher income risks and may hold larger proportions of wealth for other reasons than precautionary savings, the correlation between wealth and income risk are simply an artifact of pooling together wage workers and entrepreneurs. In fact, controlling for entrepreneurial status, Hurst, Lusardi, Kennickell, and Torrealba (2005) find much lower levels of precautionary savings.

The results in this paper highlights the importance of conditioning on occupational status as entrepreneurial savings may constitute a significant share of total wealth for entrepreneurs. However, as we have seen *potential* entrepreneurs will start to save several periods prior to entry. Therefore, not only the current occupation is important for wealth accumulation also expected future occupations. Hence to mitigate a potential heterogeneity bias, we must appropriately control for expectations about future occupations too. Since these expectations are generally unobserved (and time varying), such conditioning is in general very difficult. Hence, to appropriately account for the composition of household savings, a full structural estimation of the present model can be a useful and perhaps necessary identification strategy.

3.2. **The Case with Business Start-up Costs.** The purpose of this section is to study how the presence of start up costs influence occupational choice and saving incentives. Figure 7 displays the numerical solution of the value function under the following model specification: Wage-workers deciding to become entrepreneurs incur an entry cost in the

FIGURE 7. VALUE FUNCTION, $\phi^{entry} = 0.2$
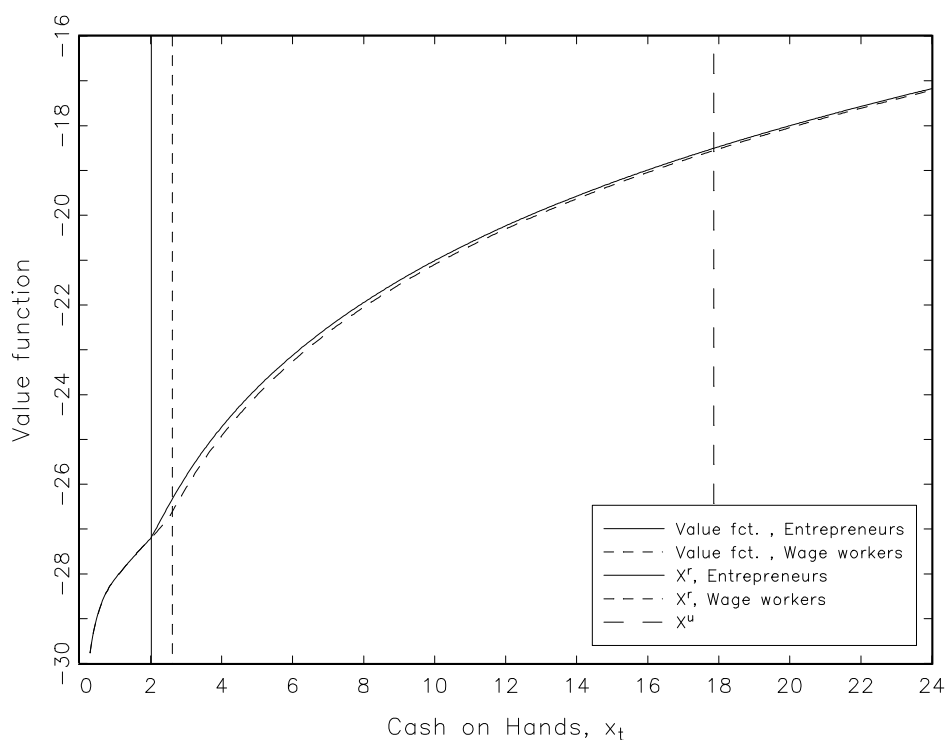


order of 20 pct, of expected wage income, i.e. $\phi^{entry} = 0.2$. The parameters of the model are otherwise identical to the baseline specification.

To analyze the impact of entry cost, it is instructive to provide a few comparative remarks: First, in absence of transition costs (and human capital accumulation), wage workers and entrepreneurs with the same values of cash on hand; face the exact same future opportunities. Since, the state variable $x_t$, summarizes all information about the past that bears on current and future decisions, only the variable $x_t$ is relevant for current and future saving decisions and occupational choices. Therefore, the value functions for wage-workers and entrepreneurs are identical equal in the absence of transition costs. Contrary to this, in the presence of start up costs, the value function is specific to the current occupation. In the following remark an important implication of this finding is formulated

**Remark 1.** *A sequence of occupational choices will not exhibit true state dependence unless transition costs exist. Hence, in the absence of transition costs, choosing a given occupation today, does not alter conditional choice probabilities in the future, i.e.*

$$P\left(i_{t+1}, x_{t+1} | x_t, i_t\right) = P\left(i_{t+1}, x_{t+1} | x_t\right)$$

Not only is the reservation value of cash on hand, $x_r$ higher for wage workers compared to entrepreneurs, $x_r^{we} > x_r^e$. The difference is larger than the entry cost, i.e. $x_r^{we} - x_r^e > \phi^{entry} = 0.2$. The explanation of this phenomenon is two-fold: In order to avoid losing invested entry costs, wage workers will postpone entry until they have accumulated enough assets to resist negative production chocks. Entrepreneurs, however, are willing to postpone exit to avoid paying the entry cost associated with potential re-entry to entrepreneurship. Consequently, being an entrepreneur represents an option value in terms of a wait-and-see option. If entry costs are large, the value of this option is important enough to make entrepreneurs willing to accept temporary income losses to keep their position and mitigate expected future entry payments. Therefore, compared to the case of no entry costs the reservation values $x_r^e$ decreases. I summarize these findings, in the follow in remark

**Remark 2.** *In the presence of entry costs, we will se later entry and later exit. Due to the indivisibility and irreversibility of the entry costs, wage workers wish postpone investments in a business, whereas entrepreneurs are willing to cut consumption temporarily to keep their position as entrepreneurs. The combination of irreversibility, indivisibility and uncertainty introduces a value of waiting.*[10]

Figures 8 and 9 displays the policy function for consumption and the implied expected future gross income, $E\left(y_{t+1}\right) + ra_{t+1}$. Notice that the policy function for wage workers and entrepreneurs coincide for $x_t < x_r^e$. Regardless of the current occupation, the decision maker knows that he will be wage worker in the following period. Therefore, the

---

[10]This results is a common find in models of investment under uncertainty; see e.g., Dixit and Pendyck (1994), Fafchamps and Pender (1997), and Malchow-Møller and Thorsen (2005).

FIGURE 8. OPTIMAL CONSUMPTION, $\phi^{entry} = 0.2$
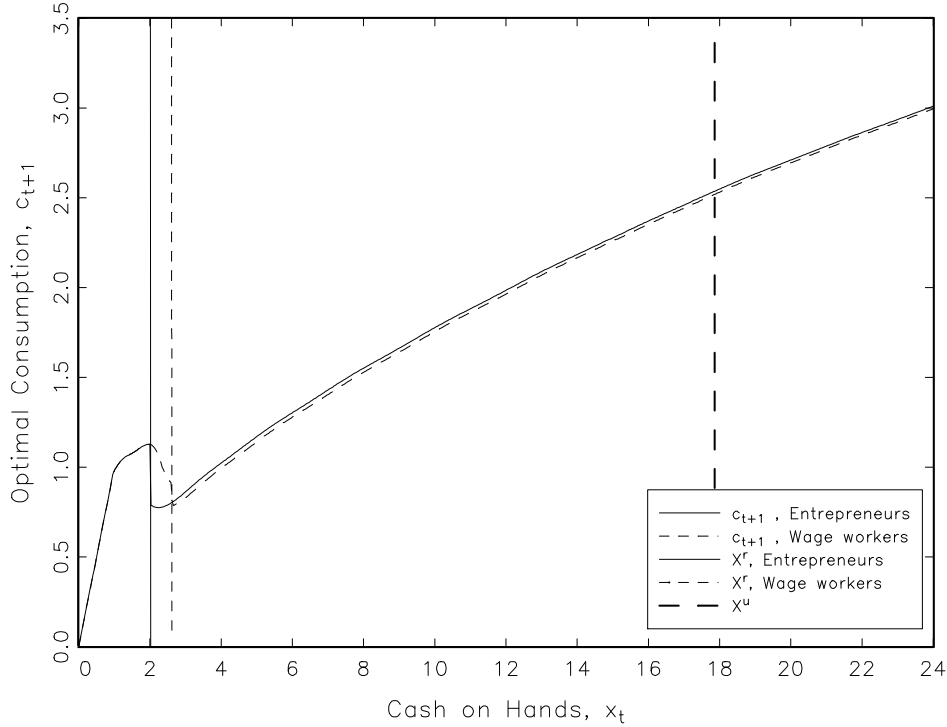


same saving motives apply in both occupations and thus the division of cash on hand between savings and consumption is with equally identical. At $x_t = x_r^e$ consumption drops discretely for entrepreneurs. In fact, entrepreneurs are willing to reduce consumption with 0.34 to maintain the business (corresponding to 34 pct of expected annual income as a wage worker or 70 pct more than the start-up cost)

The discontinuity in the policy function is due to the fact that the return to saving changes discretely at the switching point, $x_r^e$. If $x_t > x_r^e$ entrepreneurs know with certainty that they will be entrepreneurial the following period too. Therefore $dy_{t+1}/da_{t+1} = d\pi_{t+1}/da_{t+1} > 0$. If, on the other hand, $x_t < x_r^e$, they will exit entrepreneurship and become wage workers, $dy_{t+1}/da_{t+1} = 0$ for $x_t < x_r^e$.

Contrary to entrepreneurs, consumption decreases smoothly for wage workers right before the switching point $x_r^{we}$. As $x_t$ approaches $x_r^{we}$ the probability of future entry increases. Therefore, *expected* future returns to savings increase gradually. This is not

FIGURE 9. EXPECTED FUTURE GROSS INCOME, $\phi^{entry} = 0.2$
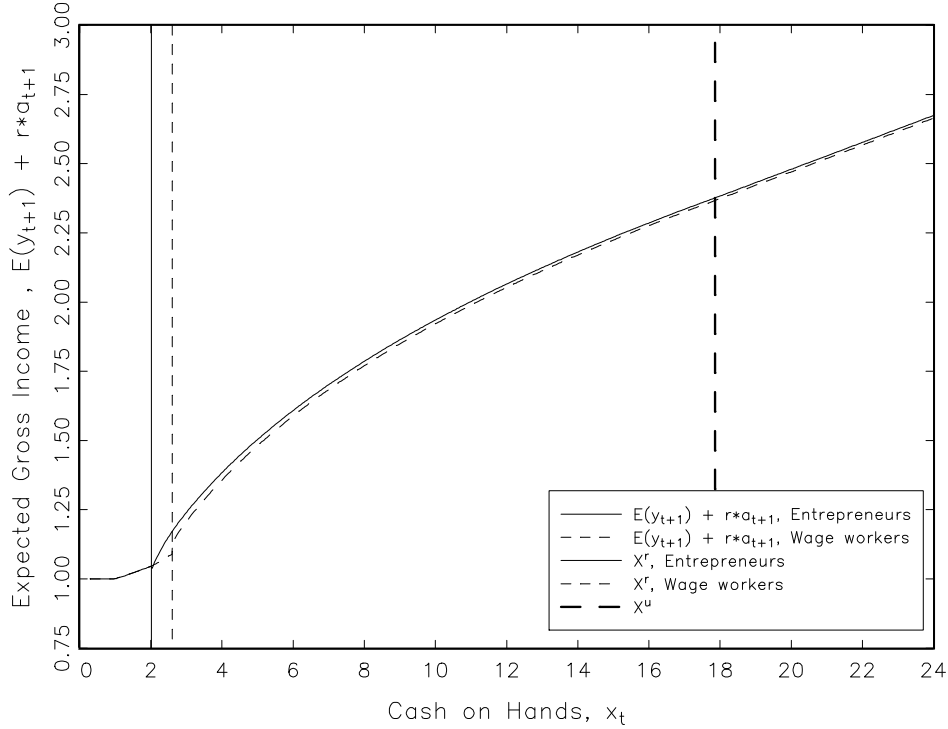


the case for entrepreneurs with $x_t \in [x_r^e, x_r^{we}]$, since the entrepreneur knows *with certainty* that he will be an entrepreneur in the following period.

Again, it is evident from the consumption function for both entrepreneurs and wage workers, that saving incentives are very strong when $x_t > x_r^e$. In fact, the more liquidity constrained, the stronger the saving incentive. Note finally that entrepreneurs can consume more, since they have paid the entry cost already. Therefore, as $x_t$ increases, the two policy functions converges due to increased ability to smooth out the entry cost over several periods.

To summarize how start up costs influence saving incentives and aspects of occupational choice, I formulate the following remark:

FIGURE 10. SIMULATED HAZARD FUNCTION, $\phi = 0.2$ AND $a_0 = 1.25$
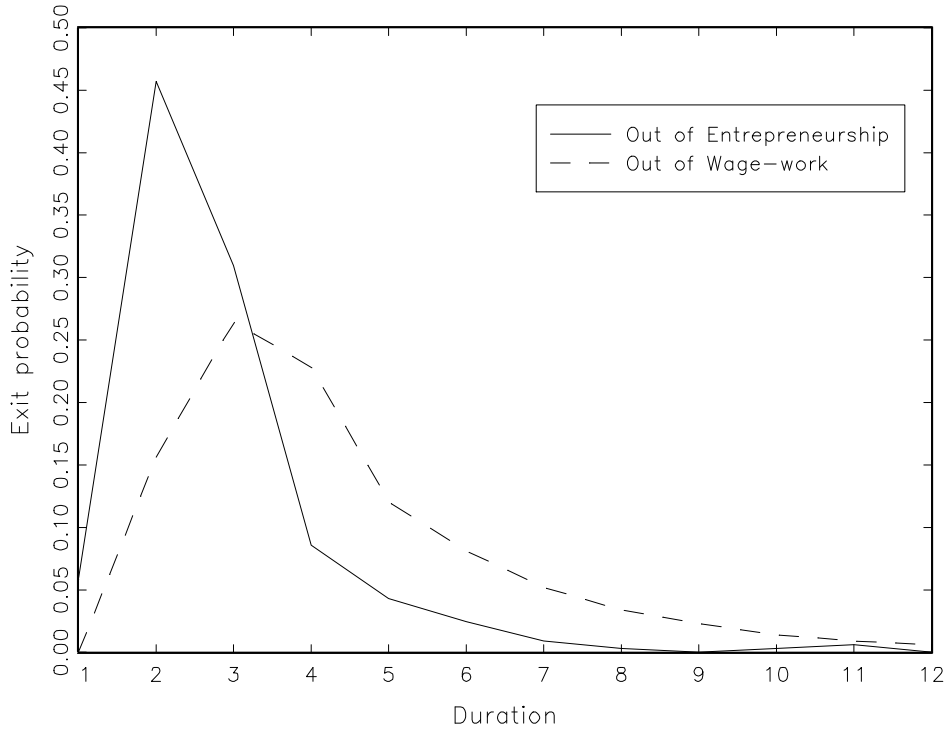


**Remark 3.** *Start up costs give an extra savings motive when credit constraints are binding: Wage workers who expect to enter entrepreneurship save to overcome entry-costs and the corresponding risk associated with entry. Entrepreneurs save to maintain their position as entrepreneurs to avoid potential costs associated with later re-entry.*

Figure 10 shows occupational specific hazard functions, for a population of individuals who enter the labor market as wage workers at an intermediate level of individual wealth, $a_0 = 1.25 \in [x_{ns}, x^e_{we}]$[11]. The presence of entry cost alters the transition patterns between entrepreneurship and ordinary wage work fundamentally: Due to start up costs, individuals will never enter entrepreneurship if there is a significant risk that they will not be able to maintain their business in the following periods. Therefore, the

---

[11]The hazard is not directly comparable with the hazard from the previous section, since two different populations are considered. Due to increased reservation values for wage workers, $x^{we}_r$, none of the wage workers from the previous simulation would ever enter entrepreneurship. Therefore we consider a different population with a higher level of cash on hand.

hazard initially increases with duration. This finding has two immediate implications
for empirical analysis

**Remark 4.** *The transition pattern depends significantly on the type of the agent, whether he faces credit constraints or not, whether he faces transitions cost etc. This suggests that estimation procedures in duration analysis should incorporate lots of heterogeneity. Not only in the intercept or scale of the hazard, but also the shape.*

**Remark 5.** *Entrepreneurial hazard functions which are initially increasing followed by negative duration dependence are consistent with the presence of start-up costs or any other phenomenon, that generates an option value for the entrepreneur.*

Jørgensen (2005) provide a careful duration analysis of Danish start-ups. Using a large and comprehensive longitudinal firm database, he is able to identify all new start-up firms in 1994 and 1998 and the entrepreneur bind the firm. After carefully conditioning firm- and individual level characteristics, he find quite robust evidence that the hazard out of entrepreneurship, is initially increasing followed by a downward sloping hazard. As predicted by the model, these results are consistent with the existence of start-up cost.

## 4. NUMERICAL RESULTS - ENTREPRENEURIAL HUMAN CAPITAL MODEL

Until now we have we have treated entrepreneurial human ability as exogenously assigned to individuals and constant through time. In this section, however, I study how intertemporal incentives are altered when individuals accumulate entrepreneurial human capital. We shall consider two extreme cases: In the fist case, $h_t$ is assumed to measure pure entrepreneurial experience. Each period, the entrepreneur gains one unit of entrepreneurial human capital while wage workers gain enough human capital to precisely offset the depreciation in human capital, when $h_t = 1$, i.e. $\Delta(e) = 1$ and $\Delta(we) = 1 - \gamma$.[12] Since individuals accumulate entrepreneurial human capital only when they are entrepreneurs, entrepreneurial experience (or learning by doing) is treated as a

---

[12]This is done for numerical convenience: Since the assumption ensures that $h_t$ can never fall below 1, we can restrict attention to a compact subset of the states pace, $h_i \in [1, 1/(1 - \gamma)]$.

behaviorally determined investment decision: Some individuals may find it optimal to incur a temporary income loss in exchange for increased future returns to their business.

At the other extreme, $h_t$ measures pure work experience such that productivity is assumed only to increase during wage work, i.e. $\Delta(we) = 1$ and $\Delta(e) = 1 - \gamma$. Knowledge spill-overs are thus assumed to be more important within firms than between firms. Admittedly, this assumptions is somewhat stylized, but it captures the idea that people learn more by working with and for other people.

When human capital accumulation does not differ across occupations, i.e. when $\Delta(we) = \Delta(e)$, human capital, $h_t$, is deterministic. Thereby, $h_t$ is exogenous in the sense, that the occupational choice does not alter human capital accumulation. The implications of the model under this parametrization are not very different from the case without human capital accumulation, except that the hazard out of entrepreneurship exhibits a higher degree of negative duration dependence due to the trending productivity. When duration in entrepreneurship increases, productivity increases as well. But not due to duration in entrepreneurship, due to the course of time. Since the primary focus of this section is how saving and occupational choice is altered by human capital investment, I will not pursue the case of deterministic accumulation any further

The rest of the parameter values are chosen to be similar to the baseline parameters in the simple model with no transition costs, although with a few modifications. I let $\alpha^h = 0.06$ such that the expected increase in productivity is 6 pct. for a percentage increase in $h_t$. I let $\gamma = 0.95$, i.e. $h_t$ depreciates with 5 pct. each period. I choose $\theta = 0.80$ such that the limit of $h_t^{a_h}$ approximately equals the baseline parameter value of $\theta$ in the simple model without human capital accumulation, i.e. $\theta^{\max} = \theta \left[ max_i \left( \Delta(i) \right) / \left( 1 - \gamma \right) \right]^{\alpha_k} = 0.96$. The baseline parameter values used in the numerical simulations of the human capital model are listed in Table 4

4.1. **Learning by doing.** Consider first the scenario where individuals accumulate human capital only while they are entrepreneurs. Since entrepreneurial earnings are monotonely increasing in human capital, $h_t$, the reservation value of cash on hand, $x_r$,

TABLE 2. Baseline Parameter Values

| $r$ | $\eta$ | $\delta$ | $\theta$ | $\alpha^h$ | $\alpha^k$ | $\sigma_\varepsilon^e$ | $\lambda$ | $w$ | $\sigma_\varepsilon^{we}$ | $\gamma$ | $\Delta\left(e\right)$ | $\Delta\left(we\right)$ | $\phi^{exit}$ | $\phi^{entry}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.05 | 1.5 | 0.07 | 0.80 | 0.06 | 0.33 | 0.3 | 1 | 1 | 0.1 | 0.95 | 1 | 0 | 0 | 0 |
| | | | | | | | | | | | 0 | 1 | | |

FIGURE 11. RESERVATION VALUE OF CASH ON HAND, $x_r$
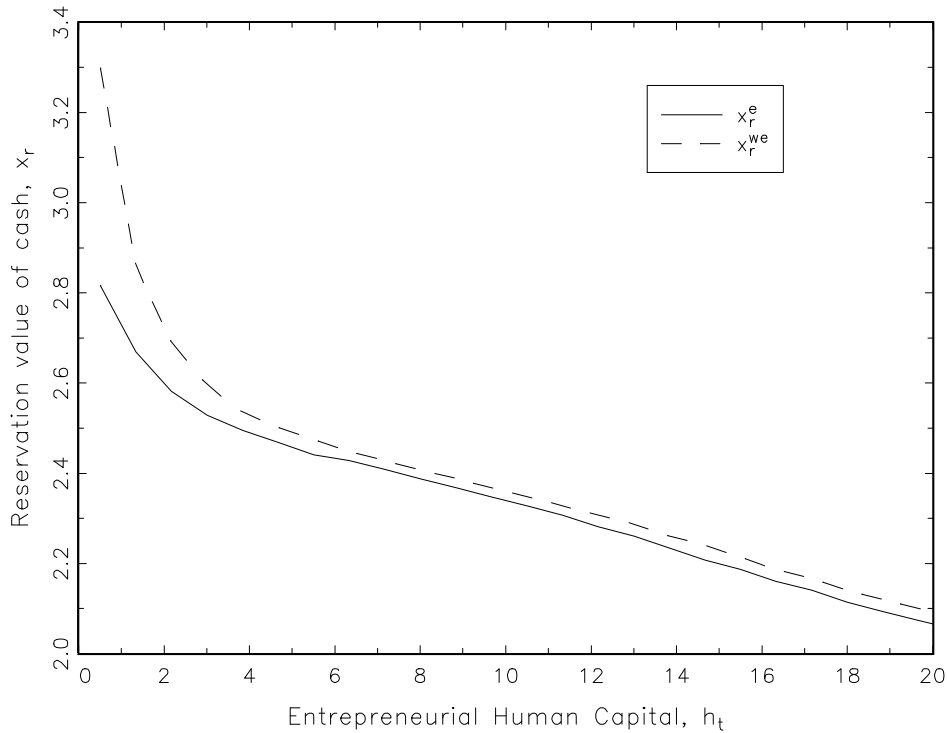


is decreasing in $h_t$. This relationship between $x_r$ and $h_t$ is displayed in Figure 11 for wage workers and entrepreneurs respectively. First note that the reservation value of cash on hand depends not only on the level of human capital, but also on the current occupation. In general, wage workers have larger reservation values of cash, since human capital in the following period is lower. One can say that the exit from entrepreneurship is associated with an indirect cost due to the forgone opportunity to accumulate human capital. As a result the occupational choice is state dependent. Due to the diminishing return to human capital, the gab between the two reservation values $x_r^{we} - x_r^e$ diminishes with human capital.

FIGURE 12. EXPECTED FUTURE GROSS INCOME - WAGE WORKERS



Wage workers
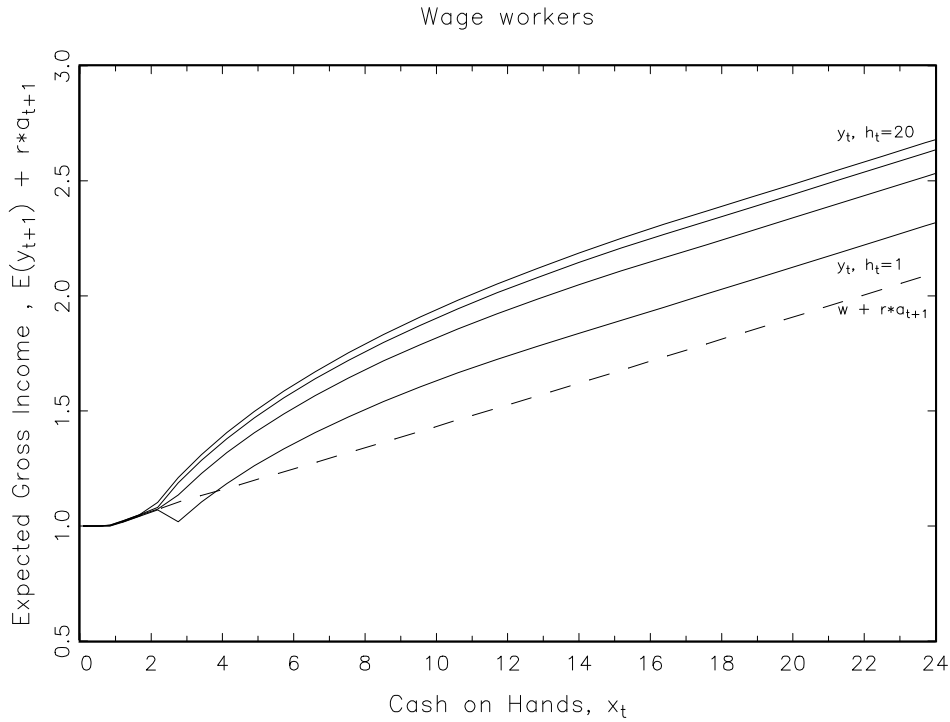
Figure 12 displays the relationship between wealth and expected future gross income for different values of human capital. Contrary to the case without human capital investments, individuals with a low level of human capital are willing to incur a temporary income loss, in exchange for an increase in expected future entrepreneurial earnings. Wage workers with a relatively low level of human capital and capital less than $x_r^{we}(h_t)$ earns the fixed wage, $w = 1$ plus interest of liquid assets holdings until he chooses to become entrepreneur at the switching point, $x_r^{we}(h_t)$.

Figure 13 displays the policy function for consumption and occupational choice for entrepreneurs and wage workers respectively. The graphs are plotted for two different level of human capital, $h_t = 1$ and $h_t = 20$. It is evident from Figure 13 that individuals with a high level of human capital have much stronger saving incentives due to the higher return to investments. Indeed, for low levels of wealth, individuals with a high level of human capital will save more out of current wealth. In contrast, individuals with a low level of human capital have relatively low returns to investments in physical capital. As

FIGURE 13. OPTIMAL CONSUMPTION

Entrepreneurs



Wage workers
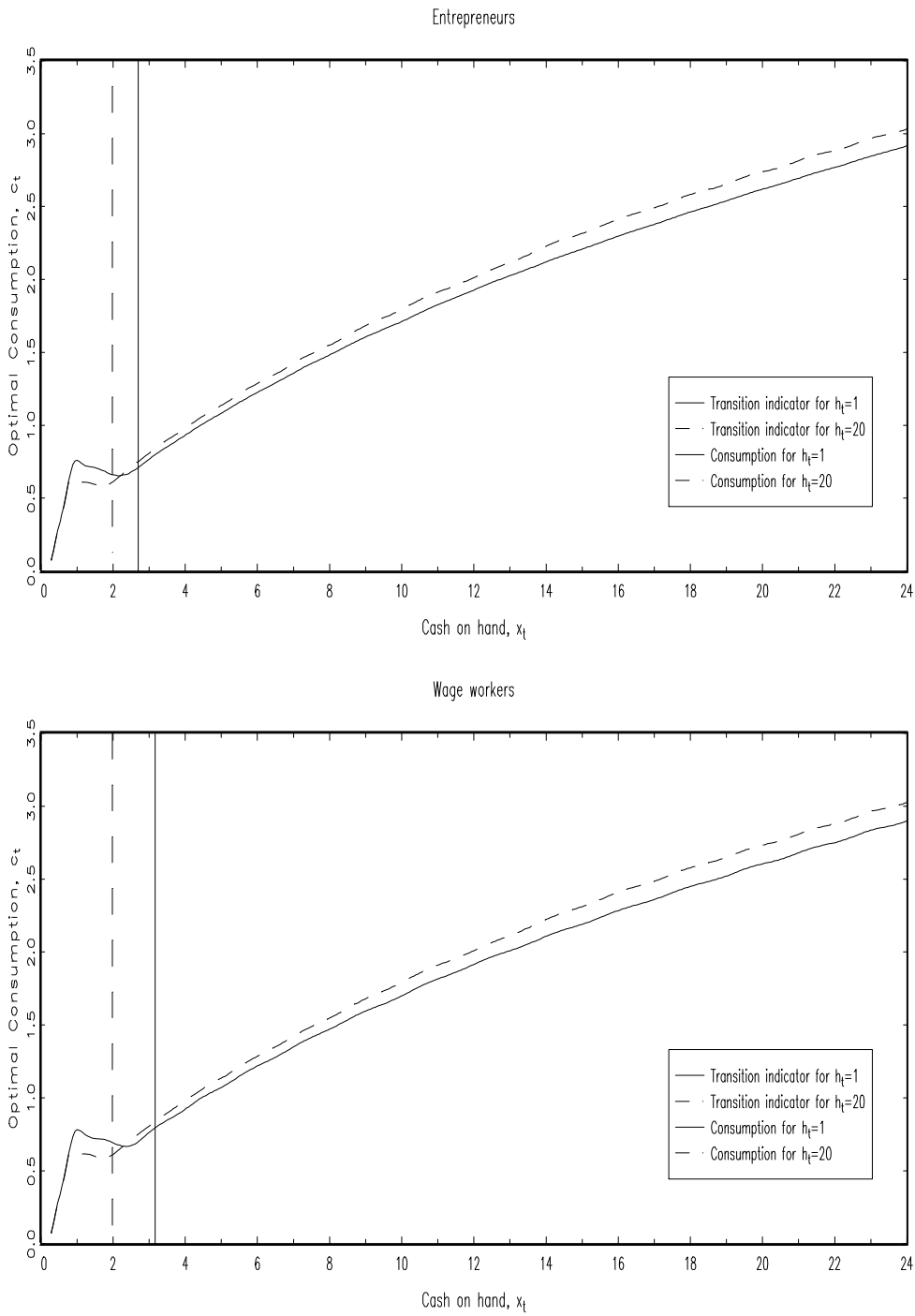


a consequence, individuals with relatively low levels of human capital have a additional incentive to entry: To gain more human capital.

FIGURE 14. SIMULATED HAZARD FUNCTION, $a_0 = 0.75$
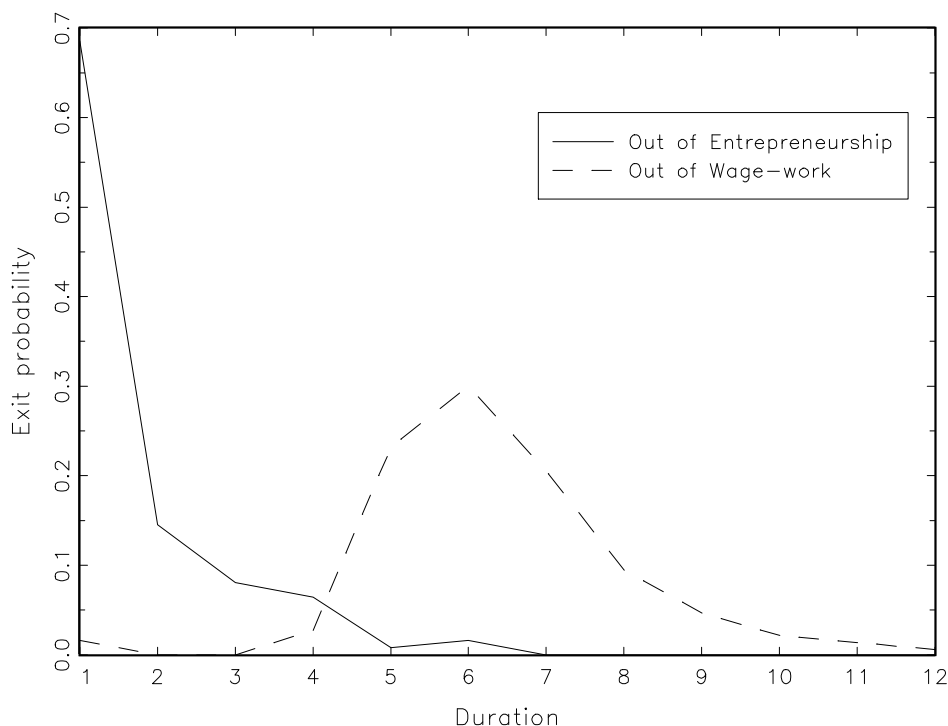


There is one important thing to note about the gab between the two reservation values $x_r^{we} - x_r^e$. The fact that individuals can only accumulate human capital in entrepreneurship creates a wedge between the two occupations: Not only is the reservation values of cash much higher for wage workers with no entrepreneurial experience, the saving incentives for constrained wage workers are weaker too. This combination reinforces significantly the impact of existing liquidity constraints on entry behavior .

In Figure 14 simulated hazard functions are displayed for entrepreneurs and wage workers respectively: The transition pattern is not qualitatively different from the case without human capital accumulation. However, the negative duration dependence for the hazard out of entrepreneurship is amplified by human capital accumulation in entrepreneurship: Since $h_t$ increases with duration in entrepreneurship, the reservation value of cash, $x_r^e(h_t)$ will decrease with duration. As a consequence, the exit probability for entrepreneurs decreases rapidly with duration. A final thing to note is that wage workers

FIGURE 15. SIMULATED SEQUENCE OF INCOME, ASSETS AND CONSUMPTION, $\alpha_0 = 0.75$
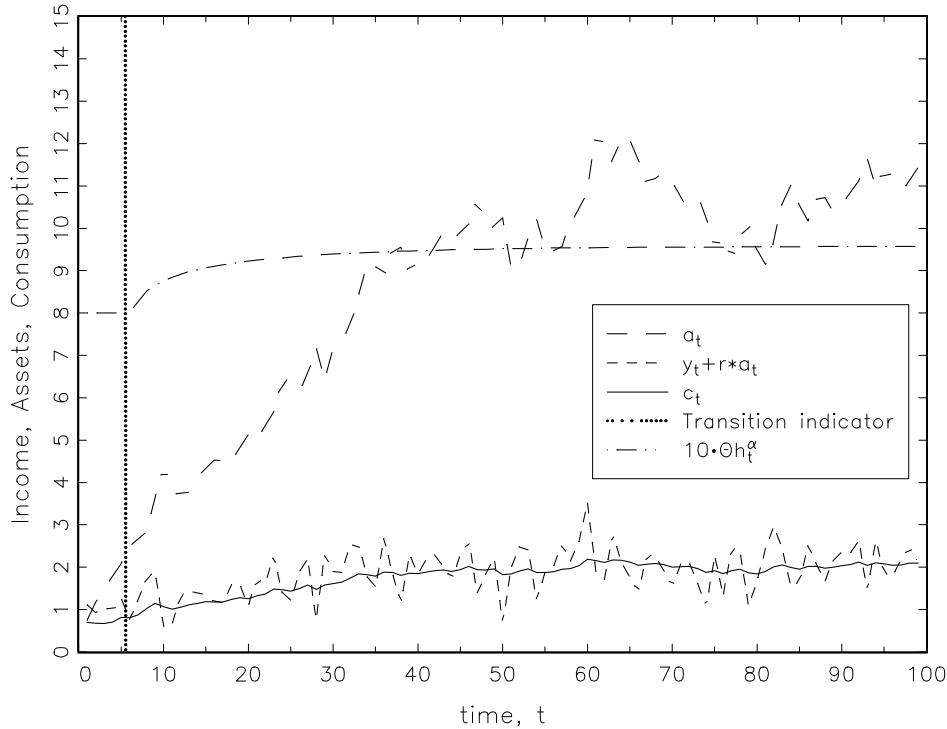


enter entrepreneurship later: On average, they will have to save for about 6-7 periods before entry.

Figure 15 graphs simulated sequences of assets, $a_t$, consumption, $c_t$, gross income, $y_t + ra_t$, and the implied value of entrepreneurial productivity multiplied by 10, $10\theta h_t^{\alpha_h}$. The simulation is done for an individual that enters the labor market with an initial level of assets of $a_0 = 0.75$.

## 4.2. Human Capital Accumulation in Wage Work.

Now turn to the opposite extreme, where human capital accumulation primarily takes place during wage work. As seen in Figure 16, this accumulation scheme holds qualitatively different implications for the implied transition pattern between the two occupations. The first thing to note is that the dependence of the current occupation is reversed relative to the picture in Figure 11. Since individuals can enhance their future entrepreneurial productivity only

FIGURE 16. RESERVATION VALUE OF CASH ON HAND
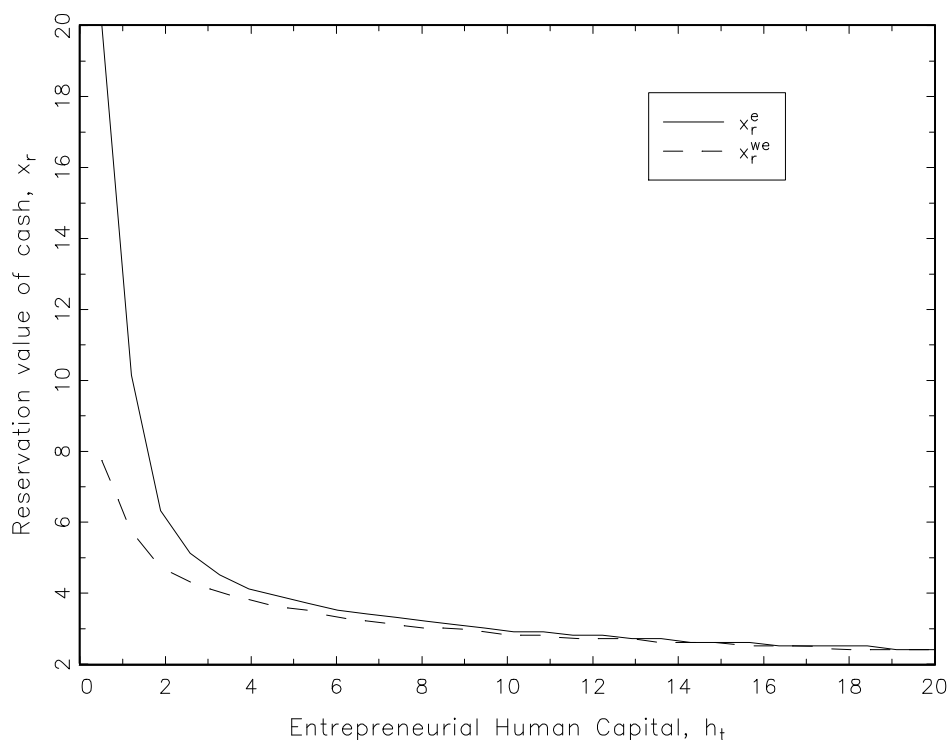


during wage work, human capital will be lower for entrepreneurs in the following period. As a result, entrepreneurs will have larger reservation values of cash on hand.

Figure 17 depicts expected future gross income for wage workers for different values of $h_t$ (the solid lines) along with expected gross income if an individual (inoptimally) chooses to remain wage worker for all values of cash on hand (the dashed line). The picture is again reversed: Individuals are willing to *stay wage workers* although expected entrepreneurial earnings are higher. Hence, to balance the return to investments in human capital and physical capital, wage workers accept a temporary income loss in exchange for an increase in expected future entrepreneurial earnings. In fact, the expected income function jumps discretely in the switching point. The size of this jump represents the amount that individuals are willing to trade for an additional unit of human capital. Due to the decreasing returns to $h_t$, this amount decreases with $h_t$.

Figure 18 displays the policy functions for consumption and occupational choice for entrepreneurs and wage workers for two different levels of human capital, $h_t = 1$ and

FIGURE 17. EXPECTED FUTURE GROSS INCOME - WAGE WORKERS
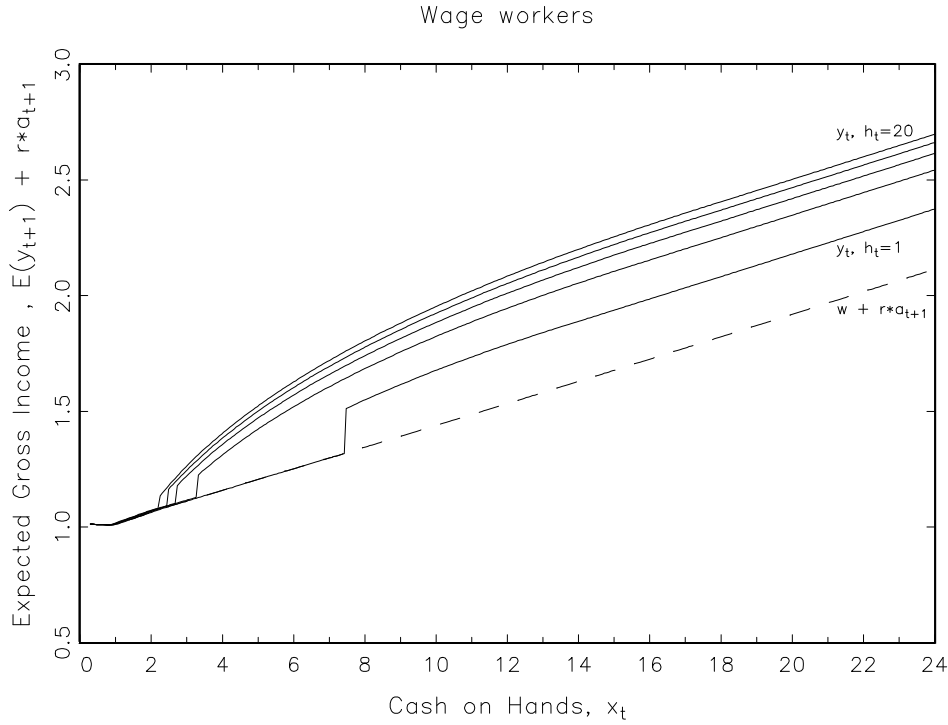
Wage workers



$h_t = 20$. The first thing to note is that entrepreneurs with $h_t = 1$, will switch to wage work regardless of their current level of wealth. Despite this transition to wage work, the entrepreneurial saving incentive remains since individuals with low levels of $h_t$ expect entry (or re-entry) to entrepreneurship after a couple of periods in wage work.

Figure 19 graphs simulated sequences of assets, $a_t$ consumption, $c_t$ gross income, $y_t + ra_t$, and the implied value of entrepreneurial productivity multiplied by 10, $10\theta h_t^{\alpha_h}$. The simulation is done for an individual that enters the labor market with an initial level of assets of $a_0 = 0.75$. Initially, the wage worker accumulates assets and human capital for a couple of periods. Hereafter, he switches to entrepreneurship to reap the benefits of his investments. At this point, the entrepreneur keeps saving to overcome binding liquidity constraints. However, since entrepreneurial human capital depreciates while access to financial assets simultaneously increases, the return to reinvestments in human capital will eventually become large enough to make the entrepreneur switch to wage work and thereby being able to accumulate more human capital. Note that the wage

FIGURE 18. OPTIMAL CONSUMPTION
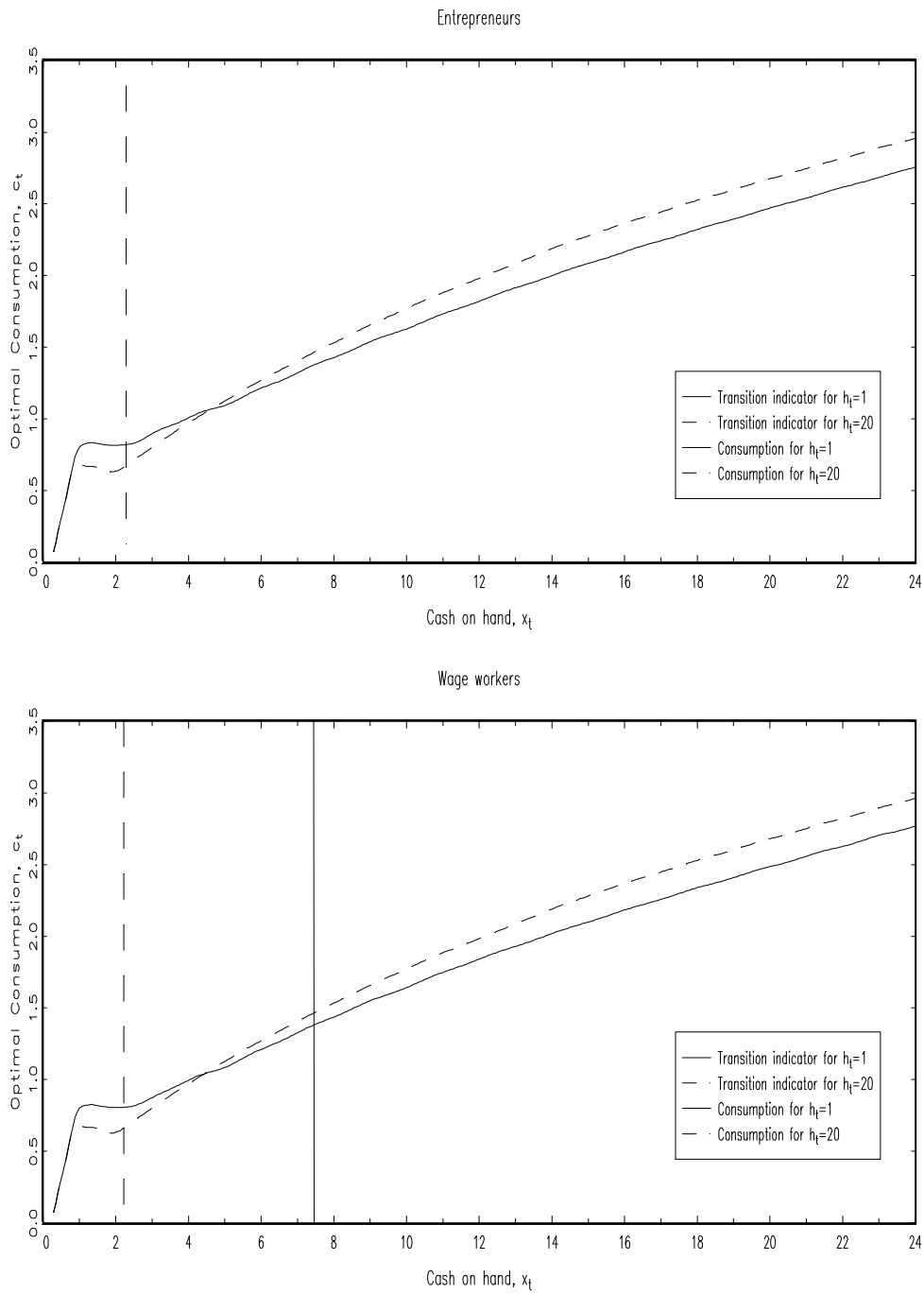


workers re-entry entrepreneurship relatively fast (after only one period in wage work).
Even more, the accumulated level of human capital is lower compared to the first entry.

FIGURE 19. SIMULATED SEQUENCE OF INCOME, ASSETS AND CONSUMPTION, $\alpha_0 = 0.75$
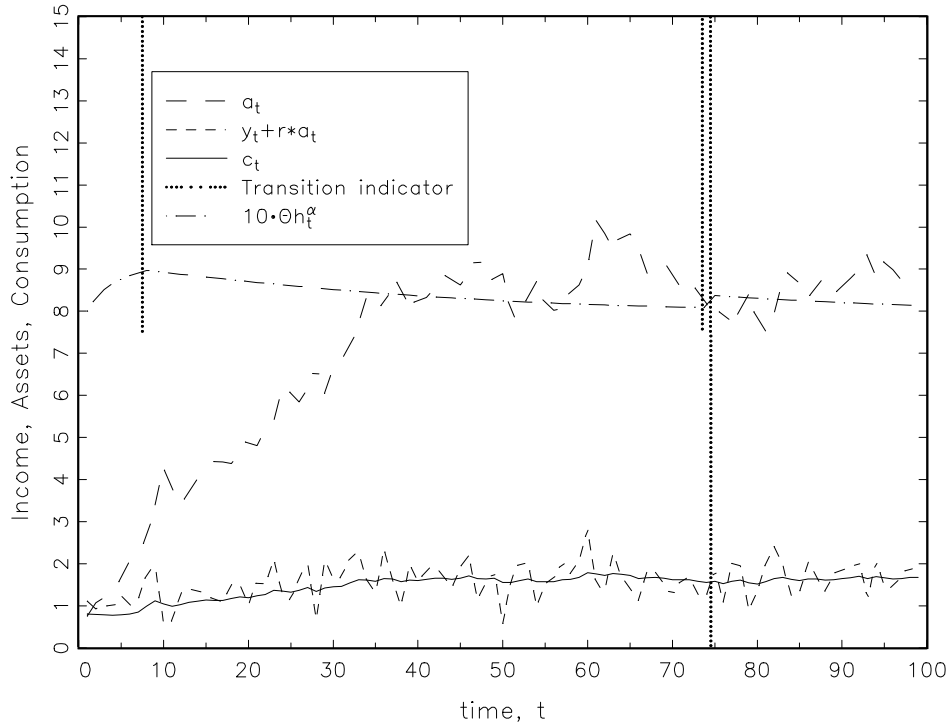


The reason is, that this agent faces a very large opportunity cost in terms of forgone entrepreneurial earnings, since accumulated assets are now much higher.

Figure 20 displays simulated hazard functions for a population that enters the labor market with an initial level of assets $a_0 = 0.75$. Consider first the hazard out of entrepreneurship. Again, we observe a decreasing hazard, although with one modification. After approximately 20 periods, the exit probability starts to increase again. At this point entrepreneurs have been exposed to human capital depreciation for several periods. Since the business is very capital intensive at this point, the return to human capital investments is high enough to induce a shift to wage work. From the hazard out of wage work, it is evident that these re-entries to wage work have a duration of only one period (the left peak). The hazard function has a second peak around 6-8 periods of wage work. The concentration here is due to the initial spell in wage work. Hence, on

FIGURE 20. SIMULATED HAZARD FUNCTION, $\alpha_0 = 0.75$
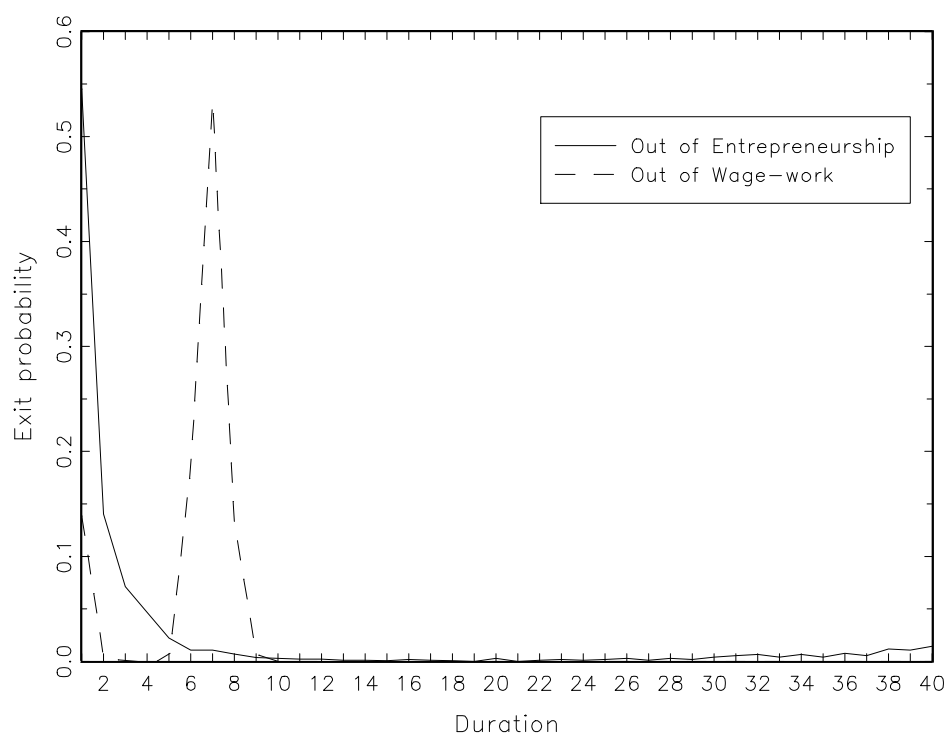


average, wage workers will need around 7 periods of wage work, before entrepreneurial activity is undertaken

If entrepreneurial human capital is primarily accumulated in entrepreneurship importance of credit constraints is amplified significantly, whereas the opposite is the case if individuals primarily acquire entrepreneurial skills in wage employment. If the former is the case, the entrepreneur will have to accept longer sequences of relatively low returns, before the business really becomes profitable. Due to the desire to smooth consumption and the lack of ability to borrow, this will in fact deter entry. If on the other hand human capital is primarily acquired as wage worker, it will be optimal to postpone entry until a sufficiently high amount of human capital is accumulated. This will be optimal even when credit constraints are not binding.

Whether human capital accumulation is in fact most pronounced in wage work or entrepreneurship is still an unresolved issue - and basically it is an empirical question. As we have seen above it has important policy implications. In particular for credit

market interventions, business start-up schemes and policies that help entrepreneurs to circumvent binding credit constraints. The answer to this question may be complicated by entrepreneurial human capital generally being unobserved and by the existence differences between e.g. industries and educational groups. However, a full estimation of the present structural model can actually be used to answer this question. This is a very intersing project, however, which I will try address in future work.

## 5. Conclusions and Directions for Future Research

To get a better understanding of the intertemporal incentives that underlie much of the behavior by entrepreneurs, I have developed an intertemporal model for saving, consumption, human capital accumulation and occupational choice in the presence of liquidity constraints, income uncertainty, and entry costs. I have done this by merging the set-up from the existing static models of entrepreneurship with the approach taken in the literature on intertemporal saving and consumption. Furthermore, I incorporate aspects from the literature on investment under uncertainty and the literature on human capital formation.

Using this model, I provide a theoretical foundation for analyzing a number of unresolved issues. I here summarize the seven key findings:

*First*, a prominent feature of the model is that it generates a well-defined transition pattern between entrepreneurship and ordinary wage work. The model predicts that workers may transit back and forth between wage employment and entrepreneurship. The latter occupation becomes more attractive as the worker accumulates sufficient wealth (and human capital) as this gives her a higher (expected) return in entrepreneurship because she can then acquire more physical capital. Wage work, on the other hand, becomes relatively more attractive when individual savings are depleted, e.g., following a series of negative shocks to entrepreneurial income which forces the individual to use her savings (i.e. sell her capital investments) to maintain consumption. As a result, the model predicts continued cycling between these occupations. This is an important feature of the model, as these sequential entry and exits is a phenomenon frequently

observed in the data. In fact, more than 25 per cent of the self-employed in Denmark have multiple completed spells of self-employment.

*Second*, the start-up costs provide an extra savings motive for wage workers in the presence of credit constraints. Wage workers who wish to (or expect to) enter entrepreneurship save to be able to finance these entry costs and to overcome the subsequent risk of losing the paid entry costs if she has to return to wage work.

*Third*, the credit constraints also induce entrepreneurs to accumulate savings and may thus explain why we empirically observe a concentration of assets among the entrepreneurs. When entrepreneurs are credit constrained, the accumulated savings determine how much physical capital they can acquire and thus the expected pay-off from entrepreneurship. Furthermore, entrepreneurs save to maintain their position as entrepreneurs to avoid potential costs associated with later re-entry.

*Fourth*, the fixed cost of entering entrepreneurship introduces a so-called option value into the model. Entrepreneurs are willing to cut consumption temporarily to keep their position as entrepreneurs, because they know that if they exit and wish to return to wage employment at a later stage, they will have to incur the sunk entry cost again. Similarly, wage workers wish to postpone investments in a business until they have enough savings to not only finance the entry cost but also to resist several subsequent negative shocks to income in entrepreneurship without existing. As a result, we will see both later entry and later exit in the presence of entry costs.

In an international comparison, we should expect Danish entrepreneurs to face relatively small entry costs, as the red tape connected with setting up a business is relatively limited in Denmark (see Fonseca, Lopez-Garzia, and Pissarides (2001)). This is also fully consistent with the relatively high entry and exit rates in Denmark.

*Fifth*, in the absence of transition costs and human capital accumulation in entrepreneurship, the probability of exiting entrepreneurship (the hazard function) will generally be declining in the elapsed duration of the entrepreneurship spell (the age of the firm).

The reason is that entrepreneurs accumulate more assets and more human capital while running a firm, making them more resistant to negative chocks. If individuals

that enter with a relatively low level of initial assets (or human capital), they will be vulnerable to negative demand chocks. Depending on the realized sequence of demand chocks, entrepreneurs will succeed in accumulating enough capital to resist future negative shocks. Hence, in addition to the primary function as working capital, assets serve as a buffer stock during downturns. Therefore, the probability of exit to wage work will decrease with duration in entrepreneurship, i.e. the hazard function is said to exhibit true negative duration dependence. This conclusion is strengthening in the presence of human capital accumulation in entrepreneurship.

In the presence of entry costs, however, the probability of exit from entrepreneurship is initially increasing in the elapsed duration and then decreasing. The entry cost alters the transition patterns between entrepreneurship and ordinary wage work fundamentally: Due to start up costs, individuals will never enter entrepreneurship if there is a significant risk that they will not be able to maintain their business in at least the following period or two. Therefore exit probability is very low for short durations. For longer durations, however, the entrepreneurs become more resistant to negative chocks due to a larger amount of accumulated capital (and possibly also human capital). When this effect dominates, hazards start to decline again.

The prediction that the hazard function is initially increasing is largely consistent with what we observe in the Danish register data on firm start-ups. Hence, the model provides a plausible explanation of observed exit behavior for Danish entrepreneurs.

*Sixth*, if entrepreneurial human capital is primarily accumulated in entrepreneurship, the consequences of credit constraints are amplified. Due to the lack of ability to borrow, individuals will have to accept long sequences of relatively low returns, before the business really becomes profitable. This will deter entry, if individuals have a desire to smooth consumption.

*Seventh*, the entrepreneurial saving motive is affected by the tightness of credit constraints in a non-monotone way: Initially, as credit constraints become less binding, potential highly productive entrepreneurs with relatively low asset holdings will find it optimal to save more. The reason is that the perspectives for (future) entrepreneurship

become better for this particular group when it is possible to borrow for investments in larger projects. As credit constraints is further loosened, and eventually become irrelevant, the incentive to save for entrepreneurial reasons disappears. This has important implications for an understanding of the effects of changes in credit policy, e.g. government loan guaranties. While such a policy would induce relatively productive individuals to increase their savings, less productive individuals will decrease their savings. Thus, such a policy would increase wealth inequality in the economy and could be associated with increased probability of default. On the other hand, relaxing the credit constraints increases the probability of entry relatively more for productive individuals - due to the increased willingness to use savings to overcome credit constraints.

The insight of the model also has important *implications for future research*: First, analysis of precautionary savings using the cross sectional correlation between income uncertainty and wealth must take into account that households may hold large proportions of wealth which is related to expectations about the decision to become entrepreneurs in the future. Hence, to mitigate a potential heterogeneity bias it is not sufficient to control for the current occupation, we must appropriately control for expectations about future occupations too. Since these expectations are generally unobserved (and time varying), such conditioning is in general very difficult. Therefore, to appropriately account for the composition of household savings, a full structural estimation of the present model can be a useful and perhaps necessary identification strategy.

Moreover, since transition patterns depend significantly on the type of the agent, i.e. whether he faces credit constraints or not, whether he faces transitions cost etc., estimation procedures in duration analysis should incorporate lots of heterogeneity; not only in the intercept or scale of the hazard, but also in the shape.

To fully understand the importance of credit constraints, entry costs and entrepreneurial risks for the importance of different saving motives it is necessary to estimate the distribution of entrepreneurial ability in the population, the significance of different

human capital accumulation schemes, returns to factors of production, preference para-meters, etc. Hence, a full structural estimation of the model is needed. This is a very interesting project that is subject to ongoing research.

However, a full estimation of the present model on micro data is a non-trivial task. For each evaluation of the likelihood function or the moments used for identification, we will have to solve a complex dynamic programming problem. Therefore, algorithms used to solve to the model must be developed further to make the estimation feasible.

In sum, the model this paper has several predictions which are consistent with ob-served behavior. Furthermore, the model adds considerably to our understanding of the effects of credit constraints on observed behavior as well as the potential consequences of government intervention in this area.

## 6. Appendix - Proof of Propositions

*Proof of Proposition 1.* Consider an individual who chooses to become entrepreneur in the following period but is constrained by capital, i.e. $i_{t+1} = e$ and $k_{t+1} = \lambda a_{t+1}$. For a marginal increase in future assets, this individual can expect to increase entrepreneurial earnings with the following amount

$$\Xi \equiv E\left[\frac{dy_{t+1}\left(i_{t+1}, h_{t+1}, a_{t+1}, \varepsilon_{t+1}\left(i_{t+1}\right)\right)}{da_{t+1}}\right] = \alpha_k \theta a_{t+1}^{\alpha_k - 1} h_{t+1}^{\alpha_h} \lambda^{\alpha_k} - r\lambda > 0$$

Note that, $\Xi$ is concave in $\lambda$ since $\alpha_k \in ]0, 1[$ and positive if the individual is credit constrained. For unconstrained entrepreneurs $\Xi = 0$.

Furthermore, $\Xi$ is increasing in $\lambda$ if $\theta h_{t+1}^{\alpha_h} \alpha_k^2 a_{t+1}^{\alpha_k - 1} \lambda^{\alpha_k - 1} > r$. For a given value of $r$, $a_{t+1}$ and $\lambda$, this condition hold for large enough $\theta h_{t+1}^{\alpha_h}$. Hence, relatively productive individuals with a relatively low asset holdings, will increase their savings when $\lambda$ increases.

On the other hand, $\Xi$ is decreasing in $\lambda$ if $\theta h_{t+1}^{\alpha_h} \alpha_k^2 a_{t+1}^{\alpha_k - 1} \lambda^{\alpha_k - 1} < r$. For a given value of $r$, $\theta h_{t+1}^{\alpha_h}$ and $a_{t+1}$, this condition holds for large enough $\lambda$. Hence, when credit constraints become less binding, the expected return to increased savings approaches zero. $\square$

*Proof of Proposition 2.* Since $h_{t+1}(h_t)$ is monotone increasing with slope $h'_{t+1}(h_t) = \gamma < 1$, then $h_{t+1}(h_t)$ will cross the 45 degree line from above in a unique fixed point $\tilde{h}$ Thereby for any $h_{high} \geq \tilde{h} < \infty$ we must have that $h_{t+1}(h_{high}) \leq h_{high}$ (since $h_{t+1}(h_{high})$ is below the 45 degree line for all $h_{high} \geq \tilde{h} < \infty$ ) $\square$

*Proof of Proposition 3.* We wish to show that there exists a $a^* < \infty$ such that if $a_t \leq a^*$ then $a_{t+1} \leq a^*$. Hence it is sufficient to prove that exists a $a^* < \infty$ such that $a_t \geq a_{t+1}$ holds for all $a_t \geq a^*$. That is,

$$(6.1) \qquad a_t \geq a_{t+1} = (1+r)\left(a_t + y_t - c_t - \phi\left(s_t, s\right)\right) \Rightarrow$$

$$c_t \geq \frac{r}{1+r} a_t + y_t - \phi\left(s_t, s_{t+1}\right)$$

should hold for all $a_t \geq a^*$.

Since $y_t$ is bounded from above by $y^{\max}$ and $\phi\left(s_t, s_{t+1}\right)$ is bounded from below by

zero, a sufficient condition for the inequality (6.1) can be formulated as

$$c \geq \frac{r}{1+r} a + y^{\max}$$

where

$$y^{\max} = \max \left\{ \begin{array}{c} w, \\ \theta \left( \lambda a_t \right)^{\alpha_k} \left( h^{high} \right)^{\alpha_h} \varepsilon^{\max} - r \lambda a_t, \\ \theta \left[ \frac{\theta \alpha_k}{r} \left( h^{high} \right)^{\alpha_h} \right]^{\frac{\alpha_k}{1-\alpha_k}} \left( h^{high} \right)^{\alpha_h} \varepsilon^{\max} - r \left( \frac{\alpha_k \theta}{r} \left( h^{high} \right)^{\alpha_h} \right)^{1/(1-\alpha_k)} \end{array} \right\}$$

is a finite since $\alpha_k < 1$ and $r > 0$

Consider for a moment the corresponding deterministic model, where individuals are endowed with initial assets $A_t$. In this case, optimal consumption can be expressed as (see Caroll (1997))

$$c_t^{det} = \left( 1 - \left( \frac{1}{1+\delta} \right)^{\frac{1}{\eta}} \left( \frac{1}{1+r} \right)^{1-\frac{1}{\eta}} \right) \left( a_t + \sum_{i=t+1}^{\infty} (1+r)^{t-i} \left( y \left( i_i, h_i, a_i \right) - \phi \left( i_i, i_{i+1} \right) \right) \right)$$

Since income is bounded from below by zero and $\phi \left( s_t, s_{t+1} \right)$ is bounded from above by $\phi$ Consumption in the stochastic model can newer fall below

$$\tilde{c}_t^{det} = \left( 1 - \left( \frac{1}{1+\delta} \right)^{\frac{1}{\eta}} \left( \frac{1}{1+r} \right)^{1-\frac{1}{\eta}} \right) \left( a_t - \frac{1}{r} \phi \right)$$

Therefore we must have that $c > \tilde{c}^{det}$ and we can therefore formulate a sufficient condition for (6.1) given by

$$\left( \left( \frac{1+r}{1+\delta} \right)^{\frac{1}{\eta}} \frac{1}{1+r} - 1 \right) \frac{1}{r} \phi > \frac{1}{1+r} \left( 1 - \left( \frac{1+r}{1+\delta} \right)^{\frac{1}{\eta}} \right) a_t + y^{\max}$$

Since the left hand side is constant and the right hand side is increasing in $a_t$ there must exist some $a^*$ such that for $a_t \geq a^*$, the inequality is a true statement. Or equally true, there must exist some $a^* < \infty$ such that such that $a_t \geq a_{t+1}$ holds for all $a_t \geq a^*$. This completes the proof $\qquad \square$

## References

BEN-PORATH, Y. (1967): "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy*, 75(4), 352–365.

BLANCHFLOWER, D. G., AND A. J. OSWALD (1998): "What Makes an Entrepreneur?," *Journal of Labor Economics*, 16, 26–60.

BLINDER, A. S., AND Y. WEISS (1976): "Human Capital and Labor Supply: A Synthesis," *Journal of Political Economy*, 84(3), 449–472.

BROCK, W. A., AND D. S. EVANS (1986): *The Economics of Small Business: Their Role and Regulation in the US Economy*. Holmes and Meier, New York.

BROWNING, M., AND A. LUSARDI (1996): "Household Saving: Micro Theories and Micro Facts," *Journal of Economic Literature*, 34(4), 1797–1855.

BUERA, F. J. (2003): "A Dynamic Model of Entrepreneurship with Borrowing Constraints," Chicago University, mimeo.

CAGETTI, M., AND M. DENARDI (2006): "Entrepreneurship, Frictions and Wealth," *Journal of Political Economy*, 144, 644–71, Federal Reserve Bank of Minneapolis Working paper no. 620.

CAROLL, C., AND A. SAMWICK (1998): "How Important is Precautionary Savings?," *Review of Economics and Statistics*, 80(3), 41–71.

CAROLL, C. D. (1997): "Buffer-Stock Savings and the Life-Cycle/Permanent Income Hypothesis," *Quarterly Journal of Economics*, 112(1), 1–55.

CRAMER, J. S., AND C. M. V. PRAAG (2001): "The Roots of Entrepreneurship and Labour Demand: Individual Ability and Low Risk Aversion," *Economia 2001, pp. 45-62*.

DE MEZA, D., AND C. SOUTHEY (1996): "The Borrower's Curse: Optimism, Finance and Entrepreneurship," *The Economic Journal, Volume 106, Issue 435 (March, 1996), pp. 375-386*.

DEATON, A. (1991): "Savings and Liquidity Constraints," *Econometrica*, 59, 1121–1248.

DENARDO, E. V. (1967): "Contraction Mappings in the Theory Underlying Dynamic Programming," *SIAM Review*, 9(2), 165–177.

DIXIT, A., AND R. S. PENDYCK (1994): *Investment under Uncertainty*. Princeton University Press, Princeton, New Jersey.

DUNN, T., AND D. HOLTZ-EAKIN (2000): "Financial Capital, Human Capital and the Transition to Self-Employment: Evidence from Intergenerational Links," *Journal of Labour Economics 18 pp. 282-305 (2000)*.

EVANS, D. S., AND B. JOVANOVIC (1989): "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints," *Journal of Political Economy*, 97, 808–827.

FAFCHAMPS, M., AND J. PENDER (1997): "Precautionary Savings, Credit Constraints, and Irreversible Investments: Theory and Evidence from Semiarid India," *Journal of Business and Economic Statistics*, 15(2), 180–194.

FLAVIN, M., AND S. NAKAGAWA (2005): "A Model of Housing in the Presence of Adjustment Costs: A Structural Interpretation of Habit Persistence," Unpublished Manuscript, University of California, San Diego.

FONSECA, R., P. LOPEZ-GARZIA, AND C. A. PISSARIDES (2001): "Entrepreneurship, Start-Up Costs and Employment," *European Economic Reveiew*, 45, 692–705.

FRIEDMAN, M. (1957): *A Theory of the Consumption Function.* . Princeton University Press.

HAMILTON, B. H. (2000): "Does Entrepreneurhip Pay? An Empirical Analysis of the Returns to Self-Employment," *Journal of Political Economy*, 108(3), 604–631.

HARRISON, G. W., M. I. LAU, AND E. E. RUTSTROM (2004): "Estimating Risk Attitudes in Denmark: A Field Experiment," *CEBR Discussion Paper*, 2004-10.

HARRISON, G. W., M. I. LAU, AND M. B. WILLIAMS (2002): "Estimating Individual Discount Rates in Denmark: A Field Experiment," *American Economic Review*, 92(5), 1606–1617.

HOLMES, T. J., AND J. A. SCHMITZ, JR (1990): "A Theory of Entreprebeurship and Its Application to the Study of Business Transfers," *Journal of Political Economy*, 98, 265–294.

HOLTZ-EAKIN, D., D. JOULFAIAN, AND H. S. ROSEN (1994): "Entrepreneurial Decisions and Liquidity Constraints," *The RAND Journal of Economics*, 25, 334–347.

HURST, E., AND A. LUSARDI (2004): "Liquidity Constraints, Household Wealth, and Entrepreneurship," *Journal of Political Economy*, 112(2), 319–347.

HURST, E., A. LUSARDI, A. KENNICKELL, AND F. TORREALBA (2005): "Precautionary Savings and Entrepreneurship," unpublished manuscript, University of Chicago.

JØRGENSEN, R. (2005): "A Duration Analysis of Danish Start-Ups," *CEBR Student Paper*, 2005-08.

KEANE, M. P., AND K. I. WOLPIN (1997): "The Career Decisions of Young Men," *Journal of Political Economy*, 105(3), 473–522.

KIHLSTROM, R. E., AND J.-J. LAFFONT (1979): "A Generel Equilibrium Entrepreneurial Theory of Firm Formation Based on Risk Aversion," *The Journal of Political Economy*, 87(4), 719–748.

KIHLSTROM, R. E., AND J. J. LAFFONT (1983): "Taxation and Risk Taking in General Equilibrium Models with Free Entry," *Journal of Public Economics*, 21, 159–181.

KIMBALL, M. S. (1990): "Precautionary Saving in the Small and in the Large," *Econometrica*, 58(1), 53–73.

MALCHOW-MØLLER, N., J. R. MARKUSEN, AND J. R. SKAKSEN (2005): "Wage Compression, Learning, and Self-Employment," *CEBR Discussion Paper*, 2005-02.

MALCHOW-MØLLER, N., AND B. J. THORSEN (2005): "Repeated Real Options: Optimal Investments and a Good Rule of Thumb," *Journal of Economic Dynamics and Control*, 29, 1025–1041.

QUADRINI, V. (2000): "Entrepreneurship, Saving and Social Mobility," *Review of Economic Dynamics*, 3, 1–40.

RUST, J. (1987): "Optimal Replacement of GMC Bus Engines: An Emprirical Model of Harold Zurcher," *Econometrica*, 55(5), 999–1033.

————— (1996): "Numerical Dynamic Programming in Economics," Chapter 14 in H. Amman, D. Kendrick and J. Rust (eds.) Handbook of Computational Economics.

STOKEY, N. L., AND R. E. LUCAS (1989): *Recursive Methods in Economic Dynamics*. Harvard University Press, Cambridge.

CENTRE FOR ECONOMIC AND BUSINESS RESEARCH (CEBR)

*E-mail address*: bsc@cebr.dk

*URL*: www.cebr.dk/bsc

# Foreign Firms, Domestic Wages

Nikolaj Malchow-Møller
*Centre for Economic and Business Research (CEBR) and*
*University of Southern Denmark*

James R. Markusen
*University of Colorado, Boulder and*
*Centre for Economic and Business Research (CEBR)*
*NBER   CEPR*

Bertel Schjerning
*Centre for Economic and Business Research (CEBR) and*
*Center for Applied Microeconometrics (CAM)*

This version, March 16, 2007

## Abstract

Foreign-owned firms are often hypothesized to generate productivity "spillovers" to the host country, but both theoretical micro-foundations and empirical evidence for this are limited. We develop a heterogeneous-firm model in which ex-ante identical workers learn from their employers in proportion to the firm's productivity. Foreign-owned firms have, on average, higher productivity in equilibrium due to entry costs, which means that low-productivity foreign firms cannot enter. Foreign firms have higher wage growth and, with some exceptions, pay higher average wages, but not when compared to similarly large domestic firms. The empirical implications of the model are tested on matched employer-employee data from Denmark. Consistent with the theory, we find considerable evidence of higher wages and wage growth in large and/or foreign-owned firms. These effects survive controlling for individual characteristics, but, as expected, are reduced significantly when controlling for unobservable firm heterogeneity. Furthermore, acquired skills in foreign-owned and large firms appear to be transferable to both subsequent wage work and self-employment.

91

1.    Introduction

The last decades have witnessed a significant increase in the amount of foreign direct investment (FDI) and the activities of multinational enterprises (MNEs). This has led to considerable academic and political interest in the role of FDI and MNEs for both source and host countries. One aspect of this concerns the potential for FDI and MNEs as important channels for productivity transfers to host countries. MNEs are hypothesized to possess superior knowledge, better production technology or better management techniques compared to the average domestic firm.

However, the microeconomic foundations of these ideas are weak, and empirical analysis is limited and indirect. In this paper, the productivity advantages of foreign firms are assumed to affect domestic workers directly. We assume that firms have different productivity levels, and that workers learn more (increase their productivity more) when they are employed by higher productivity firms. Similarly, they earn more in subsequent wage employment and self-employment if they have previously worked for a high-productivity firm.

Our model builds on Melitz's (2003) (also Helpman et al., 2004) model of industry structure with heterogeneous firms, and blends this with the learning-on-the-job models of Ethier and Markusen (1996), Markusen (2001), Fosfuri et. al. (2001), and Glass and Saggi (2002). Small numbers of domestic and foreign firms get high productivity "draws" and a potentially unlimited number of other domestic and foreign firms get low productivity draws. Foreign firms of both types face higher fixed costs of entry. All high-productivity firms can enter, but only low-productivity domestic firms can enter: they enter until profits are zero, which excludes the foreign low-productivity firms due to the latter's higher fixed costs.

The consequence of this is that foreign firms *on average* are larger and have higher productivity than domestic firms. But it also provides the hypothesis that, corrected for firm

91

size, foreign firms are not more productive than domestic firms.

Our model does not rely on externalities or on foreign firms identifying and hiring better workers. All workers are ex ante identical and earn the same present value of income over a two-period working lifetime. Skills learned in the first period when employed by a high-productivity firm are transferable to other high-productivity firms and, to a less degree, to low-productivity firms. There is thus no ex post hold-up problem as in Antrás (2003): workers in the second period of their career are paid their full productivity, but first period workers joining high-productivity firms receive a discounted wage reflecting their later higher earnings in wage work or self-employment.

The model allows us to solve for outputs, wage levels and wage growth in high and low-productivity firms, and in domestic and foreign firms. In the base case, workers joining high-productivity firms receive a higher average wage and higher wage growth over their careers, but a lower initial wage. The foreign-firm effect disappears when correcting for firm size.

Then we conduct some experiments. Increasing the productivity of a worker who transits from a high- to a low-productivity firm or increasing the probability of getting a favorable draw on suitability for self-employment *lowers* the average wage premium for workers in high-productivity firms and that premium can go *negative*. Thus the model does not trivially produce a result that workers in high-productivity or foreign firms earn more.

Imposing a minimum wage which prevents high-productivity firms from capturing rents on inexperienced workers or imposing a progressive income tax *raises* the average earnings of workers in high-productivity (larger) and foreign firms relative to those for low-productivity (smaller) and domestic firms.

A number of empirical implications concerning wages and wage growth can be derived

from the theory model. Some of these are tested on matched employer-employee data from

Denmark in the final part of the paper.

Specifically, both wage levels and wage growth are higher in foreign-owned and large

firms. Also subsequent earnings as wage worker or self-employed increase with experience

from large and foreign-owned firms. As consistent with our theory, these effects survive

controlling for both observable and unobservable worker differences, but, as expected, they

disappear when controlling for unobservable firm characteristics which proxy for the

unobservable firm type.

In summary, most of the hypotheses advanced by the simple model are verified in the

estimations. One difference is that empirically, wage *levels* and self-employment earnings in

foreign firms are still greater than in domestic firms when controlling for firm size, though the

difference is greatly reduced compared to estimates that do not control for firm size.

Consistent with the simple model, however, wage *growth* in foreign firms is not greater when

controlling for firm size. We comment on this residual positive effect of foreign firms in the

concluding section.


2.      Some Relevant Literature

Empirically, it is a well-established fact that foreign-owned firms (or MNEs more

generally) pay higher wages on average than domestically-owned firms. Existing studies can

be grouped under two headings: (i) studies based on firm-level data, as in, e.g., Feliciano and

Lipsey (1999) and Aitken et al. (1996); and (ii) studies based on matched employer-employee

data, as in Martins (2004) and Heyman et al. (2004). The advantage of using matched

employer-employee data is that it can be explicitly analyzed whether part of the wage

differential is due to individual differences among the employees.

While a number of studies have shown that part of the overall "wage-gap" between foreign-owned and domestically-owned firms can be attributed to a higher average quality of workers in foreign-owned firms, a considerable part can only be explained by different firm characteristics than the average domestically-owned firm. Hence, the existing evidence points to a productivity advantage in foreign-owned firms which is somehow transformed into higher wages of the employees; see Lipsey (2002) for a recent review.

A number of studies have also analyzed how these productivity and wage advantages have influenced the productivity and/or wages of other firms, see, e.g., Haddad and Harrison (1993), Haskel et al. (2002), Almeida (2003), and Javorcik (2004) (see Keller 2004 for a more general approach). While we consider the productivity transfers that occur from foreign-owned firms to domestic firms via worker mobility and entrepreneurship, the empirical literature has to a large extent concentrated on wage and productivity spillovers (and transfers) *between* firms; see Lipsey (2002) for a review. While a positive effect has been found in the case of firm-to-market spillovers (higher average wages), the evidence is more mixed when it comes to firm-to-firm spillovers. However, studies of productivity spillovers between plants within industries have generally found positive effects of foreign-owned companies; see Lipsey (2002).

Only few empirical studies have analyzed productivity transfers via worker mobility; see Martins (2005) and Görg and Strobel (2005), where the latter, using data from Ghana, considers transfers via worker mobility to self-employment. We return to these studies in the empirical part of the paper.

Similarly, very few studies have tried to provide a theoretical foundation for such productivity transfers. Glass and Saggi (2002) thus build a model where workers employed by MNEs immediately get access to their superior technology. Hence, MNEs must pay a wage

premium to prevent workers from moving to other companies bringing along information about this technology. In Fosfuri et al. (2001), Ethier and Markusen (1996), and Markusen (2001) on the other hand, workers only get access to the superior technology following a period of training by the MNE. Hence, workers are not immediately paid a higher wage in MNEs. In both types of models, however, productivity transfers arise when workers employed (and trained) by MNEs move to domestic firms. Markusen and Trofimenko (2006) provides a more explicit model of skill transfer from foreign experts to domestic workers. Specifically, they assume that working with foreign experts is an alternative to studying as a means of obtaining skills.

As a final point, we should mention Yeaple (2005) who provides an alternative to the Meltiz framework that we borrow here. Yeaple assumes that firms are ex ante identical while workers are not (both opposite to the present paper) and that there are alternative technologies to choose from. In general equilibrium, some firms choose technologies that make them larger and they pay higher wages because they hire more skilled workers. These larger firms are also the exporters (easy generalized to establishing subsidiaries). It strikes us that this alternative approach generates at least some predictions close to ours, and clearly deserves empirical investigation.

## 3.    A model of entry, productivity, and industry structure

A principal objective of this theory section is to develop a plausible model that, at least in some circumstances, generates (1) higher average earnings in larger and/or foreign-owned firms, (2) a steeper earnings profile for the average worker in larger or foreign-owned firms, and (c) wage workers and newly self-employed workers (entrepreneurs) earn more, on average, if they previously worked in a large and/or foreign-owned firm. But we want to

generate these results while assuming that (a) all workers are ex ante identical (foreign firms are not merely selecting the best workers) and (b) foreign firms are not arbitrarily more productive than domestic firms. The model will draw heavily on the contribution of Melitz's (2003) model of industry structure with heterogeneous firms with monopolistic competition. This is combined with a learn-on-the-job model of Markusen (2001).

We are attempting to keep the model relatively simple, and so will make a number of restrictive assumptions.

(1) There are two types of domestic and foreign firms: high-productivity (HP) firms and low- or moderate-productivity (MP) firms that produce differentiated goods, denoted X. Foreign firms face an added fixed cost of entering a foreign market with a subsidiary.[1]

(2) An unlimited number of domestic and foreign firms take productivity draws. A small number in each country draw high productivity, the rest all draw moderate productivity. Note that this avoids a more ad hoc assumption that foreign firms are inherently better.

(3) The number of high-productivity firms is sufficiently small and/or the domestic market is sufficiently large, such that all high-productivity foreign and domestic firms can enter the domestic market.

(4) The "residual" demand is then satisfied by a limited number of moderate-productivity domestic firms entering up to the point where a zero-profit condition holds for moderate-productivity domestic firms. Foreign MP firms cannot enter in competition with the domestic MP firms: the former face a higher fixed cost.

(7) The model is quasi-dynamic. Firms are long lived, but fixed costs are per period,

---

[1]We originally called the MP firms LP for low productivity. But a lower-case L is confused with the number 1 in the notation, so we switched. The MP terminology is consistent with the Lake Woebegone principle that "all children are above average".

and demand is stationary. There are no investment or borrowing decision or any other intertemporal features except that MP firms decide whether or not to enter in a given period. Thus, we can analyse a single period in this "steady-state" environment.

(8) The model also has a quasi-overlapping-generations feature. Each worker has a two-period career, and all workers begin their careers as identical inexperienced workers. Workers who join MP firms do not improve their productivity over time while workers who join HP firms have both higher productivity in their first period and learning results in an even higher productivity in the second period of their career. Skills are assumed not to be firm-specific, so experienced workers are priced in a competitive market, and their wage path is such that new workers are indifferent between joining MP and HP firms.

(9) We allow workers to transit from an HP to an MP firm (the opposite transition possibility did not seem to add anything interesting so we dropped it). These transiting workers have a lower productivity than if they stay in the HP firm but a higher productivity than new workers or workers with one period in an MP firm (who do not learn).

(10) In the second period of their careers, workers take a draw which determines whether they will be good or bad as self-employed entrepreneurs in period 2. Among workers who get favorable draws, those who worked in HP firms will have a higher productivity than those who worked in MP firms.

(11) Finally, the model is largely partial equilibrium. There is an unlimited supply of new workers available at a fixed wage, and a given worker disappears after two periods. Expenditure on X goods is fixed, and those who go to self-employment disappear off to another industry. Both the exogenous number of HP and the endogenous number of MP firms hire experienced and inexperienced workers in a competitive market. The steady-state or stationarity assumption is that the number of experienced workers available is equal to the

number of inexperienced workers hired by HP firms.

Our notation is as follows.

$r_i^h$     -     labor productivity (in physical units of X output) in HP firms, where i = 1 is an inexperienced worker and i = 2 is an experienced worker. Workers in MP firms do not learn and their productivity in both periods is normalized to $r_i^m = 1$.

$r_t^m$          a worker with one-period of experience in an HP firm can transit to an MP firm, with $r_t^m$ denoting that worker's productivity. We assume that $1 < r_t^m < r_2^h$ . In other words, a worker transiting from an HP firm to an LP firm carries only part of the HP firm's productivity advantage with him/her. This will be a variable and discussed more below.

$w_i^h$     -     wage of an inexperienced worker (i = 1) and an experienced worker (i = 2) in an HP firm. If there are transiting workers, they are indifferent in equilibrium to transiting and so a worker employed by an HP firm in period 1 earns $w_2^h$ in period 2 regardless of whether the worker is in an HP firm or transits to an MP firm.

$n_d^h, n_f^h$ -     number of HP firms of domestic (d) and foreign (f) origin respectively. These are *constants* (all existing HP firms can enter).

$n^m$     -     number of MP firms, determined by free entry. This is a *variable*.

$p^h$     -     price of a representative differentiated good produced by an HP firm.

$p^m$     -     price of a representative good produced by an MP firm.

$X_1^h, X_2^h$     outputs of an HP firm produced by inexperienced and experienced workers, respectively.

$X^m, X_t^m$     outputs of an MP firm produced by (first or second period) inexperienced workers and produced by transit workers from HP firms, respectively.

$\alpha$     -     the share of workers who, at the beginning of period 2 of their career, learn that they have a higher productivity as self-employed

$v$     -     multiplier on the wage of an experienced worker that gives self-employment earings in period 2 for workers who get a favourable draw on self-employment

productivity (e.g., self-employment earnings are $w_2^h * v$ for a worker from an HP firm).

$\delta$ - the discount factor, $0 < \delta = 1/(1+r) < 1$, where r is some rate of interest/discount

Consumers have Dixit-Stiglitz preferences over an endogenous number of differentiated goods, and spend a fixed amount of income I on X sector goods. $\sigma$ denotes the elasticity of substitution between varieties. Each period's demands do not depend on prices in the other period. Demand for good i (k) is given by

$$X_i = p_i^{-\sigma}\left[\sum_k p_k^{1-\sigma}\right]^{-1} I \qquad (1)$$

Under the so-called "large-group" assumption, individual firms are assumed to be too small to influence the price index term in square brackets, and hence each firm's perceived elasticity of demand is just $\sigma$ and the optimal markup is $1/\sigma$.

The equilibrium output of each high-productivity firm, whether foreign or domestic, is determined by marginal revenue product equal to the wage. Outputs by experienced and inexperienced workers are identical (homogeneous), but these worker types differ by productivity. There are two first-order conditions for output from inexperienced workers ($X_1^h$) and for output from experienced workers ($X_2^h$). We adopt a complementarity representation of our model in which all equations are written as weak inequalities each with an associated non-negative complementary variable. The pricing inequalities for output from inexperienced and experienced workers followed by associated complementary variables are given by

$$p^h (1 - 1/\sigma)r_1^h \leq w_1^h \qquad\qquad X_1^h \qquad\qquad (2)$$

$$p^h (1 - 1/\sigma)r_2^h \leq w_2^h \qquad\qquad X_2^h \qquad\qquad\qquad (3)$$

MP firms can hire inexperienced workers, including those who have already been employed by an MP firm for one period (no productivity increase) and transit workers from HP firms. The latter must be paid the wage $w_2^h$.[2] Similar to equations (2) and (3), the two pricing equations and complementary quantity variables are

$$p^m (1 - 1/\sigma) \leq 1 \qquad\qquad X^m \qquad\qquad\qquad (4)$$

$$p^m (1 - 1/\sigma)r_t^m \leq w_2^h \qquad\qquad X_t^m \qquad\qquad\qquad (5)$$

Assume that the fixed costs for domestic MP firms require $F^m$ number of inexperienced workers at wage = 1 or $F^m/r_t^m$ number of transit workers at wage $w_2^h$. When there are transit workers in equilibrium, (4) and (5) imply that $w_2^h/r_t^m = 1$ and so fixed costs for MP firms are always given by $F^m$ regardless of whether or not there are transit workers, and the firm is indifferent between the two types. Given this indifference, we assume that different worker types are used in proportion to their overall contributions to the output of the firm.

Fixed costs for domestic and foreign HP firms require $F_d^h/r_1^h$ and $F_f^h/r_1^h$ units of inexperienced workers, respectively, at a wage of $w_1^h$, or $F_d^h/r_2^h$ and $F_f^h/r_2^h$ units of

---

[2]The reason that MP firms may employ transit workers at wage $w_2^h$ even though they have lower productivity than when continuing to work in HP firms is that the latter are larger and hence have lower equilibrium prices $(p^h < p^m)$.

experienced workers at a wage of $w_2^h$. By virtue of (2) and (3), the firm is indifferent between

using inexperienced and experienced workers and, given this indifference, we assume that the

two types are used in proportion to their overall contributions to output of the firm.

Furthermore, we assume that $F_f^h > F_d^h$.

Turning now to wages of HP workers, the wage of second period experienced workers

will be determined by a supply = demand relationship. Stationarity requires that the number of

workers used in HP firms in the first period of their careers, minus losses to self-employment,

equal the demand for experienced (second period) HP workers by HP firms and by MP firms

hiring transit workers. To incorporate our assumptions about fixed costs, let

$$s_t^m = X_t^m/(X_t^m + X^m) \tag{6}$$

be the proportion of output by transit workers in the total employment of MP firms. Similarly,

let

$$s_1^h = X_1^h/(X_1^h + X_2^h) \qquad\qquad s_2^h = X_2^h/(X_1^h + X_2^h) \tag{7}$$

be the shares (in output) of inexperienced and experienced workers in HP firms. The
stationarity relationship determining $w_2^h$ is

$$
\begin{aligned}
&w_2^h \\
&(1-\alpha)[n_d^h(s_1^h F_d^h/r_1^h + X_1^h/r_1^h) + n_f^h(s_1^h F_f^h/r_1^h + X_1^h/r_1^h)] \ \geq
\end{aligned}
\tag{8}
$$

$$n_d^h(s_2^h F_d^h/r_2^h + X_2^h/r_2^h) + n_f^h(s_2^h F_f^h/r_2^h + X_2^h/r_2^h) + n^m(s_t^m F^m/r_t^m + X_t^m/r_t^m)$$

where the left-hand side is the use of inexperienced workers by HP firms minus losses to self-

employment and hence "supply" of second-period experienced workers. The right-hand side

is demand for second-period experienced workers by HP and MP firms.

Working backwards, the first-period wage for workers hired by HP firms will be given

by a condition that the entering worker is indifferent over his or her two-period career to being

hired by an HP firm or an MP firm. Note that if $w_2^h > 1$, this in turn implies that the first-

period HP workers will accept a wage $w_1^h < 1 = w^m$. The resulting indifference condition

takes into account the expected value of a good self-employment draw (probability $\alpha$):

$$w_1^h \geq 1 - \delta(1-\alpha)(w_2^h - 1) - \delta\alpha(w_2^h v - v) \qquad\qquad w_1^h \qquad (9)$$

As noted earlier, we assume free entry and exit of MP firms. This gives us a zero-profit

condition, where the complementary variable is the number of firms active in equilibrium.

$$p^m X^m + p^m X_t^m \leq F^m + X^m + w_2^h X_t^m / r_t^m = F^m + X^m + X_t^m \qquad n^m \qquad (10)$$

Finally, there are supply-demand equations for X output with complementary variables

being the X prices. Because of symmetry within firm types, we can reduce (1) to the supply-

demand equalities for a representative good for each firm type. These two equations have

prices as complementary variables:

$$X_1^h + X_2^h = (p^h)^{-\sigma}\left[(n_d^h + n_f^h)(p^h)^{1-\sigma} + n^m(p^m)^{1-\sigma}\right]^{-1} I \qquad\qquad p^h \qquad (11)$$

$$X^m + X_t^m = (p^m)^{-\sigma}\left[(n_d^h + n_f^h)(p^h)^{1-\sigma} + n^m(p^m)^{1-\sigma}\right]^{-1} I \qquad\qquad p^m \qquad (12)$$

Finally, consider the productivity of a transit worker, $r_t^m$. If this is fixed, the model

has a bang-bang property with respect to the productivity of transiting workers (at some

critical value all experienced workers go to MP firms and HP firms employ only inexperienced

workers). In order to smooth this, we assume that $r_t^m$ is a decreasing function of the share of

transit workers in the workforce of MP firms. The idea is that the first HP worker hired has a

big effect on productivity, but subsequent workers are less able to exploit their skills in low-tech production. Our final equation thus gives the productivity of a transit worker as

$$r_t^m = \gamma + \rho(1 - s_t^m) \qquad (\gamma + \rho) > 1 \qquad r_t^m \qquad (13)$$

$\gamma$ is then the minimum productivity of a transit worker, taken on if all workers in MP firms are transit workers (this never happens in our simulations). $(\gamma + \rho)$ is the maximum productivity, attained for the first transit worker employed, assumed strictly greater than one as noted earlier, the latter being the productivity of an inexperienced worker.

Our model given by (2)-(13) thus constitutes thirteen non-linear inequalities (there are two equations in (7)) in thirteen non-negative variables. We solve this model analytically in Appendix A to this paper. But in what follows from this point, we will just report some simulation results using the non-linear complementarity solver in GAMS, working directly with (2) - (13).

Before looking at some numerical outcomes, let us quickly summarize some general results. We do not think that these depend on the specific parameter values chosen (except of course they do depend on the inequality assumptions among parameters), but we can have no general proof in this regard. Most of these are shown analytically in the appendix.

(1) HP firms are bigger than MP firms in equilibrium both in terms of physical output and in value terms; the value difference is smaller, since the HP firm's higher output commands a lower price in equilibrium. Productivity differences are amplified in output and value difference, so a 75% productivity advantage generates about 350% more output and about 175% more value (revenue) depending on other parameters.

(2) HP firms pay a lower wage in the first period and a higher wage in the second period relative to MP firms. Obviously, the wage profile over time is steeper in an HP firm

than in an MP firm.  Higher wages paid to experienced workers in HP firms are not due to selection, but to higher learning within the HP firm.

(3) In our base cases, HP firms pay a higher average wage to its work force than MP firms, due to discounting.   (In our numerical solutions, we use a high discount rate motivated by the view that one-period of the worker's career may be at least ten years.)  However, this is not a general result, and it can be reversed by high equilibrium transit rates and high self-employment opportunities.  These alter the composition of experienced and inexperienced workers in HP firms and so alter the average wage.

(4) Foreign firms will be observed to pay a higher average wage than domestic firms. However, this is due to the composition of the two groups, with domestic firms' wages being an average of those in low and high-productivity firms.

(5) Combining this with finding (1), it follows that, corrected for firm size, foreign firms do not pay experienced workers more than (large) domestic firms.  Similarly, the higher wage to experienced workers in large firms is not due to selection, but to the fact that firm size is just a reflection of productivity.

(6) It follows directly from the assumption that self-employment earnings are greater for a worker previously employed by an HP firm than by an MP firm that self-employed workers with a background in larger or (uncorrected for size) foreign firms earn more.

We now turn to some simulations, first presenting a "base" case, in which there is no self-employment and no transiting workers.  Key parameter values are as follows:

Self-employment probability: $\quad\quad\quad \alpha = 0.0$
Discount factor: $\quad\quad\quad\quad\quad\quad\quad\quad \delta = 0.5$
Elasticity of substitution among varieties: $\quad \sigma = 3.0$
Income level (picked to give $n^m = 50$): $\quad\quad I = 116.6667$
Fixed costs: $\quad\quad\quad\quad\quad\quad\quad\quad\quad F^m = F_d^h = 0.5 \quad\quad F_f^h = 0.6$
Exogenous number of HP firms: $\quad\quad\quad n_d^h = n_f^h = 5$

Productivity multiplier in self-employment: $\quad v \; = \; 1.2$

Normalized wages in MP firms: $\qquad\qquad w_1^{\,m} \; = \; w_2^{\,m} \; = \; 1$

Productivities: $\qquad\qquad\qquad\qquad\quad r_1^{\,h} \; = \; 1.5 \qquad\qquad r_2^{\,h} \; = \; 2.0$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad r_1^{\,m} \; = \; 1.0 \qquad\qquad r_2^{\,m} \; = \; 1.0$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \gamma \; = \; 0.94 \qquad\qquad \rho \; = \; 0.25;$ [3]

Tables 3.1-3.4 present results. The first column is identical in all simulations, and this is our benchmark solution. HP firms have a higher average wage and higher wage growth (calibrated to zero growth for MP firms). The higher average wage is due entirely to discounting in this case. The output of an HP firm is 4.63 times the output of an MP firm. Income was picked so that the benchmark number of MP firms is 50.

Table 3.1 presents simulations that gradually increase $\gamma$ with the initial (endogenous value) of $r_t^{\,m}$ given by 1.19 from (13), where $\rho$ is set at 0.25. Raising $\gamma$ makes it more attractive for MP firms to hire workers who have spent one period in HP firms. Fixed values of MP wages, income and so forth give the model a critical value of $r_t^{\,m} = 1.2$ at which workers start to "defect" to MP firms. They would all jump at a higher value, and hence our formulation in (13) above "smooths" this. As indicated in the top two lines of Table 3.1, increases in $\gamma$ are just offset by a falling share of continuing MP workers (rising share of transiting workers) and so $r_t^{\,m}$ stays constant at 1.2 until all experienced HP workers are hired away (we don't run the values out that far in the table, but it can happen).

Table 3.1 shows that the increased productivity of transit workers does not affect wages in HP firms or wage growth of a given worker. However, it does affect the *average*

---

[3] $r_t^{\,m} \; = \; 1.19 = 0.94 + 0.25$ is then the initial calibrated value of the transit productivity.

*wage* paid by HP firms, because there are a lower and lower proportion of experienced workers in these firms. The bottom row of the Table indicates the share of first-period HP workers who transit to MP firms. At some point, the average wage paid by foreign firms falls *below* the wage (=1) paid by domestic firms. HP firms become "nurseries" where inexperienced workers learn skills that they take to MP firms. These results emphasize that the model by no means trivially produces an outcome in which average wages are higher in HP or foreign firms versus domestic firms, but wage growth and wages corrected for experience continue to be higher in HP firms (and hence the average foreign firm).

Table 3.2 conducts an experiment in which the probability of a favorable draw on self-employment, $\alpha$, is increased, starting at the benchmark value of 0. While $\alpha$ is the same for workers who chose HP and MP firms in the first period of their careers, the HP workers get a bigger absolute bonus (bonus is a proportion of what would have been their second-period wage). In equilibrium, this forces down the wages in HP firms proportionately in both periods of a worker's career: higher expected self-employment earnings reduce the wages needed to make workers indifferent to joining HP firms. While wage growth of HP workers is unaffected, the average wage in HP firms decreases and can fall below that in MP firms, and this is also true in this case corrected for experience as just noted. Similar comments then apply to foreign versus domestic firms.

Because of this fall in HP firm wages, MP firms can hire away workers. For the parameters we use, this begins to happen at $\alpha = 0.06$ in Table 3.2.[4] Employment in HP firms shifts toward inexperienced workers, but total output per firm rises due to the lower wages.

Table 3.3 considers a minimum wage. We begin the simulations with this

---

[4]In order to get this to happen, we had to use a very large value of $v = 6$; so those that do get a favorable draw are very well off indeed.

unconstrained and equal to the free-market wage of 0.90 for HP workers in their first period, so these two are of course the same solution. Then we gradually raise the minimum until it hits $w = 1$ in the right-hand column, the wage of inexperienced workers available to the industry. In addition to having the obvious effect of raising the wages of inexperienced workers in HP firms, it cuts output by HP firms (price rises), and so has the equilibrium effect of raising the wages of experienced workers in HP firms as well. Wage growth is in fact unaffected, but the average wage paid in HP firms is considerably higher in the right-hand column of Table 3.3. HP workers collect pure rents. There is no transit of workers to MP firms, which cannot afford these pricey experienced workers. Thus, minimum wages is a reason to expect to observe both higher wages and higher wage growth in high-productivity firms. The HP firms remain large, but less so.

Table 3.4 imposes a progressive income tax. We keep this very simple by assuming that the tax rate on wages less than or equal to one is zero, and that there is a constant tax rate t on wages in excess of one (the so-called "flat tax"). Table 3.4 shows that this acts somewhat like the minimum wage, but by making experienced rather than inexperienced labor more expensive for HP firms. The resulting fall in demand pushes up (before tax) wages for both experienced and inexperienced workers, although the growth rate remains the same. The average wage in HP firms is significantly higher than in MP firms and hence similarly much higher in foreign than in domestic firms.

However, care must be taken in the presence of income or payroll taxes. Results depend very much on which wage is reported in the data: the producer (before tax) cost or the household (take-home) wage. If it is the producer cost, then the income tax increases the average wage paid in HP firms relative to MP firms as we have just indicated. If it is the consumer (take home) wage that is measured, the difference in the average wage between HP

and MP firms is smaller. The growth in take-home wage is reduced by the tax. The growth rate in the take-home wage is reported in the fourth row of Table 3.4, and so we see that the profile of the take-home wage is flatter (higher initial wage, lower take-home wage) than in the base case.

We suspect that data is generally reporting producer cost of labor, or gross wage before tax, and hence here we have another reason why the average wage and wage growth is higher in HP firms and hence higher in foreign than in domestic firms. In the following sections, we use gross wages before taxes.

4.      Data and Empirical Strategy

In this and the following section, we confront the empirical predictions of our theoretical model with the real world using matched employer-employee data from Denmark. The data come from the Integrated Data Base (IDA) for Labor Market Research compiled by Statistics Denmark, combined with firm level information about foreign ownership, size, turnover, and exports. IDA contains register based annual data since 1980 on all individuals with Danish residence. It provides detailed information on individual background variables such as education and family characteristics as well as detailed records of previous labor market performance, including occupations and income.

All workers are linked to workplaces (plants) which in turn can be linked to firm level information, which, *e.g.*, allows us to identify all employees in foreign-owned firms in Denmark. Information about foreign ownership is currently available only for the years 2000-2002.[5] As a consequence, in the regressions including foreign ownership, we have to rely on a

---

[5] A firm is classified as foreign owned by Statistics Denmark if foreigners ultimately own more than 50% of the firm, and the foreign direct investment amounts to more than DKK 10 million.

panel for the years 2000 to 2002, while for the regressions without foreign ownership, we can extend the panel to the period 1981-2003. Note that information about occupation in a given year is based on the individual's occupation in the last week of November. Hence, we cannot observe worker flows within a given year.

Our theoretical model implies that the firm type (HP or MP) affects (a) the current wage level; (b) the current wage growth; and (c) the wage level in future occupations, since acquired skills are, at least partly, transferable. However, as we cannot directly observe whether a firm is an HP or an MP firm, we use the following two additional implications of our theory to derive our testable hypotheses: (d) All foreign-owned firms are of type HP; and (e) HP firms are larger than MP firms. Combining these five predictions results in the following three sets of empirical hypotheses:

First, with respect to the relationship between *the current firm type and the wage level*, our model predicts that in a cross section, we should observe higher wages for workers in foreign-owned and/or large firms, as these firms should all be HP firms. Furthermore, the effect of foreign-ownership should disappear (or at least be reduced) when controlling for firm size, as all HP firms are large but not necessarily foreign-owned. In other words, size should be a better proxy for HP than foreign ownership.

Furthermore, as a positive relationship between firm size (or foreign ownership) and wages is hypothesised to reflect unobservable firm productivity differences and not just an accumulation of more able workers by larger (or foreign-owned) firms, we expect this relationship to survive when we control for observable worker characteristics as well as unobservable time-invariant worker differences (individual fixed effects). For the same reason, we expect the positive relationship to be reduced significantly when controlling for unobservable time-invariant firm differences (firm fixed effects).

Second, with respect to the relationship between *current firm type and wage growth*,

we expect a similar set of results: In a cross section, we should observe higher wage growth for workers in foreign-owned and/or large firms, where the effect of foreign-ownership should disappear (or be reduced) with the inclusion of firm size. A positive relationship should again survive controlling for worker differences, and be reduced significantly when controlling for unobservable time-invariant firm differences.

Third, when it comes to the relationship between *previous firm type and the current wage level*, our data only allow us to analyse the effects of previous experience from large firms, as foreign ownership is only observed since 2000. Thus, in a cross section, we should observe higher wages for workers with previous experience from large firms, and this relationship should not disappear when controlling for worker differences. Instead, it should increase with the amount of previous experience from a large firm. Similarly, we should observe higher earnings for new self-employed with a background in a foreign-owned and/or a large firm; an effect which should increase with the amount of experience from such firms.

We test these three sets of hypotheses formally in the following section by regressing wages and wage growth rates on worker and firm characteristics. However, before turning to that, we take a look at some descriptive statistics.

Table 4.1 presents the number of firms as well as the total employment of foreign-owned (F) and domestically-owned (D) firms in Denmark in the years 2000-2002 divided into different size classes. While the total stock of firms averaged approximately 245,000, only slightly more than 1% of these were foreign owned in the years 2000-2002. However, as also shown in the Table, the foreign firms were considerably larger on average, which implies that they accounted for 12-15% of total employment.[6] Note that this relationship between size and

---

[6] The employment figures in Table 4.1 are based on firm-level information about full-time employees. Note that part of the difference between foreign- and domestically-owned firms may be due to the fact that some of the smaller foreign-owned firms are not classified as foreign-owned in the data, as it requires FDI of a certain amount (see footnote above).

ownership is fully consistent with the implications of our theory.

In Table 4.2, we provide a first check of the relationship between firm type, wage levels, and wage growth. The Table contains the average wages and average wage growth rates for employees in foreign-owned and domestically-owned firms, respectively, as well as in different size classes. The income measure used is an hourly (nominal) wage computed by Statistics Denmark. As hypothesized, the average wages reveal a significant wage gap between domestically- and foreign-owned firms (more than 16% in each of the three years) as well as between small and large firms (10-12%) .

The Table also shows that average wage growth is higher in larger firms. As an example, the difference in wage growth rates between small (<50 employees) and large (>500 employees) firms was 1.0 percentage points in 2001-2, which corresponds to 36% higher annual wage growth in large firms. The difference between foreign-owned firms and domestically-owned firms is much smaller.

While the numbers from Table 4.2 are fully in line with the predictions of our theory model, they do not control for any background characteristics of the individuals, such as education, age and experience. We return to this in the next section.

Note that the Danish labor market is characterized by a high degree of flexibility as firing costs are extremely low. In that vein, Denmark compares better to US and UK labor markets than to the labor markets of the larger European countries. At the same time, the Danish welfare state takes care of the unemployed through for example particularly high compensation rates which is why the Danish model is often termed "Flexicurity". Thus, Danish labor market data seem particularly useful for analyzing productivity transfers through worker mobility.

The data also reveal that a considerable amount of individuals flow between foreign- and domestically-owned firms, and between small and large firms, each year. As an example,

around 20% of those employed in a foreign-owned firm in a given year move to another firm the following year (see Table 4.3).  Out of these, around two thirds end up in a domestically-owned firm. Of those employed in large (>500 employees) firms, a similar share move to another firm, but this time only around 1/3 end up in a small firm the following year (see Table 4.4). Thus, judged from the mere amount of observed mobility, there is considerable potential for transfers of acquired skills across firms.

5.      Empirical Results

This section provides a formal test of the three sets of empirical hypotheses derived in the previous section.  In Section 5.1, we test the hypothesis that wages are higher in large and/or foreign-owned firms, whereas in Section 5.2, we concentrate on the hypothesis of higher wage growth in these firms.  Finally, in Section 5.3 we turn to the hypothesis of transferability of skills acquired in previous employments by focusing on (i) the relationship between the current wage level and previous experience from large firms; and (ii) the relationship between the earnings of new self-employed and previous experience from large and foreign-owned firms. Variable definitions can be found in Appendix B.

5.1      Wage Levels and Current Firm Type

In this section, we test the hypothesis of higher wage levels in large and/or foreign-owned firms by regressing individual wages on worker and firm characteristics.  Columns 1-4 of Table 5.1 report the results of ordinary least squares (OLS) regressions  for various specifications of the right hand side.  The dependent variable in all regressions is the log hourly wage.

In the first column, the individual wage is regressed only on a dummy for current employment in a foreign-owned firm, as well as a set of year dummies.  The positive and strongly significant coefficient confirms the relationship from Table 4.2.  In column 2, the log

of firm size is also included. As hypothesised, the coefficient to this variable is also positive and strongly significant, and its inclusion reduces the coefficient to the foreign-ownership dummy, but does not eliminate it. In column 3, a dummy taking the value one if the worker is employed in an exporting firm is included as an additional proxy for HP firm type. Not surprisingly, the estimated coefficient to this dummy is also positive, and it decreases the estimated coefficient to firm size by approximately 1/4.

Addition of (observable) worker characteristics in column 4, such as age, experience, gender, region, and industry dummies, lowers, but far from eliminates, the effect of foreign ownership and firm size. Summarizing the OLS results in Table 5.1, the move from column 1 to column 4 reduces the foreign-ownership premium by about half (0.070 versus 0.134). Acknowledging that the theory presents a pure case in which the foreign-ownership premium should be reduced to zero, we feel that these results are consistent with the theory but still leave something of an unexplained premium. On the other hand, the fact that adding worker characteristics still leaves a large firm-size premium is perfectly consistent with the theory.

While the positive coefficients to firm size and foreign ownership found in columns 1-4 support the empirical predictions of our theory, namely that large and/or foreign-owned firms pay higher wages due to unobserved productivity advantages, they could alternatively reflect: (a) that large and foreign-owned firms pick or attract workers who are more "able" in some unobservable way; and/or (b) that growing firms or firms taken over by foreigners increase the wages of their workers.

To test the importance of (a), we run the regressions in columns1-4 including individual fixed effects to eliminate any time-invariant unobservable worker differences. The results are contained in columns 5-8 of Table 5.1. While this significantly reduces the coefficient to the foreign ownership dummy from around 10% to 1%, it only slightly affects the firm size effect. The elasticity of wages with respect to firm size is still found to be around 0.01 and strongly

significant. Thus, the effects of firm size (and to some extent foreign-ownership) survive when controlling for all individual characteristics, including the unobservable ones, as our theory would predict.

Note, however, that with the short panel, the individual fixed-effects regression can only pick up the short run effects of a change in ownership status or firm size. This may explain why the coefficient to foreign ownership drops, as we would not expect the full wage premium to materialize until after a couple of years in a foreign firm.[7] In this light, it may be surprising that the coefficient to firm size only drops slightly, but it could reflect either rapid learning and/or the importance of minimum wages and/or progressive taxes. It could also reflect that large firms in general tend to attract less "able" workers which in itself would exercise a negative effect on the firm-size coefficient in the ordinary least squares regression.

In any case, the results are consistent with a wage premium in large (and to some extent foreign owned) firms already in the short run, which is not due to a different mix of workers in these firms. We shall get back to the long run effects in the following sections when considering wage growth (Section 5.2) and the effects in a longer panel (Section 5.3).

Finally, to test the importance of (b), we add firm fixed effects instead of individual fixed effects in column 9 of Table 5.1. Our theory predicts that if we control for firm heterogeneity by using firm fixed effects, the estimated effect of firm size (and foreign ownership) should be reduced significantly. If, instead, (b) is the explanation for the positive relationship between firm size (and foreign ownership) and wages, the estimated relationship should not be affected by the inclusion of firm fixed effects.

From column 9, we observe that the positive coefficients to firm size and foreign ownership do in fact disappear (they even become negative) with the inclusion of firm fixed

---

[7] An alternative explanation for the drop in the coefficient to foreign ownership is of course that these firms tend to attract better (high-wage) workers.

effects. This provides additional support for the hypothesis that the higher wages in large and foreign-owned firms are indeed caused by unobserved productivity advantages.

The result that foreign-owned firms pay higher wages is consistent with a number of existing studies; see, *e.g.*, Doms and Jensen (1987), Aitken et al. (1996), Feliciano and Lipsey (1999), Griffith and Simpson (2003), and Girma and Görg (2006). The few existing studies based on matched employer-employee data, as in the current study, also find that the overall "wage-gap" between foreign-owned and domestically-owned firms is reduced significantly when controlling for firm and worker characteristics; see Heyman et al. (2004) and Martins (2004).

The positive relationship between firm size and wages has also been documented in a number of studies, including Idson and Feaster (1990) and Bayard and Troske (1999). As opposed to our study, Abowd et al. (1999) find that individual fixed effects remove most of the relationship. Brown and Medoff (1989) and Evans and Leighton (1989) also find that controlling for individual heterogeneity by estimating the relationship in first differences reduces but does not eliminate the positive effect of firm size.

## 5.2    Wage Growth and Current Firm Type

In this section, we test the hypothesis of higher wage growth in large and/or foreign-owned firms by regressing individual wage growth within job spells on worker and firm characteristics. Using the change in log hourly wages as the dependent variable, Columns 1-4 in Table 5.2 report the results of ordinary least squares estimates for various specifications of the right hand side.

The results show that the coefficient to firm size is significantly positive, and robust to the inclusion of observable worker characteristics such as age, education and experience. The estimated coefficients imply that, *e.g.*, a doubling of firm size should be associated with

approximately 0.2 percentage points higher wage growth. With an annual wage growth rate of 2%, this corresponds to 10% higher wage growth.

Foreign ownership also has a positive effect on wage growth when firm size is not included (column 1). However, this effect disappears with the inclusion of firm size in columns 2-3, as predicted by our theory. In fact, foreign ownership appears in itself to have a negative effect on wage growth when controlling for firm size, although the effect is much less statistically significant than that of firm size. In Column 3, we control for a number of observable worker characteristics as in Table 5.1. This in fact slightly increases the firms-size coefficient.

As in Section 5.1, we would like to control for unobservable worker differences as well to determine whether the positive relationship between firm size and wage growth could be the result of a different composition of worker types in these firms, *e.g.*, workers with higher learning potential and therefore higher wage growth. However, as we only observe one job spell for each individual due to the short nature of our panel, it is not possible to include individual fixed effects. Furthermore, the potential problem of unobservable worker differences seems much less pertinent when considering wage growth rates instead of wage levels. For example, unobservable time-invariant ability differences that affect the wage level do not affect wage growth.

Finally, in column 4 of Table 5.2, we add firm fixed effects to test whether the positive relationship between firm size and wage growth just reflects that growing (or shrinking) firms increase (or decrease) their wage growth rates. This does not appear to be the case, as the coefficient to firm size becomes strongly negative with the inclusion of firm fixed effects.[8]

---

[8] An explanation for the negative coefficient to firm size in column 4 may be that expanding firms hire workers who have less wage growth in their first years of employment than already employed workers. This difference between new and old workers may in turn result from imperfect ex-ante information about the new workers' productivities and thus their potential for wage growth. Furthermore, if firms that downsize tend to lay off the unproductive workers (those with less wage

Instead, the result strongly indicates that the higher wage growth in large firms does in fact reflect unobservable firm differences, as our theory predicts.

While a number of studies have previously dealt with the relationship between firm size (or foreign ownership) and wage *levels*, much fewer studies have considered the effects on wage *growth*, and the existing results are mixed. Using cross-section data, Pearce (1990) finds larger effects of tenure in large firms, whereas Baron et al. (1987) using panel data (for two years) find a negative relationship between size and wage growth. In a somewhat different context, Møen (2005) finds higher wage growth in R&D intensive firms using an estimation method similar to the one we use above.

### 5.3    Earnings and Previous Experience

In this section, we turn to the hypothesized relationship between previous experience from large and/or foreign-owned firms and current earnings, which should be brought about by a transfer of acquired skills from one occupation to the next. This is done by extending the estimations from Table 5.1 with measures of previous experience. To construct such measures, we need a longer panel, and we therefore extend it to the period 1981-2003. This implies that we cannot include measures of foreign ownership (and exports) as this variable is unavailable prior to the year 2000 (1995 for exports). Furthermore, instead of firm size, we use plant size, as the link between plants and firms is also incomplete for the earlier years.

Columns 1-4 of Table 5.3 report the results of ordinary least squares wage regressions for various specifications of the right hand side, whereas columns 5-6 include individual fixed effects. The regressions show the same picture as in Table 5.1. The coefficient to the log of plant size is significantly positive in all regressions and robust to the inclusion of individual fixed effects. As previously, it disappears with the inclusion of firm (plant) fixed effects

---

growth) first, we will get a negative coefficient to firm size when controlling for firm fixed effects.

(column 7).

Furthermore, while tenure in the current firm in general increases the wage level, there is an additional positive effect of tenure if the current plant is large (column 3). This supports the finding from Table 5.2 that wage growth is higher in larger firms. Thus, our empirical findings are fully consistent with not only a short-run wage premium but also higher subsequent wage growth in large (and foreign-owned) firms.

Perhaps even more interestingly, there is an additional positive effect of experience if this experience is from a large plant. This can be seen from a comparison of the coefficients to "experience" and "experience from large plants". In fact, while each extra year of experience initially adds around 3% to the wage (an effect which is declining to around 1.5% per year after 10 years of experience due to the quadratic term in the regressions), this effect is increased by another 1.4 percentage points if the experience is from a large plant. This strongly supports the hypothesis of transferable skills. This is further confirmed by the finding that this effect is preserved with the inclusion of plant fixed effects in column 7. The effects of previously acquired skills do not disappear when controlling for current plant characteristics.

An additional test of the transferability of skills is to consider the earnings of new self-employed. Table 5.4 contains OLS estimates of the relationship between earnings of new self-employed and their previous experience from large and foreign-owned firms. Specifically, in columns 1 and 2, the log of annual earnings is regressed on a number of individual characteristics as well as dummies for the individual being employed in a foreign-owned and/or a large firm the year before. While employment in a foreign-owned firm increases income as self-employed by 13.5%, this effect is reduced to 7.3% when also controlling for the size of the firm.[9] The elasticity of firm size, on the other hand, is found to be around 2%.

---

[9] Note that the smaller t-values in the estimations in Table 5.4 reflect the much smaller sample sizes when considering exclusively the new self-employed.

In columns 3-5, we add the *amount* of experience from foreign-owned and large firms - computed as the tenure in the last job spell if that was in a firm that was large or foreign-owned, respectively, in the final year of employment. In the case of foreign-owned firms, this moves the effect away from the simple dummy variable to the experience variable, indicating that not just experience from a foreign-owned firm, but also the amount of this experience matters for self-employment income. However, this effect is not statistically significant when it comes to experience from large firms.

Very few papers in the literature have considered the effects of previous workplace (or firm) characteristics on current wages and self-employment income. Møen (2005), however, finds that a higher R&D intensity in previous employment increases the positive effect on wages of previous experience. Martins (2005) also finds some evidence of higher wages for workers moving from foreign to domestic firms compared to their colleagues in domestic firms. In a developing country context, Görg and Strobel (2005), using data from Ghana, find positive effects on self-employment earnings of previous experience from foreign-owned firms.

In sum, this section has provided considerable support for the empirical implications of our theoretical model. Wage levels and wage growth are higher in large and/or foreign-owned firms, and these effects survive controlling for worker characteristics but are reduced significantly when controlling for unobservable firm characteristics, as predicted by our model. Furthermore, skills appear to be transferable to both subsequent wage work and self-employment.


6.      Summary and Conclusions

The paper is motivated by the interests by both researchers and policy makers in possible beneficial effects of foreign companies on local companies and workers. We focus on workers, and on the direct effects of working for a high-productivity firm on the individual's

productivity in subsequent wage work or self-employment. Using a much-simplified version of Meltiz's heterogeneous firms model, our theory model predicts that foreign firms pay higher average wages, their workers have higher average wage growth, and they earn more in subsequent self-employment for workers who switch.

However, this is due to foreign firms having a higher average productivity, in turn due to the inability of low-productivity foreign firms to enter due to fixed entry costs. High productivity firms are larger, and hence our model also predicts that foreign firms are not more productive than larger domestic firms. In other words, most of the favorable effects of foreign firms disappear when correcting for firm size.

Several experiments with the model indicate that the average wage premium in foreign firms is reduced as either (1) the productivity in MP firms of workers who switch to MP firms from HP firms is increased or (2) the absolute expected value of self-employment earnings when switching from an HP firm grows relative to switching to self-employment from an MP firm. This fall in the average wage premium is due to a decrease in the share of HP firm workers who are experienced and also in (2) by a willingness to work for less in the first period in a HP firm due to the higher expected payoff in self-employment. In these cases, the HP firms are partly performing the function of "nurseries", training inexperienced workers who work on the cheap and then leave for MP firms or self-employment.

On the other hand, the theory model also concludes that the wage premium is increased by either (1) a minimum wage which prevents HP firms from paying a low initial wage and/or (2) a progressive income tax that hits the second-period earnings of workers in HP firms (or transiting to MP firms). Both of these factors seem empirically relevant, and should lead to the observation of a higher wage premium simultaneously with higher wage growth in HP firms.

Our theory model can be used to derive a number of testable empirical hypotheses about wage levels, wage growth, and productivity transfers. These are tested using matched

data between individual Danish workers and firms in the second part of the paper.

Consistent with our theory, we find that working for a foreign owned firm significantly (1) increases the worker's wage, (2) increases the worker's wage growth, and (3) increases subsequent earnings of new self-employed. We also find that effects (1) and (3) are significantly reduced, but not eliminated, by controlling for firm size. Effect (2) disappears completely when controlling for firm size. Firm (plant) size is also in itself found to have a significantly positive effect on subsequent earnings as wage employed (where information on foreign ownership is not available). Thus the empirical results are certainly close to our theoretical predictions although, the foreign ownership effect does not disappear in all cases.

However, as also consistent with our theory, the effects of firm size and foreign ownership do disappear with the inclusion of firm (or plant) fixed effects, while they survive controlling for worker differences (observable and unobservable).

There are several plausible reasons for the residual foreign-ownership effect after controlling for firm size and observable worker characteristics. One is simply that the top end of the productivity distribution of foreign firms is higher than that for domestic firms and so the average entering foreign firm has a higher average productivity than the higher-productivity domestic firms. This is pretty ad hoc and again, within the heterogenous-firms approach, those higher productivity foreign firms would then have higher outputs, so the effect on wages should disappear controlling for size. Of course, the effect of size may be non-linear, offering a second explanation for the residual foreign-ownership effect, and we are considering that.

Second, and perhaps related, is that foreign firms are somehow able to pick the best workers and the characteristics in question are unobservable. This obviously calls for fixed effects at the worker level, but the short nature of our panel creates difficulties. All we are able to identify with our fixed-effects regressions is essentially the initial wage premium from switching to a foreign firm. We estimate this to be positive, but quite small . But this is

perfectly consistent with the theory model, which predicts a low or negative initial wage

premium in equilibrium which balances higher earnings later on. Thus the question of whether

or not the foreign firms have a better ability to select must await further research.[10]

---

[10]As noted in our literature review, the alternative approach of Yeaple (2005) with ex ante identical firms and heterogeneous workers deserves a careful and thorough examination in the empirical context.

Appendix A: analytical solutions

This appendix gives analytical solutions to the model for the case where there is some

(but not full) transiting in equilibrium, i.e. $X_t^m > 0$ and $X_2^h > 0$.

Wages for workers in HP firms can be solved for from (2), (3), and (9).

$$\frac{w_1^h}{w_2^h} = \frac{r_1^h}{r_2^h} \quad \text{from (2) and (3), so using these in (9) gives} \tag{A1}$$

$$w_1^h = \frac{(1 + \beta)}{(1 + \beta(r_2^h/r_1^h))} < 1 \qquad w_2^h = \frac{(1 + \beta)}{(r_1^h/r_2^h + \beta)} > 1 \tag{A2}$$

where $\beta = \delta(1 - \alpha + \alpha v)$. $\beta = \delta$ in the base case with no self-employment probability.

The price of a representative good from an HP firm is solved for from (3) given that we know $w_1^h$ from (A2).

$$p^h = \frac{\sigma}{\sigma - 1} \frac{1 + \beta}{(r_1^h + \beta r_2^h)} \tag{A3}$$

And the price of a representative good from an MP firm is solved for from (4).

$$p^m = \frac{\sigma}{\sigma - 1} > p^h \tag{A4}$$

The output of a representative good from an MP firm is solved for from (10) given that

we know the output price from (A4).

$$X^m + X_t^m = (\sigma - 1)F^m \tag{A5}$$

The total output of a representative good from an HP firm is solved for from the

consumer's marginal rate of substitution condition.

$$\frac{X_1^h + X_2^h}{X^m + X_t^m} = \left[\frac{p^h}{p^m}\right]^{-\sigma} \tag{A6}$$

Given that we know $X^m + X_t^m$, $p^h$, $p^m$, we then have

$$X_1^h + X_2^h = \left[\frac{r_1^h + \beta r_2^h}{1 + \beta}\right]^\sigma (\sigma - 1)F^m > X^m + X_t^m \tag{A7}$$

We then have one remaining variable, $n^m$, and parameter, $I$. The expenditure-income equation is

$$p^m n^m (X^m + X_t^m) + p^h (n_d^h + n_f^h)(X_1^h + X_2^h) = I \tag{A8}$$

All of the endogenous variables except $n^m$ are now known and so this gives one remaining equation in one unknown. In our base case numerical solution, we chose an initial value of $n^m = 50$ and this then calibrates, given our choice that $n_d^h + n_f^h = 10$ and the endogenous values of prices already solved for, to a value of $I = 116.6667$. This is given by the equation

$$I = n^m \frac{\sigma}{\sigma - 1}(\sigma - 1)F^m + (n_d^h + n_f^h)\frac{\sigma}{\sigma - 1}\frac{1 + \beta}{(r_1^h + \beta r_2^h)}\left[\frac{r_1^h + \beta r_2^h}{1 + \beta}\right]^\sigma (\sigma - 1)F^m$$

or just

$$I = n^m \sigma F^m + (n_d^h + n_f^h)\sigma\left[\frac{r_1^h + \beta r_2^h}{1 + \beta}\right]^{\sigma - 1} F^m \tag{A9}$$

The calibrated value of $I$ is then held constant in the subsequent analysis, and (A9) can be inverted to give the equilibrium value of $n^m$.

To get average wages within the firm, we have to push the analysis further and must solve for the shares in (6) and (7). This requires us to make use of (8). From (13),

$$s_t^m = 1 - (r_t^m - \gamma)/\rho \qquad \text{where} \qquad r_t^m = w_2^h = \frac{(1 + \beta)}{(r_1^h/r_2^h + \beta)} \tag{A10}$$

and $s_t^m \geq 0$ by virtue of the non-negativity constraint on $X_t^m$.

Briefly, we have the share of transit workers in MP firms along with the equilibrium

output per firm in (A5)

$$s_t^m = \frac{X_t^m}{X_t^m + X^m} = 1 - \frac{(1+\beta)}{(r_1^h/r_2^h+\beta)}\frac{1}{\rho} + \frac{\gamma}{\rho} \qquad X^m + X_t^m = (\sigma-1)F^m \tag{A11}$$

These two equations can be solved to get $X_t^m$ .

$$X_t^m = \left[\rho + \gamma - \frac{1+\beta}{r_1^h/r_2^h+\beta}\right]\frac{(\sigma-1)}{\rho}F^m \tag{A12}$$

Finally, (A11), (A12), and the number of MP firms from (A9) are inserted into

equation (8). The only remaining variables in (8) are then $X_1^h$ and $X_2^h$ . As a consequence, (8)

can be reduced to

$$(1-\alpha)X_1^h/r_1^h - X_2^h/r_2^h = T/R \tag{A13}$$

$$R = n_d^h(F_d^h/Q^h+1) + n_f^h(F_f^h/Q^h+1) \qquad Q^h = X_1^h + X_2^h$$

$$T = n^m(s_t^m F^m/r_t^m + X_t^m/r_t^m)$$

Equation (A13) and (A7) allow us to solve for two equations in two unknowns.

$$X_1^h = \frac{Q^h/r_2^h + T/R}{(1-\alpha)/r_1^h + 1/r_2^h} \qquad X_2^h = \frac{(1-\alpha)Q^h/r_1^h - T/R}{(1-\alpha)/r_1^h + 1/r_2^h} \qquad X_1^h + X_2^h = Q^h \tag{A14}$$

The above equations hold as long as there is transit in equilibrium. The minimum value

such that below this value there is no transit, is given by setting (A12) equal to zero.

$$\gamma_{min} = \frac{1+\beta}{r_1^h/r_2^h+\beta} - \rho \tag{A15}$$

The maximum value such that above this value all workers transit from HP firms is given by setting the number of first-period workers who do not go into self-employment equal to the number of transit workers. This is just equation (8) without the first two terms on the right hand side. Using that

$$\frac{n^m}{r_t^m}\left(s_t^m F^m + X_t^m\right) = \frac{(1-\alpha)}{r_1^h}\left[\left(n_d^h + n_f^h\right)\left(\frac{r_1^h + \beta r_2^h}{1+\beta}\right)^{\sigma}(\sigma-1)F^m + n_f^h F_f^h + n_d^h F_d^h\right] \quad s_1^h = 1$$

and that $X_1^h$ equals (A7), we get:

$$(A16)$$

which using (A11) and (A13) can be solved to.

$$\gamma_{\max} = \gamma_{\min} + \frac{\rho(1-\alpha)}{r_1^h}\left[\frac{\left(n_d^h + n_f^h\right)\left(\frac{r_1^h + \beta r_2^h}{1+\beta}\right)^{\sigma}(\sigma-1)F^m + n_f^h F_f^h + n_d^h F_d^h}{I - \sigma F^m\left(n_d^h + n_f^h\right)\left(\frac{r_1^h + \beta r_2^h}{1+\beta}\right)^{\sigma-1}}\right]\frac{1+\beta}{\left(r_1^h / r_2^h + \beta\right)} \quad (A17)$$

With respect to the experiments conducted in the paper, note that the share of transiting workers (or alternatively we report the share of first-period HP workers who transit) is increasing in $\gamma$, which is the experiment in Table 3.1. Also, $\beta$ is increasing $\alpha$, and $\nu$, the probability of a successful self-employment draw and the self-employment premium, respectively. The share of transiting workers is increasing in $\beta$ and therefore in $\alpha$ and $\nu$. The former is the experiment in Table 3.2, so both analytical findings are confirmed in the simulations.

Appendix B: Definitions of variables used in regressions:

*Variables available in both panels*

| Variable | Definition |
|---|---|
| Hourly wage (Continuous variable) | Wages divided by number of hours worked. The number of hours is imputed from mandatory pension payments, which are determined by the number of hours in employment per week. These estimates are computed by Statistics Denmark. |
| Gross annual earnings (Continuous variable) | All taxable income. |
| Experience: (Continuous variable). | This variable is a continuous measure of actual labor market experience based on the number of days in employment over the worker's career. Experience is measured in number of years of full time work. |
| Years of education (Count variable). | Scheduled number of years of completed of education. Examples: High-school = 12 years; Master degree = 18 years. |

*Variables only available in the short panel*

| Variable | Definition |
|---|---|
| Foreign (Dummy variable). | Takes the value one for workers employed in firms where foreigners ultimately own more than 50% of the firm, and FDI amounts to more than DKK 10 million. Zero otherwise. |
| Firm size (Count variable). | The average number of full-time employees (within a year) in the firm where the individual is employed. The firm is defined as the legal entity which employ the worker. |
| Exporter (Dummy variable). | Takes the value one for workers employed in firms with exports. Zero otherwise. |

*Variables only available in the long panel*

| Variable | Definition |
|---|---|
| Plant size (Dummy variable). | The number of employees in the last week of November at the workplace (plant) where the individual is employed. A workplace is defined by its address. |
| Tenure (Count variable) | The number of years employed at the current workplace. Tenure is reset to zero when the individual changes workplace. |
| Large (Dummy variable) | Takes the value one for workers employed at workplaces with more 50 employees. Zero otherwise. The dummy variable *Large* is then interacted with *Experience* and *Tenure*. |
| Experience from large plants (Count variable) | Total number of years of employment at workplaces with more 50 employees, measured from the beginning of the individual's career to the current date. |

REFERENCES

Abowd, J.M., F. Kramarz, and D.N. Margolis (1999): "High Wage Workers and High Wage Firms", *Econometrica*, 67, 251-333.

Aitken, B., A. Harrison, and R.E. Lipsey (1996): "Wages and Foreign Ownership: A Comparative Study of Mexico, Venezuela, and the United States", *Journal of International Economics*, 40, 345-371.

Almeida, R. (2003): "The Effects of Foreign Owned Firms on the Labor Market", IZA Discussion Paper No. 785.

Antrás, P. (2003): "Firms, Contracts, and Trade Structure"*, Quarterly Journal of Economics*, 118, 1374-1418.

Barron, J.M., D.A. Black, and M.A. Loewenstein (1987): "Employer Size: The Implications for Search, Training, Capital Investment, Starting Wages, and Wage Growth", *Journal of Labor Economics*, 5, 76-89.

Bayard, K. and K.R. Troske (1999): "Examining the Employer-Size Wage Premium in the Manufacturing, Retail Trade, and Service Industries Using Employer-Employee Matched Data", *American Economic Review*, 89, 99-103.

Brown, C. and J. Medoff (1989): "The Employer Size-Wage Effect", *Journal of Political Economy*, 97, 1027-1059.

Doms, M.E. and J.B. Jensen (1987): "Comparing Wages, Skills, and Productivity between Domestically and Foreign-Owned Manufacturing Establishments in the United States", in Baldwin, R.E., R.E. Lipsey and J.D. Richardson (eds.): *Geography and Ownership as Bases for Economic Accounting.* University of Chicago Press, Chicago.

Ethier, W.J. and J.R. Markusen (1996): "Multinational Firms, Technology Diffusion and Trade", *Journal of International Economics*, 41, 1-28.

Evans, D. and L. Leighton (1989): "Why do Smaller Firms Pay Less?", *Journal of Human Resources*, 24, 299-318.

Feliciano, Z.M. and R.E. Lipsey (1999): "Foreign Ownership and Wages in the United States, 1987-1992", NBER Working Paper No. 6923.

Fosfuri, A., M. Motta, and T. Rønde (2001): "Foreign Direct Investments and Spillovers Through Workers' Mobility", *Journal of International Economics*, 53, 205-222.

Girma, S. and H. Görg (2006): "Evaluating the Foreign Ownership Wage Premium using a Difference-in-Differences Matching Approach", *Journal of International Economics* (forthcoming).

Glass, A.J and K. Saggi (2002): "Multinational Firms and Technology Transfer", *Scandinavian Journal of Economics*, 104, 495-513, 2002.

Görg, H. and E. Strobl (2005): "Spillovers from Foreign Firms through Worker Mobility: An Empirical Investigation", *Scandinavian Journal of Economics*, 107, 693-709.

Griffith, R. and H. Simpson (2003): "Characteristics of Foreign-Owned Firms in British Manufacturing", NBER Working Paper No. 9573.

Haddad, M, and A, Harrison (1993): "Are there Positive Spillovers from Direct Foreign Investment: Evidence from Panel-Data for Morocco", *Journal of Development Economics*, 42, 51-74.

Haskel, J.E., S.C. Pereira, and M.J. Slaughter (2002): "Does Inward Foreign Direct Investment Boost the Productivity of Domestic Firms?", NBER working paper 8724.

Helpman, Elhanan, Marc J. Melitz and Stephen R. Yeaple (2004): "Exports versus FDI with Heterogeneous Firms", *American Economic Review* 94, 300-316.

Heyman, F., F. Sjöholm, and P.G. Tingvall (2004): "Is there Really a Foreign Ownership Wage Premium?", FIEF Working Paper No. 199.

Idson, T.L. and D.J. Feaster (1990): "A Selectivity Model of Employer-Size Wage Differentials", *Journal of Labor Economics,* 8, 91-122.

Javorcik, B.S. (2004): "Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers through Backward Linkages", *American Economic Review*, 93, 605-627.

Keller, W. (2004), "International Technology Diffusion", *Journal of Economic Literature* 44, 752-782.

Lipsey, R.E. (2002): "Home and Host Country Effects of FDI", NBER Working Paper No. 9293.

Markusen, J.R. (2001): "Contracts, Intellectual Property Rights, and Multinational Investment in Developing Countries", *Journal of International Economics*, 53, 189-204.

Markusen, J.R. and N. Trofimenko (2006): "Teaching Locals new tricks: Foreign Experts as a source of Productivity Transfers", NBER working paper 12872, http://spot.colorado.edu/~markusen.

Martins, P.S. (2004): "Do Foreign Firms Really Pay Higher Wages? Evidence from Different Estimators", IZA Discussion Paper No. 1388.

Martins, P.S. (2005): "Inter-Firm Employee Mobility, Displacement, and Foreign Direct Investment Spillovers", working paper.

Melitz (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity", *Econometrica*, 71, 1695-1725.

Moen, J. (2005): "Is Mobility of Technical Personnel a Source of R&D Spillovers?", *Journal of Labor Economics*, 23, 81-114.

Pearce, J.E. (1990): "Tenure, Unions, and the Relationship between Employer Size and Wages", *Journal of Labor Economics*, 8, 251-269.

Yeaple, S. (2005): "A Simple Model of Trade, Wages, and Firm Heterogeneity", *Journal of International Economics*, 65, 1-20.

## Table 3.1: increase in parameter $\gamma$ ( determining $r^m_t$ ), producitivty of HP firm transiting to MP firms.

| | 0.940 | 0.950 | 0.960 | 0.970 | 0.980 | 0.990 |
|---|---|---|---|---|---|---|
| $\gamma$ | | | | | | |
| $r^m_t$ | 1.190 | 1.200 | 1.200 | 1.200 | 1.200 | 1.200 |
| WH1 | 0.900 | 0.900 | 0.900 | 0.900 | 0.900 | 0.900 |
| WH2 | 1.200 | 1.200 | 1.200 | 1.200 | 1.200 | 1.200 |
| average wage growth in an HP firm | 0.333 | 0.333 | 0.333 | 0.333 | 0.333 | 0.333 |
| average wage domestic firm | 1.009 | 1.009 | 1.012 | 1.016 | 1.019 | 1.023 |
| average wage foreign firm | 1.050 | 1.050 | 1.038 | 1.026 | 1.015 | 1.004 |
| total output of an MP firm | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| total outpout of an HP firm | 4.630 | 4.630 | 4.630 | 4.630 | 4.630 | 4.630 |
| share of first period HP workers who transit to MP firms | | | 0.119 | 0.226 | 0.323 | 0.410 |

## Table 3.2: increase in the probability of a favorable self-employment draw, $\alpha$

| $\alpha$ | 0.000 | 0.020 | 0.040 | 0.060 | 0.080 | 0.100 |
|---|---|---|---|---|---|---|
| WH1 | 0.900 | 0.894 | 0.889 | 0.884 | 0.879 | 0.875 |
| WH2 | 1.200 | 1.192 | 1.185 | 1.179 | 1.172 | 1.167 |
| average wage growth in an HP firm | 0.333 | 0.333 | 0.333 | 0.333 | 0.333 | 0.333 |
| average wage domestic firm | 1.009 | 1.007 | 1.006 | 1.005 | 1.006 | 1.006 |
| average wage foreign firm | 1.050 | 1.042 | 1.034 | 1.025 | 1.011 | 0.998 |
| total output of an MP firm | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| total outpout of an HP firm | 4.630 | 4.720 | 4.805 | 4.887 | 4.964 | 5.038 |
| share of first period HP workers who transit to MP firms | | | | 0.017 | 0.089 | 0.151 |

## Table 3.3: increase in the minimum wage, which impacts only on wh1

| $w^h_1$ | 0.9 | 0.92 | 0.94 | 0.96 | 0.98 | 1 |
|---|---|---|---|---|---|---|
| WH1 | 0.900 | 0.920 | 0.940 | 0.960 | 0.980 | 1.000 |
| WH2 | 1.200 | 1.227 | 1.253 | 1.280 | 1.307 | 1.333 |
| average wage growth in an HP firm | 0.333 | 0.333 | 0.333 | 0.333 | 0.333 | 0.333 |
| take home wage growth in an HP fi | 0.333 | 0.333 | 0.333 | 0.333 | 0.333 | 0.333 |
| average wage domestic firm | 1.010 | 1.014 | 1.018 | 1.020 | 1.023 | 1.025 |
| average wage foreign firm | 1.050 | 1.073 | 1.097 | 1.120 | 1.143 | 1.167 |
| total output of an MP firm | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| total outpout of an HP firm | 4.630 | 4.334 | 4.063 | 3.815 | 3.586 | 3.375 |

share of first period HP workers
who transit to MP firms

## Table 3.4: progressive income tax (tax on earnings above w = 1)

| tax rate | 0 | 0.04 | 0.08 | 0.12 | 0.16 | 0.20 |
|---|---|---|---|---|---|---|
| WH1 | 0.900 | 0.914 | 0.927 | 0.940 | 0.953 | 0.964 |
| WH2 | 1.200 | 1.219 | 1.237 | 1.254 | 1.270 | 1.286 |
| average wage growth in an HP firm | 0.333 | 0.333 | 0.333 | 0.333 | 0.333 | 0.333 |
| take home wage growth in an HP fi | 0.333 | 0.282 | 0.235 | 0.190 | 0.149 | 0.111 |
| average wage domestic firm | 1.010 | 1.013 | 1.016 | 1.018 | 1.019 | 1.021 |
| average wage foreign firm | 1.050 | 1.066 | 1.082 | 1.097 | 1.111 | 1.125 |
| total output of an MP firm | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| total outpout of an HP firm | 4.630 | 4.419 | 4.230 | 4.060 | 3.905 | 3.764 |

**Table 4.1: Firm Types and Employment**

| Firm Size (# employees) | Absolute numbers | | | | | |
| | 2000 | | 2001 | | 2002 | |
| | Domestic | Foreign | Domestic | Foreign | Domestic | Foreign |
| --- | --- | --- | --- | --- | --- | --- |
| | | | Firms | | | |
| 0-49 | 241,946 | 1,966 | 240,393 | 2,037 | 237,605 | 2,119 |
| 50-499 | 2,632 | 573 | 2,631 | 579 | 2,586 | 564 |
| 500+ | 195 | 58 | 194 | 52 | 186 | 55 |
| Total | 244,773 | 2,597 | 243,218 | 2,668 | 240,377 | 2,738 |
| | | | Employment | | | |
| 0-49 | 551,159 | 23,245 | 543,985 | 23,738 | 537,882 | 24,250 |
| 50-499 | 302,658 | 84,765 | 304,132 | 88,435 | 299,516 | 86,621 |
| 500+ | 322,316 | 74,561 | 321,540 | 75,598 | 298,222 | 84,360 |
| Total | 1,176,133 | 182,571 | 1,169,657 | 187,771 | 1,135,620 | 195,231 |

Note: The table includes all full-time workers in the private sector. The division of firms into size classes is based on the average number of employees over the year.

**Table 4.2: Firm Types, Average Wages and Wage Growth**

| | Average wages | | | Wage growth | |
| | 2000 | 2001 | 2002 | 2000-1 | 2001-2 |
| --- | --- | --- | --- | --- | --- |
| Domestic | 174.89 | 183.02 | 186.93 | 5.00% | 3.23% |
| Foreign | 203.77 | 212.33 | 218.32 | 5.01% | 3.34% |
| | | | | | |
| Small (0-49) | 167.68 | 175.62 | 178.75 | 4.79% | 2.70% |
| Medium (50-499) | 186.13 | 194.70 | 199.68 | 5.02% | 3.46% |
| Large (500+) | 185.41 | 193.98 | 199.93 | 5.24% | 3.66% |

Note: The table includes all full-time workers in the private sector, aged 20-65 years. The division of firms into size classes is based on the average number of employees over the year. Average wages are hourly wages in DKK.

**Table 4.3: Worker Flows, by Ownership of the Firm**

*Workers employed in foreign-owned firms*

| Status the following year | 2000 | | 2001 | |
|---|---|---|---|---|
| Same firm, foreign owned | 107,615 | 63.2% | 119,121 | 67.5% |
| Same firm, domestically owned | 18,387 | 10.8% | 7,191 | 4.1% |
| New firm, foreign owned | 11,409 | 6.7% | 13,779 | 7.8% |
| New firm, domestically owned | 18,398 | 10.8% | 20,293 | 11.5% |
| Self-employment | 1,003 | 0.6% | 840 | 0.5% |
| Unemployment/non-employment | 9,677 | 5.7% | 11,695 | 6.6% |
| Public sector | 3,676 | 2.2% | 3,593 | 2.0% |
| Total | 170,165 | 100.0% | 176,512 | 100.0% |

*Workers employed in domestically-owned firms*

| Status the following year | | | | |
|---|---|---|---|---|
| Same firm, domestically owned | 805,641 | 67.8% | 820,971 | 69.5% |
| Same firm, foreign owned | 21,507 | 1.8% | 17,463 | 1.5% |
| New firm, domestically owned | 205,376 | 17.3% | 182,443 | 15.4% |
| New firm, foreign owned | 31,920 | 2.7% | 29,605 | 2.5% |
| Self-employment | 11,556 | 1.0% | 10,758 | 0.9% |
| Unemployment/non-employment | 79,812 | 6.7% | 89,129 | 7.5% |
| Public sector | 32,936 | 2.8% | 30,787 | 2.6% |
| Total | 1,188,748 | 100.0% | 1,181,156 | 100.0% |

Note: The table includes all full-time workers in the private sector, aged 20-65. A firm is considered foreign owned if foreigners ultimately own more than 50% of the firm, and FDI amounts to more than DKK 10 million.

**Table 4.4: Worker Flows, by Firm Size**

*Workers employed in large firms*

| Status the following year | 2000 | | 2001 | |
|---|---|---|---|---|
| Same firm, large | 301,653 | 64.6% | 312,685 | 67.6% |
| Same firm, small | 14,129 | 3.0% | 9,457 | 2.0% |
| New firm, large | 66,526 | 14.2% | 52,905 | 11.4% |
| New firm, small | 37,796 | 8.1% | 37,030 | 8.0% |
| Self-employment | 2,623 | 0.6% | 2,292 | 0.5% |
| Unemployment/non-employment | 29,577 | 6.3% | 33,589 | 7.3% |
| Public sector | 14,968 | 3.2% | 14,512 | 3.1% |
| Total | 467,272 | 100.0% | 462,470 | 100.0% |

*Workers employed in small firms*

| Status the following year | | | | |
|---|---|---|---|---|
| Same firm, small | 661,659 | 69.4% | 668,127 | 69.8% |
| Same firm, large | 9,547 | 1.0% | 7,268 | 0.8% |
| New firm, small | 139,910 | 14.7% | 134,660 | 14.1% |
| New firm, large | 37,152 | 3.9% | 35,623 | 3.7% |
| Self-employment | 10,682 | 1.1% | 9,968 | 1.0% |
| Unemployment/non-employment | 67,485 | 7.1% | 76,385 | 8.0% |
| Public sector | 26,734 | 2.8% | 25,215 | 2.6% |
| Total | 953,169 | 100.0% | 957,246 | 100.0% |

Note: The table includes all full-time workers in the private sector, aged 20-65. A firm is considered larger if the average number of employees over the year is lager than 500.

## Table 5.1: Effects of Current Firm Type on Wage Levels (short panel)

| | Dependent variable: log(hourly wage) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| | OLS | OLS | OLS | OLS | Indv. FE | Indv. FE | Indv. FE | Indv. FE | Firm FE |
| Foreign | 0.134 | 0.110 | 0.101 | 0.070 | 0.020 | 0.013 | 0.011 | 0.012 | -0.006 |
| | (216.47)** | (173.98)** | (154.12)** | (133.02)** | (36.53)** | (24.82)** | (17.84)** | (19.69)** | (3.95)** |
| Log(firm size) | | 0.019 | 0.014 | 0.011 | | 0.012 | 0.012 | 0.011 | -0.021 |
| | | (235.16)** | (146.48)** | (124.55)** | | (99.38)** | (91.09)** | (81.74)** | (22.31)** |
| Exporter | | | 0.052 | 0.048 | | | 0.010 | 0.011 | 0.001 |
| | | | (109.51)** | (107.57)** | | | (24.40)** | (26.26)** | (0.54) |
| Age | | | | 0.041 | | | | | 0.036 |
| | | | | (218.69)** | | | | | (231.62)** |
| $Age^2$ | | | | -0.0005 | | | | | -0.0004 |
| | | | | (211.37)** | | | | | (216.38)** |
| Experience | | | | 0.016 | | | | 0.051 | 0.015 |
| | | | | (147.45)** | | | | (61.19)** | (163.55)** |
| $Experience^2$ | | | | -0.0002 | | | | -0.0011 | -0.0002 |
| | | | | (89.00)** | | | | (185.28)** | (102.39)** |
| Years of education | | | | 0.049 | | | | | 0.045 |
| | | | | (557.64)** | | | | | (565.46)** |
| Female | | | | -0.198 | | | | | -0.187 |
| | | | | (515.89)** | | | | | (474.04)** |
| | | | | | | | | | |
| Time effects | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| Individual effects | | | | | yes | yes | yes | yes | |
| Firm effects | | | | | | | | | yes |
| Industry effects | | | | yes | | | | yes | |
| Regional effects | | | | yes | | | | yes | |
| | | | | | | | | | |
| Observations | 3,584,768 | 3,565,987 | 3,259,986 | 3,215,543 | 3,584,768 | 3,565,987 | 3,259,986 | 3,259,986 | 3,215,543 |
| Number of firms | | | | | | | | | 111,232 |
| Number of individuals | | | | | 1,449,600 | 1,445,909 | 1,341,082 | 1,341,082 | |
| R-squared | 0.02 | 0.03 | 0.03 | 0.35 | 0.09 | 0.09 | 0.09 | 0.10 | 0.26 |

Note: The table is based on a panel from 2000-2002, which includes all full-time workers in the private sector aged 20-65. Foreign=1 for workers employed in firms where foreigners ultimately own more than 50% of the firm, and FDI amounts to more than DKK 10 million. Firm size refers to the average number of full-time employees in the firm within a year. Exporter=1 for workers employed in firms with exports. Robust t statistics are in parentheses. * significant at 5%; ** significant at 1%.

# Table 5.2: Effects of Current Firm Type on Wage Growth

| | Dependent variable: dlog(hourly wages) | | | |
|---|---|---|---|---|
| | (1) OLS | (2) OLS | (3) OLS | (4) Firm FE |
| Foreign | 0.07 (2.05)* | -0.19 (5.69)** | -0.17 (4.94)** | -1.80 (10.66)** |
| Log(firm size) | | 0.21 (42.62)** | 0.26 (47.51)** | -1.80 (15.61)** |
| Age | | | -0.76 (54.21)** | -0.77 (62.87)** |
| Age$^2$ | | | 0.0072 (45.12)** | 0.0075 (52.69)** |
| Experience | | | -0.10 (12.29)** | -0.10 (14.41)** |
| Experience$^2$ | | | 0.0022 (12.28)** | 0.0019 (11.28)** |
| Years of education | | | 0.07 (12.92)** | 0.04 (6.73)** |
| Female | | | 0.38 (14.70)** | 0.15 (5.44)** |
| | | | | |
| Time effects | yes | yes | yes | yes |
| Firm Effects | | | | yes |
| Industry effects | | | yes | |
| Regional Effects | | | yes | |
| Observations | 1,728,255 | 1,723,877 | 1,704,278 | 1,704,278 |
| Number of firms | | | | 86,794 |
| R-squared | 0.00 | 0.00 | 0.02 | 0.02 |

Note: All coefficients are multiplied by 100. The table is based on a panel from 2000-2002, which includes all full-time workers in the private sector aged 20-65. Foreign=1 for workers employed in firms where foreigners ultimately own more than 50% of the firm and FDI amounts to more than DKK 10 million. Firm size refers to the average number of full-time employees in the firm within a year. Robust t statistics are in parentheses. * significant at 5%; ** significant at 1%.

# Table 5.3: Effects of Previous Plant Size on Wage Levels (long panel)

| | Dependent variable: log(hourly wage) | | | | | | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| | OLS | OLS | OLS | OLS | Indv. FE | Indv. FE | Plant FE |
| Log(plant size) | 0.041 | 0.022 | 0.022 | 0.022 | 0.020 | 0.019 | 0.002 |
| | (294.59)** | (167.70)** | (110.71)** | (111.03)** | (117.46)** | (108.87)** | (2.62)** |
| Age | | 0.059 | 0.059 | 0.059 | | 0.070 | 0.050 |
| | | (246.29)** | (245.27)** | (244.98)** | | (78.96)** | (256.76)** |
| Age$^2$ | | -0.0007 | -0.0007 | -0.0007 | | -0.0008 | -0.0006 |
| | | (205.42)** | (204.37)** | (203.70)** | | (238.31)** | (215.08)** |
| Experience | | 0.035 | 0.037 | 0.033 | | 0.025 | 0.028 |
| | | (246.55)** | (211.82)** | (173.76)** | | (120.83)** | (165.70)** |
| Experience$^2$ | | -0.0010 | -0.0011 | -0.0010 | | -0.0007 | -0.0008 |
| | | (205.40)** | (167.10)** | (141.56)** | | (116.88)** | (127.89)** |
| Tenure | | 0.0114 | 0.0043 | -0.0030 | | 0.0021 | 0.0058 |
| | | (61.46)** | (15.81)** | (10.64)** | | (11.48)** | (20.57)** |
| Tenure$^2$ | | -0.0007 | -0.0002 | 0.0001 | | -0.0004 | -0.0002 |
| | | (57.00)** | (13.23)** | (3.53)** | | (33.55)** | (9.22)** |
| Female | | -0.175 | -0.175 | -0.176 | | | -0.157 |
| | | (401.36)** | (400.98)** | (403.36)** | | | (343.82)** |
| Years of education | | 0.043 | 0.043 | 0.043 | | | 0.036 |
| | | (428.51)** | (428.36)** | (424.76)** | | | (348.20)** |
| Large x Experience | | | -0.006 | -0.005 | | | -0.006 |
| | | | (30.25)** | (24.59)** | | | (32.04)** |
| Large x Experience$^2$ | | | 0.0002 | 0.0001 | | | 0.0002 |
| | | | (20.17)** | (14.45)** | | | (22.32)** |
| Large x Tenure | | | 0.014 | 0.015 | | | 0.008 |
| | | | (40.64)** | (43.43)** | | | (24.04)** |
| Large x Tenure$^2$ | | | -0.0009 | -0.0010 | | | -0.0007 |
| | | | (39.08)** | (42.98)** | | | (28.81)** |
| Experience from large plants | | | | 0.014 | | | 0.012 |
| | | | | (67.33)** | | | (64.53)** |
| (Experience from large plants)$^2$ | | | | -0.0005 | | | -0.0003 |
| | | | | (40.78)** | | | |
| Time effects | yes | yes | yes | yes | yes | yes | yes |
| Individual effects | | | | | yes | yes | |
| Firm effects | | | | | | | yes |
| Industry effects | yes | yes | yes | yes | yes | yes | yes |
| Regional effects | yes | yes | yes | yes | yes | yes | yes |
| Observations | 3,553,239 | 3,553,239 | 3,553,239 | 3,553,239 | 3,553,239 | 3,553,239 | 3,553,239 |
| Number of plants | | | | | | | 175,413 |
| Number of individuals | | | | | 570,497 | 570,497 | |
| R-squared | 0.20 | 0.39 | 0.39 | 0.39 | 0.35 | 0.37 | 0.29 |

Note: The table is based on a panel from 1981-2003. The panel includes a 50% sample of all full-time workers in the private sector aged 20-65 who entered the Danish labor market in 1981 or later. Plant size is the total number of employees in the last week of November. Large =1 for workers employed in firms where plant size > 50. Robust t statistics are in parentheses. * significant at 5%; ** significant at 1%.

# Table 5.4: Effects of Previous Firm Type on Earnings of New Self-Employed

| | Dependent variable: log(gross annual earnings) | | | | |
|---|---|---|---|---|---|
| | (1)<br>OLS | (2)<br>OLS | (3)<br>OLS | (4)<br>OLS | (5)<br>OLS |
| $Foreign_{t-1}$ | 0.135<br>(6.11)** | 0.073<br>(3.11)** | -0.044<br>(1.07) | 0.067<br>(2.85)** | -0.031<br>(0.72) |
| Log(firm $size_{t-1}$) | | 0.022<br>(7.44)** | 0.022<br>(7.44)** | 0.016<br>(4.16)** | 0.018<br>(4.51)** |
| Age | 0.003<br>(0.62) | 0.008<br>(1.48) | 0.008<br>(1.48) | 0.008<br>(1.45) | 0.008<br>(1.45) |
| $Age^2$ | 0.00001<br>(0.17) | -0.00004<br>(0.60) | -0.00004<br>(0.60) | -0.00004<br>(0.57) | -0.00004<br>(0.58) |
| Experience | 0.047<br>(14.90)** | 0.046<br>(13.98)** | 0.046<br>(13.84)** | 0.046<br>(13.85)** | 0.046<br>(13.79)** |
| $Experience^2$ | -0.001<br>(11.09)** | -0.001<br>(10.52)** | -0.001<br>(10.49)** | -0.001<br>(10.51)** | -0.001<br>(10.49)** |
| Years of schooling | 0.047<br>(15.41)** | 0.046<br>(14.39)** | 0.046<br>(14.37)** | 0.046<br>(14.38)** | 0.046<br>(14.36)** |
| Female | -0.272<br>(15.39)** | -0.267<br>(14.35)** | -0.268<br>(14.38)** | -0.268<br>(14.37)** | -0.268<br>(14.39)** |
| $Foreign_{t-1}$ x $Tenure_{t-1}$ | | | 0.047<br>(3.05)** | | 0.043<br>(2.59)** |
| $(Foreign_{t-1}$ x $Tenure_{t-1})^2$ | | | -0.002<br>(2.15)* | | -0.002<br>(2.04)* |
| $Large_{t-1}$ x $Tenure_{t-1}$ | | | | 0.012<br>(1.79) | 0.006<br>(0.88) |
| $(Large_{t-1}$ x $Tenure_{t-1})^2$ | | | | -0.0002<br>(0.67) | 0.0000<br>(0.001) |
| Time dummies | yes | yes | yes | yes | yes |
| Industry dummies | yes | yes | yes | yes | yes |
| Regional dummies | yes | yes | yes | yes | yes |
| Observations | 23,125 | 20,183 | 20,183 | 20,183 | 20,183 |
| R-squared | 0.08 | 0.09 | 0.09 | 0.09 | 0.09 |

Note: The table is based on a panel from 2000-2002 which includes all new self-employed (in their first year as self-employed) who were employed as wage-workers in the period prior to self-employment. Only individuals aged between 20 and 65 that operate in the private sector are included. $Foreign_{t-1}$=1 for individuals previously employed in foreign firms (the period prior to business start-up). Firm $size_{t-1}$ refers to size of the firm in the previous job. $Large_{t-1}$=1 for individuals previously employed in firms with more than 500 employees. $Tenure_{t-1}$ is tenure in previous job. Robust t statistics are in parentheses. * significant at 5%; ** significant at 1%.

# AN ESTIMATION PROCEDURE FOR MIXED MARKOV DECISION MODELS OF DISCRETE CHOICES.

BERTEL SCHJERNING

*First Draft, January 2005*

*Revised, August 2007*

ABSTRACT. This paper develops, implements and evaluates an algorithm for estimating a Dynamic Programming Mixed Logit (DPMXL) model. DPMXL is a highly flexible model that obviates many of the limitations of models with additively separable, conditionally independent, extreme value distributed unobservables. It can potentially allow for random taste variation, unrestricted substitution patterns, and correlation in unobserved state variables over time using rich mixtures of random components. I use Chebyshev polynomials to approximate expected value functions over (observed and unobserved) continuous state variables and simulation techniques to evaluate integrals. This strategy has several important spin-offs. First, it helps ameliorating the 'curse of dimensionality'. Second, the approach permits very fast and accurate simulation of likelihood functions. Third, we can easily allow for lots of heterogeneity, greatly expanding the range of models that can be considered. Fourth, the data on continuous state variables does not have to be discretized, mitigating serious problems with approximation errors.

When evaluating the approach, I find that a fifth order polynomial provide sufficiently accurate approximation. The approximation errors, transmitted to conditional choice probabilities, likelihood functions and structural parameter estimates, are practically eliminated at this level of approximation. Finally, I provide a Monte Carlo experiment that highlights the importance of heterogeneity bias in dynamic programming discrete choice models.

Keywords: Structural estimation, dynamic programming, nested fixed point algorithm, unobserved heterogeneity, mixture models, approximation methods

JEL codes: C15, C61, C63.

---

## 1. Introduction

Methods for solving and estimating discrete choice dynamic programming models have been one of the most active research areas over the last two decades[1]. The importance and usefulness of dynamic discrete choice models for individual economic behavior is now well established: First, the structural approach where parameters to be estimated are linked directly to individual objectives or constraints in a theory model enables us to discriminate between competing behavioral models and to draw precise inference about individual economic behavior. In turn, a structurally estimated model can be used for evaluation of policy proposals by means of counterfactual simulations.

However, researchers using these types of models, often face considerable computational burdens to overcome. Therefore model specifications often have to be very parsimonious in terms of state variables. Furthermore, inference in many interesting models are precluded by the requirement of high dimensional integration and for that reason researchers often must restrict attention to models with a very simple - some times unrealistic - error structure.

Rust (1988, 1989) introduced the conditional independence assumption, which defined a class of models where the dimensions of the dynamic programming problem were significantly reduced. As a special case, Rust considered a particular simple class of models, where unobservables are additively separable in utility, independent across alternatives, and extreme value distributed. These assumptions imply that the multidimensional integrals in the choice specific expected value functions and conditional choice probabilities have a simple closed form, thereby saving the cost of multi-dimensional integration over unobserved state variables.

In this paper, I propose an estimation procedure for mixed Markov decision models of discrete choices. The suggested approach builds on the Nested Fixed Point Algorithm developed in Rust (1987, 1988) and ideas from the literature on mixed discrete choice

---

[1]For excellent surveys of the literature on structural estimation of dynamic decision processes, see Eckstein and Wolpin (1989); Pakes (1994); Rust (1994); and Miller (1997). For an update of the recent developments in the literature, see Aguirregabiria and Mira (2007a).

models with simulation (see e.g. Train (2003)). Assuming that only a part of the un-observed state variables are additively separable in utility and multivariate iid extreme value distributed results in a Dynamic Programming Mixed Logit (DPMXL) model.

Chebyshev polynomials are used to approximate expected value functions over continuous variables and Monte Carlo integration is used to evaluate integrals in the expected value functions and to simulate likelihood functions. The use of Chebyshev polynomials has important spin-offs. First of all, it helps ameliorating the 'curse of dimensionality'; the well known exponential rise in computer time and storage requirements as the dimension of the state space increases.[2] Secondly, it permits fast and accurate evaluation of simulated likelihood functions, once value functions are approximated.

DPMXL obviates many of the limitations of Rust's model with additive separable, conditional independent, extreme value distributed unobservables and can potentially allow for i) random taste variation, ii) unrestricted substitution patterns, iii) correlation in unobserved state variables over time, iv) lots of observed and unobserved heterogeneity, and v) arbitrary shape of the distribution of unobserved factors. This greatly expands the range of models that can be considered. I present two examples of different specifications of the model. First, I consider the case of observed and unobserved individual specific heterogeneity, arbitrarily correlated across alternatives. Second, I consider the case of serially correlated unobservables.

I evaluate the method by studying the extent to which approximation errors are transmitted to conditional choice probabilities, likelihood functions and structural parameter estimates. Doing this for various degrees of approximation, I find that a fifth order polynomial provide sufficiently accurate approximations to the expected value function, where approximation errors and bias in ML estimates are practically eliminated.

Finally, to illustrate the DPMXL model with unobserved individual specific heterogeneity, I perform a Monte Carlo experiment where I evaluate the importance of neglected individual specific heterogeneity in dynamic programming models.

---

[2]This idea is not new. The use of flexible functions (such as series of polynomials) has become increasingly popular as a way of ameliorating the curse of dimensionality (see Rust (1996) for a comprehensive treatment and evaluation of these methods and Imai and Keane (2004) and Gamba and Tesser (2006) for recent applications)

The rest of the paper is organized as follows: In section 2, I briefly review some of the existing estimation procedures for dynamic structural models for discrete choices. In section 3, I formulate the assumptions behind the behavioral model of interest and the econometric specification. In section 4, I describe how parametric approximations of the expected value function are implemented in the estimation procedure to obtain maximum likelihood estimates under the conditional independence assumption. In section 5, I present two specifications of the DPMXL model. In section 6, I present some numerical results. First, I evaluate the accuracy of the approximation methods and the consequences for empirical inference. Secondly, the importance of heterogeneity bias in dynamic programming discrete choice models is investigated. Section 7 concludes.

## 2. Existing Estimation Procedures

In this section, I review some of the existing algorithms to solve and estimate dynamic structural models of discrete choices.

Rust (1988) introduced the Nested Fixed Point Algorithm (NFXP), which is one of the first and most widely adopted approaches. This procedure describes how to obtain Maximum likelihood (ML) estimates in dynamic Markov decision models. Rust's NFXP algorithm is a full solution method. To obtain ML estimates an 'outer' hill climbing algorithm searches over the parameter space, while an 'inner' algorithm is used to re-solve a fixed point problem for each evaluation of the likelihood function. NFXP has been extensively used for empirical applications in numerous areas: Optimal replacement of machines (Rust (1987) and Kennet (1994)), and retirement behavior (Rust and Phelan (1997) and Karlstrom, Palme, and Svensson (2004)).

Hotz and Miller (1993) suggested an even simpler estimator called the Conditional Choice Probability (CCP) estimator and provided an application to a model of contraceptive choices.[3] As opposed to NFXP, the CCP estimator does not require econometricians to explicitly solve the fixed point problem. Hotz and Miller show that value functions, characterizing the choice specific expected future utility, can be expressed as a (closed form) function of state variables, structural parameters and conditional choice probabilities. Replacing true conditional choice probabilities by non-parametric estimates, Hotz

---

[3]The CPP estimator has also been applied to models of price and inventory decisions, see e.g. Aguirregabiria (1999).

and Miller obtain a simple closed form expression for conditional choice probabilities, which in turn is used to construct a sample criterion used in estimation.

However, although the CCP estimator is proven to be consistent and asymptotically normal, the computational gain is followed by a cost of efficiency in finite samples. This efficiency cost is likely to grow exponentially with the dimension of the state space, since the non-parametric estimator to used estimate conditional choice probabilities is subject to (an empirical) curse of dimensionality.

Aguirregabiria and Mira (2002) suggested a new iterative nested algorithm, called the Nested Pseudo Likelihood (NPL) algorithm. They utilize the insights from Hotz and Miller (1993) to obtain a representation of the solution to the dynamic programming model in conditional probability space. This allows them to *swap the nesting* of the two algorithms in Rust's nested fixed point algorithm, such that the 'hill climbing' algorithm that maximizes the (pseudo) likelihood over the parameter space is nested in the more computationally expensive fixed point algorithm. Successive iterations thus return a sequence of estimators, which they call K-stage Policy Iteration (PI) estimators.

The PI estimator is shown to nest the CCP estimator (when $K = 1$) and to converge to the ML estimator obtained by Rust's NFXP algorithm (as $K \to \infty$). Hence, by iterating on the NPL algorithm, the researcher can obtain more efficiency at the expense of computational cost as $K$ is increased. Aguirregabiria and Mira (2002) present Monte Carlo evidence showing that NPL provides very precise estimates when $K = 2$ and therefore converges much faster (5 to 15 times) than NFXP.[4]

The techniques developed in Hotz and Miller (1993) and Aguirregabiria and Mira (2002) has proven to be particularly useful in the estimation of dynamic games of incomplete information. (see Aguirregabiria and Mira (2007b) and Aguirregabiria and Mira (2007c) for an applications of the NPL algorithm to firms' entry and exit decisions in oligopoly markets) The NPL algorithm has also been applied to single agent models (see e.g. Sanchez-Mangas (2001)).

---

[4]When evaluating the NPL algorithm Aguirregabiria and Mira (2002) consider a particular simple class of models with unobservables which are extreme value distributed, binary choice and multiplicative separability between state variables and parameters in the utility functions. Moreover, the model considered in the Monte Carlo experiment has only a single state variable.

So why not always use the NPL algorithm? The answer is that computational gains come at the cost of generality. Aguirregabiria and Mira (2002) cite difficulties in dealing with more general specifications. First, the PI and CCP estimators both maintain the conditional independence assumption, and hence they cannot handle serial correlation.[5] Secondly, as mentioned above, computation of choice specific value functions in probability space requires that the mapping from value functions to conditional choice probabilities is inverted. For other distributions than the extreme value distribution, this mapping does not have a closed form in the first place and the mapping must therefore be inverted numerically. This may imply a serious computational cost and it remains to be seen whether CCP and PI estimators will be useful when the iid extreme value assumption is relaxed.

Several authors have tried to deviate from Rust's model with additively separable, conditionally independent, extreme value distributed unobservables. A prominent example is the occupational choice model in Keane and Wolpin (1997). First, unobserved state variables are not additive separeable in utility. Second, the model allows for permanent unobserved heterogeneity in skill endowments using a finite mixture of discrete types. When relaxing the assumptions in this direction, the integrals in the expected value function no longer have closed form. Consequently, this requires high dimensional numerical integration over unobserved state variables at each point in the state space. To estimate the model, Keane and Wolpin (1997) use the simulation and interpolation method developed in Keane and Wolpin (1994).[6]

Recently, Imai, Jain, and Ching (2006) developed a procedure for Bayesian dynamic programming estimation that solves the dynamic programming problem and estimates the parameters at the same time. In contrast to most dynamic structural estimation procedures, their algorithm can be readily be applied to models where observed and

---

[5]The CCP estimator has the additional difficulty, that it must be initialized with nonparametric estimates of conditional choice probabilities. Hence, I applications with e.g. unobserved heterogeneity, consistent nonparametric estimates of choice probabilities either are not available or are very imprecise.

[6]In Keane and Wolpin (1997), Monte Carlo integration is used to simulate expectations of the value function at a subset of the state points. For every time period, they estimate a regression function based on these points and use the fitted regression to interpolate between them.

unobserved heterogeneities are continuous. This is in contrast to, e.g., the approach used in Keane and Wolpin (1997).

## 3. The Behavioral Framework

3.1. **The General Problem.** Consider a decision maker who faces a sequence of repeated choice situations. In each discrete time period, the agent chooses between a finite number of mutually exclusive choice alternatives, $d \in D = \{1, 2, .., J\}$, in order to maximize an infinite sum of expected future payoffs subjectively discounted by the discount factor $\beta \in (0, 1)$. In each time period, the decision maker receives a state- and choice-specific pay-off according to the instantaneous utility function $u(s, d)$, where $s$ denotes the current value of the state variables. Future values of the state variables, $s'$, are uncertain to the agent; and the agents beliefs about future states obeys a (controlled) Markov transition process with transition probability $p(s'|s, d)$.[7]

The optimal choice can be summarized by a stationary decision rule

$$(3.1) \qquad \delta(s) = \arg\max_{d \in D(s)} \left\{ u(s, d) + \beta \int V(s') \, p(ds'|s, d) \right\}$$

where $V(s)$ is the value function.

By Bellman's principle of optimality, $V(s)$ can be found as the unique fixed point of the Bellman equation.

$$(3.2) \qquad V(s) = \max_{d \in D(s)} \left\{ u(s, d) + \beta \int V(s') \, p(ds'|s, d) \right\}$$

From an econometric point of view, we need to introduce an 'error term', since no structural model will ever perfectly predict observed choice behavior. One reason for this is that the econometrician only observes a subset of the state variables $s$. To discriminate between observable and unobservable variables, I assume that the vector of state variables $s$ partitions into two components $s = (x, \varepsilon)$ where $x$ are state variables observable for us as researchers and $\varepsilon$ are unobserved state variables. Assume further that $\varepsilon$ have $J$ choice specific components $\varepsilon(d) : d \in D$ with support on the real line. With this notation, the

---

[7]In the class of models considered here, the value function does not depend on time per se, resulting in a stationary decision rule. We can therefore omit time subscripts and let $s'$ denote the state variable in the following period.

optimal value function $V(x, \varepsilon)$, can be expressed as the unique solution to the Bellman equation

$$(3.3) \qquad V(x, \varepsilon) = \max_{d \in D(s)} \{u(x, \varepsilon, d) + \beta EV(x, \varepsilon, i)\}$$

where $EV(x, \varepsilon, d) \equiv \int_y \int_\eta V(y, \eta) \, p(dy, d\eta | x, \varepsilon, d)$ denote the expected value function.

Empirical implementation of this model attempts to uncover the structural parameters, $\theta$, of the agent's optimization problem given panel data on observable state variables $x_{nt}$, the discrete choice variable $d_{nt}$ (and possibly also permanent conditioning variables $z_n$). The parameters of interest are the discount factor $\beta$, the parameters that index the utility function, $u(x, \varepsilon, d)$, and the transition probability of the state variable process, $p(s'|s, d)$. Estimates of these parameters can be obtained by maximizing the sample log likelihood $l(\theta) = \sum_{n=1}^{N} l_n(\theta)$, where $l_n(\theta)$ is the log likelihood contribution of individual $n$

$$\begin{aligned} l_n(\theta) &= \log \Pr(d_{nt}, x_{nt} : t = 1, 2...T_n | \theta) \\ &= \log \Pr(d_{nt} = \delta(x_{nt}, \varepsilon_{nt}, \theta), x_{nt} : t = 1, 2...T_n | \theta) \end{aligned}$$

To evaluate the likelihood function for a given value of the structural parameters, $\theta$, we must solve for the decision rule $\delta(x_{nt}, \varepsilon_{nt}, \theta)$. The conventional procedure to Maximum Likelihood (ML) estimation of dynamic programming models is to use Rust's Nested Fixed Point Algorithm: The Algorithm consist of two steps. First an *'inner'* algorithm numerically computes the solution to the dynamic programming problem and subsequently evaluates the likelihood function. Second an *'outer'* hill climbing algorithm searches over different values of $\theta$ to maximize the implied likelihood function. Thus, the inner fixed point algorithm is *nested* in the outer algorithm.

3.2. **Rust's Assumptions.** As it stands, we face a considerable computational burden when repeatedly solving for the fixed point in the Bellman equation (3.3). In many problems the state variables are continuously distributed with unbounded support. This introduces serious dimensionality problems and the requirement of high dimensional integration when evaluating $EV(x, \varepsilon, d)$. Furthermore, computation of conditional choice probabilities $P(d|x)$ used for estimation and inference, requires numerical integration over the unknown function $EV(x, \varepsilon, d)$ with respect to $\varepsilon$. A major contribution of Rust (1987)

and Rust (1988) was to introduce the conditional independence assumption that allows us to circumvent some of these problems.

**Assumption 1.** *(Conditional Independence Assumption (CI)): The transition density of the controlled process $\{x, \varepsilon\}$ factors as*

$$(3.4) \qquad\qquad p(x', \varepsilon'|x, \varepsilon, d) = g(\varepsilon'|x')f(x'|x, d)$$

Under assumption 1, the unobserved state variables, $\varepsilon$, can be integrated out from the Bellman equation (3.3) such that the unknown function, $EV$, no longer depends on $\varepsilon$.

$$(3.5) \qquad EV(x, d) = \int_y \int_\varepsilon \max_{j \in D(x,\varepsilon)} \left[ u(y, \varepsilon(j), j) + \beta EV(y, j) \right] g(d\varepsilon|x) f(dy|x, d)$$

Rust (1988) proved that the operator $\Gamma(\cdot)$ defined by the right hand side of (3.5) is a separate contraction mapping, where $EV$ is the unique fixed point: $EV = \Gamma(EV)$. This eliminates the requirement of computing the fixed point $V$ on the full state space and allow us to restrict attention to a much reduced state space spanned by $D(x, \varepsilon)$ and the domain of $x$. Furthermore, since $EV(x, d)$ is not a function of $\varepsilon$, conditional choice probabilities $P(d|x)$ can be computed without numerical integration over the unknown function $EV$.

(3.6)
$$P(d|x) = \int_\varepsilon I \left\{ d = \arg \max_{j \in D(x,\varepsilon)} \left[ u(x, \varepsilon(j), j) + \beta \int_y EV(y, j) f(dy|x, d) \right] \right\} g(d\varepsilon|x)$$

Under assumption 1, the observable state vector $x_{nt}$ is a sufficient statistic for the current choice. Therefore the log likelihood contribution for individual $n$ has a particular simple form

$$\begin{aligned} l_n(\theta) &= \sum_{t=1}^{T_n} \ln \left[ P(d_{nt}|x_{nt}; \theta) \right] + \sum_{t=1}^{T_n} \ln \left[ f(x_{nt}|x_{nt-1}, d_{nt-1}; \theta_f) \right] \\ &= l_n^1(\theta) + l_n^2(\theta_f) \end{aligned}$$

where $l_n^1$ and $l_n^2$ are the *partial* log-likelihood contribution for a single individual $n$ relating to the conditional distribution of $d_{nt}$ given $x_{nt}$ and the evolution in $x_{nt}$ given $x_{nt-1}$.

Yet a another advantage of assumption 1, is that the model can be consistently estimated by a two step procedure: The log-likelihood is additively separable in $l_n^1$ and $l_n^2$ and the transition equation does not dependent on objective function parameters, once

we have conditioned on the choice of the decision maker. Thus, we can obtain consistent partial maximum likelihood estimates of $\theta_f$ by maximizing the second part of the likelihood function, with respect to these parameters. If we denote these partial estimates, $\hat{\theta}_f^p$, consistent ML estimates of the objective function parameters, $\theta_u$ and the parameters, $\theta_g$ of the distribution of the unobserved state variables $\varepsilon$, can be obtained by maximizing $l^1\left(\hat{\theta}_f^p, \theta_u, \theta_g\right)$ with respect to $\theta_u$ and $\theta_g$. Since the Hessian of the likelihood function is block diagonal under assumption 1 we will not have to correct second step standard errors used for inference.

As a special case, Rust considered a particular simple model, where i) the pay-of function $u(x, \varepsilon(d), d)$ is additively separable in the unobserved state variables, $\varepsilon(d)$ and ii) the unobserved state variables are independent across alternatives and extreme value distributed.

**Assumption 2.** *(Additivity (AS)): The pay-off function $u(s, d)$ is additively separable in $\varepsilon(d)$*

$$u(s, d) = \tilde{u}(x, d) + \varepsilon(d)$$

**Assumption 3.** *(Extreme value assumption): The unobserved state variables, $\varepsilon$ are assumed to be multivariate iid. extreme value distributed*

These two assumptions implies that the expectation operator with respect to $\varepsilon$ in the choice specific expected value functions and the conditional choice probabilities have a particular simple closed form. Consequently, the contraction mapping $\Gamma$ defined by (3.5) reduces to

$$(3.7) \qquad EV(x, d) = \int_y \ln\left[\sum_{j \in D(y)} \exp\left[\tilde{u}(y, j) + \beta EV(y, j)\right]\right] f(dy|x, d)$$

Moreover, the conditional choice probabilities are multinomial logistic

$$(3.8) \qquad P(d|x) = \frac{\exp\left\{\tilde{u}(x, d) + \beta EV(x, d)\right\}}{\sum_{j \in D(x)}\left\{\tilde{u}(x, j) + \beta EV(x, j)\right\}}$$

This saves the cost of multi-dimensional integration over unobserved state variables.

Rust's assumptions put severe restrictions on the behavioral model. First, assumption 1, effectively assumes that all time dependence in $\varepsilon$ is transmitted entirely through the state variable $x$. This for example precludes the presence of serial dependence in $\varepsilon$. Second,

assumption 2 implies that marginal utility of observable state variables does not depend on unobservables. Third, assumption 3 restricts substitution patterns in the model, since $\varepsilon$ is assumed independent across choice alternatives. In a choice situation with several similar alternatives this assumption may well be violated.

Several authors have estimated models that depart form these assumptions. One prominent example is the occupational choice model of Keane and Wolpin (1997). They use the interpolation and simulation method developed in Keane and Wolpin (1994) to allow for non-additive chocks, correlation across alternatives and permanent unobserved heterogeneity (using a finite mixture). Stinebrickner (2000) estimate a model of teachers' occupational decisions allowing for serial dependent unobservables in the wage equation. A comprehensive discussion of the problems that are encountered when dynamic discrete choice models are specified with continuous, serially correlated state variables is given here.

## 4. Maximum Likelihood Estimation

In this section, I describe the estimation procedure to obtain maximum likelihood estimates under assumption 1. It builds on Rust's Nested Fixed Point Algorithm, on the use of parametric approximations of the expected value function (see e.g. Rust, Hall, Benítez-Silva, Hitsch, and Pauletto (2005)), and ideas from the literature on mixed discrete choice models with simulation (see Train (2003)). In particular, I use Chebyshev polynomials to approximate expected value functions over the continuous state variables and Monte Carlo simulation to integrate out unobserved state variables.

Under assumption 1, solving the model is equivalent to finding a fixed point of the functional equation (3.5). However, when operating numerically on a computer, we can only hope to evaluate the unknown function $EV$ in a finite number of points. Hence, if the domain for any of the state variables is continuous, the value function must be represented in a finite number of grid points to make the evaluation operational on the computer. Traditionally, two approaches are taken to ameliorate the consequences of this inherent curse of dimensionality: a smooth or a discrete approximation.

Usually, the vector of observable state variables is assumed to have a discrete and finite support and thus continuous variables are approximated with a finite number of grid

points, such that $s \in S = \{s_1, ..., s_m\}$. This effectively avoids the dimension of the fixed point $EV$ to be infinitely large and due to the finite domain of $s$, the integral in (3.5) can be replaced by a sum of conditional value functions weighted by the probability of the outcome of the state variable.

The discrete approximation works well for small dimensional problems: when the number of grid points $m$ increases, the solution of the discrete approximation approaches the fixed point for the original continuous problem. However, with $Q$ state variables, a discrete approximation with $m_q$ grid points for the $q'th$ dimension in the state space requires evaluations of the value function in $m = \prod_{q=1}^{Q} m_q$ points and computation of a $m \times m$ transition probability matrix for the discretized state variables. Therefore, since the number of grid points needed to obtain an accurate solution often is fairly large for continuous variables, this approach becomes impractical even for a small number of state variables.

A considerable amount of research has been carried out to increase the speed and accuracy of the numerical techniques used. One line of research has been to use smooth approximations of the value function, which can be represented by simple flexible parametric functions based on evaluation in a finite number of grid points, $m_q$. Current approximation techniques include orthogonal polynomials, splines, neural networks among others[8]. Although this line of thought is not able to break the curse of dimensionality, it adds significantly to speeding up the process of finding an accurate solution (given the dimension of the state space). Although the size of the fixed point problem is still exponentially increasing in the number of state variables, $Q$, it is usually much smaller in the case where we use a smooth approximation. As we shall see later, choosing $m_q \geq 6$ gives a solution so accurate that increasing the number of grid points does not affect expected precision of the model solution, ML estimates and the predictions of the model.

4.1. **The Chebyshev Approximated Expected Value Function.** I use orthogonal *Chebyshev Polynomials* to approximate the expected value function using a finite number of nodes $m_q$ in each of the $Q$ dimensions of the state space. This particular approximation method, was chosen for its ability to fit smooth functions, its efficiency in spanning the

---

[8]See Judd (1998), Rust (1994), Santos (1999) and Rust, Hall, Benítez-Silva, Hitsch, and Pauletto (2005) for comprehensive surveys of these methods.

state space (guaranteed by the orthogonal basis), and its ability to fit functions defined over multiple dimensional state spaces, and most important its simplicity. As we shall see in the following sections, this choice has several important spin-offs that have not yet been fully utilized.

Let $T$ denote a full set of Chebyshev polynomials up to order $n$, evaluated in a set of Chebyshev nodes, $x_i$, $i = 1, .., m$. Let further $\alpha_{EV(x,d)}$ denote the choice specific set of Chebyshev coefficients, such that the polynomial, $T\alpha_{EV(x,d)}$ approximates the expected value function, $EV$, evaluated in these nodes. Finally, let $T(x)$ denote a matrix of Chebyshev polynomials evaluated at $x$. If we replace the unknown continuous expected value function $EV$ in (3.5) with its Chebyshev approximated counterpart, $\widehat{EV}(x,d) = T(x)\alpha_{\widehat{EV}(x,d)}$ and denote the implied approximated contraction operator by $\hat{\Gamma}$ we can write

$$(4.1) \quad \hat{\Gamma}\left(\widehat{EV}\right)(x,d) = \int_y \int_\varepsilon \max_{j \in D(x,\varepsilon)} \left[u(y, \varepsilon(j), j) + \beta T(y)\alpha_{\widehat{EV}(x,d)}\right] g(d\varepsilon|x) f(dy|x,d)$$

Since polynomials are linear in parameters, the coefficients can be expressed as least squares coefficients of $EV(x,d)$ on $T$. That is $\alpha_{\widehat{EV}(x,d)} = P_T EV(x,d)$ where $P_T = (T'T)^{-1} T'$ is referred to as the Chebyshev Projection matrix. We can then express $\hat{\Gamma}$ as

$$(4.2)$$
$$\hat{\Gamma}\left(\widehat{EV}\right)(x,d) = \int_y \int_\varepsilon \max_{j \in D(x,\varepsilon)} \left\{ \left[u(y, \varepsilon(j), j) + \beta T(y) P_T \widehat{EV}(x,d)\right] g(d\varepsilon|x) \right\} f(dy|x,d)$$

Note that even though the size of the fixed point problem has become finite due to finite number of Chebyshev nodes, we can now evaluate the expected value function in any point in the state space. This has one important spin-off:. Since we can evaluate the expected value function in any point of the state space, Monte Carlo integration is particularly well suited for evaluating the expectation. We can write the simulated counterpart of (4.2) as

$$(4.3) \qquad \hat{\Gamma}\left(\widehat{EV}\right)(x,d) = \frac{1}{R}\sum_{r=1}^{R} \max_{j \in D(x,\varepsilon)} \left[u(x'_r, \varepsilon'_r(j), j) + \beta T(x'_r) P_T \widehat{EV}(x,j)\right]$$

where $\varepsilon'_r$ and $x^r_j$ are draws from the distribution for the state variable processes $g(\varepsilon'|x')$ and $f(x'|x,d)$, respectively.

Under assumption 1, the expected value function $EV$, summarizing the solution of the model, can be approximated by the fixed-point of the contraction mapping $\hat{\Gamma}\left(\widehat{EV}\right)(x,d) =$

$\widehat{EV}$ defined by the functional equation (4.3) and the simulated conditional choice probabilities, $\hat{P}(d|x)$, are given by

$$(4.4) \qquad \hat{P}(d|x) = \frac{1}{R} \sum_{r=1}^{R} I \left\{ d = \arg \max_{j \in D(s)} \left[ u(x, \varepsilon_r(j), j) + \beta \widehat{EV}(x, j) \right] \right\}$$

In the following section, I describe the solution algorithm used to find the fixed point of the functional equation (4.3). It is instructive to go a bit into details to make it transparent where the use of Chebyshev polynomials gives computational gains.

4.2. **The Fixed Point Algorithm.** There are several ways to compute the fixed point $\widehat{EV}$. In Rust's NFXP software, a combination of successive contraction iterations and the Newton-Kantorowich algorithm is used.

For a sufficiently precise Chebyshev approximation, the contraction mapping property for the operator $\Gamma$ carries over to the approximate operator $\hat{\Gamma}$.[9] This allows us to obtain the unique fixed point $\widehat{EV}$ by successive iterations on the contraction mapping, defined by

$$(4.5) \qquad\qquad \widehat{EV}_k = \Gamma\left(\widehat{EV}_{k-1}\right) = \Gamma^k\left(\widehat{EV}_0\right)$$

The idea behind the Newton-Kantorowich algorithm is to convert the problem of finding a fixed point of the contraction mapping $EV = \Gamma(EV)$ to finding a fixed point for the nonlinear operator $F = I - \Gamma$, since the contraction mapping implies $F(EV) = 0$. Taking a first order Taylor approximation of $[I - \Gamma](EV_{k+1}) = 0$ around the current fixed point, $EV_k = \Gamma(EV_k)$

$$0 = [I - \Gamma](EV_{k+1}) \; \tilde{} \; [I - \Gamma](EV_k) + [I - \Gamma'](EV_{k+1} - EV_k)$$

Solving this equation with respect to $EV_{k+1}$ gives the Newton-Kantorowich iteration

$$\begin{aligned} EV_{k+1} &= EV_k - (I - \Gamma')^{-1}(I - \Gamma)EV_k \\ &= EV_k - (I - \Gamma')^{-1}(EV_k - EV_{k+1}) \end{aligned}$$

Computation of the Newton-Kantorowich iteration requires computation of the Fréchet derivative of the contraction mapping $\Gamma$. For the parametrically approximated operator

---

[9]If the $EV$ is sufficiently smooth, the approximation error can be made arbitrarily small by increasing the degree of the approximating polynomial.

$\hat{\Gamma}\left(\widehat{EV}\right)(d,x)$, $\hat{\Gamma}'$ takes the form of a $J \cdot m \times J \cdot m$ matrix of simulated partial derivatives, where the $l,k$'th block equals a $m \times m$ matrix of derivatives of contraction mapping $\hat{\Gamma}\left(\widehat{EV}(x,d=l)\right)$ differentiated with respect to $EV(x,d=k)'$. It can be shown that the $l,k$'th block can be formulated as

$$(4.6) \qquad \frac{\partial \tilde{\Gamma}\left(\widehat{EV}\right)(x,l)}{\partial \widehat{EV}(x,k)'} = \beta \frac{1}{R} \sum_{r=1}^{R} \hat{P}\left(k|x_r'\right) T\left(x_r'\right) P_T$$

where $\hat{P}\left(k|x_r'\right)$ is the approximated choice probability conditional on a draw $x_r'$ from $f(x'|x,l)$ and based on the value $\widehat{EV}$ evaluated at the Chebyshev nodes.

While contraction iterations guarantee convergence due to the contraction mapping property, they slow down when the approximation errors $\left\|\widehat{EV}_k - \widehat{EV}\right\|$ become small (in particular for $\beta$ close to 1). In contrast, Newton-Kantorowich iterations are not guaranteed to converge, but converge in a quadratic rate in the neighborhood of the solution. The resulting fixed point algorithm, known as the *poly-algorithm*, combines these two algorithms in order to balance robustness versus speed of convergence. That is, it uses contraction iterations until the approximated value function gets within the domain of attraction and then switches to Newton-Kantorowich iterations to obtain fast convergence to the solution.

It is useful to highlight a number of computational shortcuts. First, the nodes for the Chebyshev approximation are fixed for each contraction iteration (and Newton-Kantorowich). Therefore, once the Chebyshev projection matrix $P_T = (T'T)^{-1} T'$ has been computed we will not have to compute it again. Moreover, since Chebyshev polynomials are orthogonal, $T'T$ is a diagonal matrix which is very fast to invert. In fact, $T'T$ can be expressed analytically: The $q'th$ block in $T'T$ referring to the $q'th$ continuous variable is $T_q'T_q = diag\left(\prod_{i=1}^{q} m^i, \prod_{i=1}^{q} m^i/2, .., \prod_{i=1}^{q} m^i/2\right)$. Hence, we only need to store a couple of scalars on the computer, rather than computing (and inverting) an $m \times m$ dimensional matrix.

Secondly, if also the same draws from the state variable process are used at each iteration, the Chebyshev Polynomial evaluated at these draws, $T(x_r')$, are constant too. Therefore, it is also sufficient to compute the matrix product $T(x_r') P_T$ once. Hence, repeated multiplication of an $Rm \times m$ matrix with an $m \times m$ matrix is circumvented when solving the model. This feature really speeds up the algorithm. Moreover, if the

parameters of the variable process are estimated using partial ML, $T\left(x_r'\right)P_T$ does not have to be updated during the second step of estimation.

4.3. **Maximizing the Likelihood.** In Rust's NFXP algorithm, BHHH is used to maximize the likelihood function. This algorithm is convenient for two reasons. First, since BHHH approximates the Hessian by the outer product of the gradient, it avoids numerical calculation of second derivatives of the likelihood function. Moreover, since the outer product of the gradient by construction is positive semi-definite, the BHHH algorithm will always move in the direction of the gradient - even in convex areas of the likelihood function.

A BHHH iteration is defined as

$$\theta_{k+1} = \theta_k + \left(\sum_{n=1}^{N} \nabla_\theta l_n\left(\theta_k\right) \nabla_\theta l_n\left(\theta_k\right)'\right)\left(\sum_{n=1}^{N} \nabla_\theta l_n\left(\theta_k\right)\right)$$

where $\nabla_\theta l_n\left(\theta_k\right)$ is the gradient of the log likelihood contribution for individual $n$ evaluated at current the value of the parameters, $\theta_k$. The likelihood function for the state variable process, $l_n^2\left(\theta_f\right)$, does not depend on the value function parameters and can therefore usually be expressed in closed form.

However, the derivative of $l_n^1\left(\theta\right)$ is a function of the conditional choice probabilities $P\left(d|x\right)$, simulated by $\hat{P}\left(d|x\right)$. As it stands in equation (4.4), the conditional choice probabilities are simulated by a crude accept-reject (AR) simulator. This raises two numerical issues. First, $\hat{P}\left(d|x\right)$ can be zero for any finite number of draws $R$. That is, if none of the $R$ draws of $\varepsilon_r$ result in alternative $d$ being the optimal choice for a given value of $x$, $\hat{P}\left(d|x\right)$ can actually be zero. This is problematic when estimating the model with Maximum Likelihood, because if $\hat{P}\left(d|x\right) = 0$, then $\log \hat{P}\left(d|x\right)$ is undefined.

Secondly, the choice probability is a step function in the structural parameters, $\theta$. Since simulated probabilities are not smooth in the parameters, the likelihood function is nondifferentiable. This makes gradient based estimation infeasible and derivation of standard errors nonstandard.

For some modeling choices it is actually possible to obtain smooth choice probabilities. One obvious example is Rust's model with additive separable, conditionally independent, extreme value distributed unobservables. As we have seen, this result in closed-form logit probabilities. Another example is when $\varepsilon$ is multivariate normal, the model reduces to a

dynamic programming version of the multinomial probit model. In this case, the GHK (Geweke-Hajivassiliou-Keane) recursive simulator can be used to obtain smooth choice probabilities.

One way to mitigate the difficulties with the crude AR simulator is to replace the indicator function in equation (4.4) with a smooth, strictly positive function. In the context of static multinomial probits, McFadden (1989) suggested the use of a logit kernel as smoother. The resulting simulator is referred to as the *logit-smoothed AR simulator* and the simulated probabilities have a form similar to the mixed logit model (Train (2003)). In the next section, I discuss how this simulator can be utilized in the present context.

## 5. A DYNAMIC PROGRAMMING MIXED LOGIT MODEL

The logit-smoothed AR simulator used in the present context can be seen as a Dynamic Programming Mixed Logit (DPMXL) model. That is, a model where conditional choice probabilities are the integrals of standard logit probabilities. This model can be obtained by assuming that part of the unobserved state variables are additively separable in utility and multivariate iid extreme value distributed.

DPMXL is a highly flexible model that obviates many of the limitations of Rust's model with additively separable, conditionally independent, extreme value distributed unobservables. It can potentially allow for random taste variation, unrestricted substitution patterns, and correlation in unobserved state variables over time. Moreover, the distribution of unobservables is not restricted to, e.g., the normal distribution as in the dynamic programming Probit. In fact, McFadden and Train (2000) show that any discrete choice model derived from random utility maximization has choice probabilities that can be approximated as closely as one pleases by a mixed logit model.

In the following I present two examples of how different specifications of this model can be used to allow for more flexibility. In example 1, I consider the case of individual specific heterogeneity - observed as well as unobserved. In example 2, I consider the case of serially correlated unobservables.[10] Both cases allow for correlation across alternatives in unobserved state variables.

---

[10]In the appendix, I describe in more detail how to obtain the simulated likelihood and corresponding gradients.

5.1. **Example 1: Individual Specific Heterogeneity.** Obviously, heterogeneity can both enter in utility and in transition rules. For illustrative purpose, I first focus on how to implement individual specific heterogeneity only in utility, and then briefly discuss the issues involved when implementing heterogeneity in the transition rules.

Assume that the unobserved state variables, $\varepsilon$ partitions into two components $\varepsilon = \{\mu, \eta\}$, where both $\mu$ and $\eta$ have $J$ choice specific components $\{\mu(d), \eta(d) : d \in D\}$ with support on the real line. $\eta(d)$ is assumed to be multivariate iid. extreme value distributed and $\mu(d)$ is a permanent (agent specific) component randomly distributed in the population according to the (mixing) distribution $g(\mu|x_1, z)$, where $z$ are observable agent specific components and $x_1$ is the initial observation of $x$. The introduction of individual specific heterogeneity, violates assumption 1, since $x$ is no longer a sufficient static for $\varepsilon$. However, conditional on $\mu$, $\varepsilon$ is conditionally independent. Moreover, since the additional unobserved factors introduced in the model is due to heterogeneity rather than uncertainty, the formulation of the dynamic optimization problem faced by a single decision maker, *remains unchanged given* $\mu$.

Assume further that, utility is additively separable in $\eta(d)$, such that the utility factors as $u(x, \varepsilon(d), d) = \tilde{u}(x, \mu(d), d) + \lambda \eta(d)$, where $\lambda$ is a (smoothing) weight on $\eta(d)$ chosen by the researcher. I discuss the choice of $\lambda$ later.

$\mu$ can be given many interpretations. For example $\mu$ can be a particular coefficient in the model which we want to be heterogenous among agents. Alternatively, $\mu$ can be seen as a choice specific error component arbitrarily correlated over the alternatives. In both cases, lots of heterogeneity can be obtained by letting $\mu$ be a function of observed individual characteristics, $z$, initial conditions, $x_1$, and appropriate mixtures of (discrete or continuous) random components. This allows the distribution of $\mu$ to have almost any shape.

Under the assumptions made so far, the choice probabilities conditional on $x$ and $\mu$ are multinomial logistic

$$(5.1) \qquad P(d|x, \mu) = \frac{\exp\{[\tilde{u}(x, \mu(d), d) + \beta EV(x, \mu(d), d)]/\lambda\}}{\sum_{j \in D(x)} \exp\{[\tilde{u}(x, \mu(j), j) + \beta EV(x, \mu(j), j)]/\lambda\}}$$

where $EV$ can be computed as the unique fixed point to the contraction mapping

$$(5.2) \quad \Gamma(EV)(x, \mu, d) = \int_y \ln \left[ \sum_{j \in D(y)} \exp \left[ \frac{\tilde{u}(y, \mu(j), j) + \beta EV(y, \mu(j), j)}{\lambda} \right] \right] f(dy|x, d)$$

To solve the model, $EV$ is approximated with Chebyshev polynomials over $x$ and $\mu$ for each value of $d$.

(5.3)
$$\hat{\Gamma}\left(\widehat{EV}\right)(x, \mu, d) = \frac{1}{R} \sum_{r=1}^{R} \ln \sum_{j \in D(y)} \exp \left[ \frac{\tilde{u}(x_r', \mu(j), j) + \beta T(x_r', \mu(d)) P_T \widehat{EV}(x, \mu(j), j)}{\lambda} \right]$$

The expected value function $EV$ can now be approximated by the fixed point of the contraction mapping $\hat{\Gamma}\left(\widehat{EV}\right)(x, \mu, d) = \widehat{EV}$ defined by the functional equation (5.3). Again, we obtain the fixed point using a combination of contraction iterations and Newton-Kantorowich iterations.

Given the solution of the approximated value function $\widehat{EV}$, we can express the choice probabilities conditional on $\mu$ evaluated at the data

(5.4)
$$\hat{P}(d_{nt}|x_{nt}, \mu) = \frac{1}{R} \sum_{r=1}^{R} \frac{\exp\left\{ \left[ \tilde{u}(x_{nt}, \mu(d_{nt}), d_{nt}) + \beta T(x_{nt}, \mu(d_{nt})) p_T \widehat{EV}(x, \mu, d) \right] / \lambda \right\}}{\sum_{j \in D(x)} \exp\left\{ \left[ \tilde{u}(x_{nt}, \mu(j), j) + \beta T(x_{nt}, \mu(j)) p_T \widehat{EV}(x, \mu, d) \right] / \lambda \right\}}$$

The structural parameters consist of $\theta = \{\beta, \theta_u, \theta_f\}$, and the parameters, $\theta_g$, of the mixing distribution, $g(\mu|x_1, z, \theta_g)$. We can now express the likelihood for individual $n$ conditional on $\mu_n$

$$\begin{aligned} L_i(\theta, \mu_n) &= \prod_{t=1}^{T_n} \left[ \hat{P}(d_{nt}|x_{nt}, \mu_n; \theta) \right] \prod_{t=1}^{T_n} \left[ f(x_{nt}|x_{nt-1}, d_{nt-1}; \theta_f) \right] \\ &= L_i^1(\theta, \mu_n) + L_i^2(\theta_f) \end{aligned}$$

Note that likelihood still factors into two components similar to Rust's model. However, since $\mu(d)$ varies randomly in the population, we will have to integrate it out using simulation. The use of approximation methods is particularly convenient for models with lots of heterogeneity. Once we have solved for the expected value function, simulation of the likelihood is very fast, since we then have a closed form for $\hat{P}(d_{nt}|x_{nt}, \mu_n)$.[11] The

---

[11]Alternatively, the model can be repeatedly solved for given values of $\mu_n^s$. This does not multiply the problem by the number of draws, since small changes in $\mu_n^s$ has a small effect on $EV$. Thus if the

simulated likelihood is

$$(5.5) \qquad l_i(\theta) = \ln \frac{1}{S} \sum_{s=1}^{S} L_i^1(\theta, \mu_n^s) + \ln L_i^2(\theta_f)$$

where $\mu_n^s$ is a simulation draw from $g(\mu|x_1, z, \theta_g)$.

The likelihood is still additively separable when permanent unobserved factors only enter in the utility function. The state variable process can therefore still be consistently estimated by maximizing the partial likelihood $\ln L_i^2(\theta_f)$. However, if we want to allow for heterogeneity in the transition rule for $x$, the likelihood function is no longer additively separable and structural estimation must be carried out using Full Information Maximum Likelihood (FIML). Moreover, when approximating the expected value function, the integral over the observable state variable, $x$, has to be computed conditional on the permanent component.

## 5.2. Example 2: Serially Correlated Unobservables.

Serial correlation in unobserved state variables has important implications for behavior. If a shock has a persistent effect, not only current utility is altered when the shock occurs, also the expected future value. In the previous example the sums of permanent and transitory unobserved factors, $\mu_n + \eta_{nt}$ were in fact serially correlated. However, with this formulation, time dependence is transmitted entirely through the permanent component, $\mu_n$. Hence, from the agents perspective, there is no serial correlation. In the following, I discuss how to implement serial correlation of unobserved state variables in the DPMXL framework.

As before, we assume that the unobserved state variables, $\varepsilon$, partitions into two components, $\varepsilon(d) = \{\mu(d), \eta(d)\}$ where $\eta(d)$ is assumed to be multivariate iid. extreme value distributed. We also assume that utility is additively separable in $\eta(d)$. However, $\mu(d)$ is no longer a permanent component, but varies over time, is subject to uncertainty and serially correlated. We also allow $\mu$ to be arbitrarily correlated with observable state variables $x$. Hence, the (mixing) distribution is $g(\mu'|\mu, x')$. Since $\mu$ is uncertain from the agents perspective, the serial correlation in $\mu$ alters the dynamic optimization problem faced by the decision maker, *even conditional on $\mu$*.

---

simulation draws are properly sorted, solution will typically require a single Newton-Kantorowich step. However, this will not work if $\mu$ is serially correlated as in example 2.

Again, if we condition on $x$ and $\mu$, choice probabilities, $P(d|x, \mu)$, are multinomial logistic and therefore similar to the expression in equation (5.1). However, since $\mu$ is now uncertain to the agent, the expected value function must be integrated over $\mu$ to form expectations about future utility. Hence, $EV$ is now determined as the unique fixed point to the contraction mapping

(5.6)

$$\Gamma(EV)(x, \mu, d) = \int_w \int_y \ln \left[ \sum_{j \in D(y)} \exp \left[ \frac{\tilde{u}(y, w(j), j) + \beta EV(y, w(j), j)}{\lambda} \right] \right] g(dw|\mu, dy) f(dy|x, d)$$

As before, we approximate $EV$ with Chebyshev polynomials over $x$ and $\mu$ and integrate over $\mu$ and $x$ using Monte Carlo integration.

(5.7)

$$\hat{\Gamma}\left(\widehat{EV}\right)(x, \mu, d) = \frac{1}{R} \sum_{r=1}^{R} \ln \sum_{j \in D(y)} \exp \left[ \frac{\tilde{u}(x'_r, \mu'_r(j), j) + \beta T(x'_r, \mu'_r(j)) P_T \widehat{EV}(x, \mu, d)}{\lambda} \right]$$

where $\mu'_r(j)$ is the $j'th$ component of a draw from $g(\mu'|\mu, x')$. Given the solution of $\widehat{EV}$, there is conceptually no difference to the example above. Aside from the fact that $\mu$ has to be sampled from a different distribution, everything remains unchanged. The expression for choice probabilities conditional on $\mu$, $\hat{P}(d_{nt}|x_{nt}, \mu)$, is equal to the expression in equation (5.4). The likelihood function is identical to the expression in equation (5.5), except from $\mu$ being drawn from the distribution $g(\mu'|\mu, x)$.

Note that the current formulation of $\mu$ allows for both first order serial correlation and contemporary correlation across alternatives. Hence, $\mu(i)_{nt}$ can be correlated with $\mu(i)_{nt-1}$, $\mu(j)_{nt}$, and $\mu(j)_{nt-1}$. Moreover, we can allow for a great deal of flexibility in the shape of $g(\mu'|\mu, x')$. If for example $\mu$ follows an $AR(1)$ process $\mu_{nt} = \rho\mu_{nt-1} + \epsilon_{nt}$, we can specify $\epsilon_{nt}$ as mixtures of several multivariate distributions.

5.3. **Identification and the choice of $\lambda$.** Given that the use of mixtures allows for a lot more flexibility in the model, it is relevant to ask: To what extent is the model identified? Rust (1994) has proven generic nonidentification of dynamic discrete choice models. Hence, some identifying restrictions have to be imposed in order to identify the primitives of the model. Recently, Magnac and Thesmar (2002) have shown that the

degree of underidentification consists of the discount factor, the distribution function for the unobserved state variables and preferences in a reference alternative.

Since we only observe sequences of discrete choices for which the level and scale of the value function is irrelevant, we can never hope to identify parameters that index the level or overall scale of the value function. Since the extreme value assumption is the identifying assumption with respect to the overall scale of the value function, the choice of $\lambda$ in example 1 does not have any empirical content and can be set arbitrarily. If, for example, $\mu$ is individual specific and additive in utility, changing $\lambda$ will just scale the variance of $\mu$.

However, in example 2, $\mu$ is time varying. In this case, we might want to think of $g\left(\mu'|\mu, x'\right)$ as being part of the behavioral model, whereas $\lambda\eta$ is a convenient smoothing device. Hence, we should therefore lower $\lambda$ as we get more and more observations in the sample, since simulated probabilities then get smoothed out across individuals. As $\lambda$ approaches zero, we will eventually approximate the behavioral model of interest. In this case, we need instead to restrict the $g\left(\mu'|\mu, x'\right)$ to obtain identification. At least one of the components of $\mu$ must fix the scale of the value function by normalizing the variance, while another component must fix the level by normalizing the mean (usually set to zero). Hence, the remaining factors are measured relative to the normalized components.

Magnac and Thesmar (2002) show that the degree of underidentification is even larger in the case with unobserved heterogeneity. They show that additivity in utility of individual effects is a powerful identifying restrictions. However, as we shall see in the numerical example in subsection 6.2, other parametric assumptions can be used to obtain identification. In particular, unobserved heterogeneity in coefficients to state variables are parametrically identified.

## 6. Numerical Results

6.1. **Performance of the Algorithm.** The estimation procedure applied in this paper makes use of numerical approximations in many dimensions: First, evaluation of the expectations in the Bellman equation is approximated by Monte Carlo integration. Second, the expected value function is approximated by Chebyshev polynomials. Third, conditional choice probabilities used for maximum likelihood estimation, are also approximated

by the use of simulation methods. It is therefore relevant to ask: What are the implications for statistical inference of the use of approximated likelihood functions?

The implications of simulation errors that originate from simulating the choice probabilities, are studied extensively in the literature on simulation assisted estimation.[12] However, little is known about the consequences for inference of the approximation error that originate from solving the dynamic programming model. Fernandez-Villaverde, Rubio-Ramirez, and Santos (2006) study the properties of the likelihood function when the value function is approximated numerically. They establish conditions for which convergence of the likelihood is obtained as the approximated policy functions converge to the exact policy. The implications of their results are depressing: Approximation errors in the policy function are amplified when evaluating the likelihood function: It is shown that the upper bound on the approximation error in the likelihoods is increasing linearly in the sample size. They also present an application based on simulated data that documents substantial biases in likelihood functions and parameter estimates when continuous state variables are approximated by a finite number of grid points. However, parameter estimates are shown to converge as the number of grid points is increased.

Still, our understanding of the consequences of approximation errors for empirical inference is limited. To the best of my knowledge, the case where continuous state variables are approximated by polynomials and integrals are evaluated using Monte Carlo simulation has not been investigated specifically. Hence, it remains to be seen how polynomial approximations perform. In particular, how many nodes are required to obtain reasonably precise ML estimates?

To evaluate the approximation approach, I use Rust's well known bus replacement model and data set.[13] Since the expected value function inherits some of the curvature from the specification of the utility, I adopt a non-linear specification of the cost function

---

[12]See e.g. Gouriéroux and Monfort (1996) and Train (2003)

[13]See Rust (1987). Rust's model has been extensively used in other studies to evaluate the performance of alternative algorithms and estimators, e.g., Hotz, Miller, Sanders, and Smith (1994) and Aguirregabiria and Mira (2002).

TABLE 1
APPROXIMATION ERROR IN CONDITIONAL CHOICE PROBABILITIES
ALTERNATIVE APPROXIMATIONS

| Number of nodes, $m$ | Based on ML estimates for alternative approximations | | | | |
|---|---|---|---|---|---|
| | 10 | 20 | 50 | 100 | 5000 |
| 2 | 6.31 | 6.31 | 6.31 | 6.31 | 6.31 |
| 4 | 1.26 | 1.16 | 1.20 | 1.20 | 1.17 |
| 6 | 0.16 | 0.10 | 0.09 | 0.09 | 0.10 |
| 8 | 0.17 | 0.13 | 0.11 | 0.11 | 0.13 |
| 10 | 0.07 | 0.03 | 0.01 | 0.01 | 0.02 |
| 50 | 0.05 | 0.00 | 0.03 | 0.02 | 0.00 |
| | Based on ML estimates for $R$=5000 and $m$=50 | | | | |
| | 10 | 20 | 50 | 100 | 5000 |
| 2 | 3.87 | 3.87 | 3.87 | 3.87 | 3.87 |
| 4 | 1.19 | 0.96 | 0.95 | 0.96 | 0.95 |
| 6 | 0.34 | 0.10 | 0.16 | 0.16 | 0.11 |
| 8 | 0.22 | 0.04 | 0.07 | 0.07 | 0.03 |
| 10 | 0.23 | 0.01 | 0.06 | 0.07 | 0.02 |
| 50 | 0.23 | 0.01 | 0.05 | 0.05 | 0.00 |

*Note:* All figures are measure in percentage points. Approximation error is measured as the maximum absolute deviation between the approximated choice probability and the "exact" choice probability. The "exact" solution was based on *R=5000*, and *m=50*. In the top panel, the choice probabilities were based on ML estimates based on using the alternative approximations. In lower panel, parameter estimates are held fixed at the ML estimates for the exact solution. In both cases, the choice probabilities were based on an approximation over the interval [0;500], and evaluated at the range of the state variable in a fine uniformly distributed grid (10000 points). The range of the observed state variable, $x_t$ is [0;387]

to rightfully access the performance. Therefore, the cost function is specified as $C(x) = c\sqrt{x}$.[14]

Since Rust's model does not have closed form value functions, I approximate the "exact" solution with 50 Chebyshev coefficients, $m = 50$, and 5000 Halton draws were used to evaluate integrals, i.e. $R = 5000$. This approximation is supposed to be very close to the exact solution; a 49 order polynomial is a *very* flexible function.

In Table 1, I have computed the approximation error obtained as the maximum absolute difference in conditional choice probabilities between the "exact" solution and the

---

[14]Specifically, the utility unction can be written as $u(x, d) = -I(d = 1)C(x) - I(d = 2)RC$ where $RC$ is a fixed cost of replacing a bus engine and $C(x) = c\sqrt{x}$ is maintenance operating costs as a function of elapsed mileage for the engine. The distribution for state variable, $x$, is a regenerative random walk $f(x_t|x_{t-1}, d) = h(x_t - I(d = 1)x_{t-1})$. Hence, by replacing the engine the state variable regenerates to zero.

approximated solution. This is done for various combinations of $R$ and $m$. To obtain the figures in the top panel of Table 1, the model was solved *and estimated* for each value of $m$ and $R$. Hence, these figures also include the bias from the ML estimation associated with a given approximation. In the lower panel, choice probabilities are based on the ML estimates obtained using the "exact" model. Hence, the figures in the lower panel, measure pure approximation error in the conditional choice probabilities and are not affected by the approximation bias in the ML estimates.

The results are striking: With only 6 Chebyshev nodes, choice probabilities can be estimated with an negligible approximation error (less than 0.2 percentage points). Moreover, the difference between the top and the lower panel is very small for $m \geq 6$, indicating that potential bias (if any) does not affect predictions from the model substantially.[15]

TABLE 2
BIAS IN ML ESTIMATES FOR ALTERNATIVE APPROXIMATIONS
RUST'S ENGINE REPLACEMENT MODEL
COST FUNCTION: $C(x) = c\sqrt{x}$

| Number of nodes, $m$ | | Engine replacement costs, $RC$ | | | Cost function parameter, $c$ | | |
|---|---|---|---|---|---|---|---|
| | $R$ | 10 | 20 | 5000 | 10 | 20 | 5000 |
| 2 | | -3.59 | -3.59 | -3.59 | 3.83 | 3.86 | 3.86 |
| | | (0.58) | (0.58) | (0.58) | (8.06) | (8.07) | (8.07) |
| 4 | | 0.70 | 0.64 | 0.64 | 0.38 | 0.75 | 0.73 |
| | | (1.73) | (1.72) | (1.72) | (3.84) | (3.95) | (3.94) |
| 6 | | 0.05 | -0.02 | -0.02 | -0.19 | -0.04 | -0.05 |
| | | (1.60) | (1.57) | (1.57) | (3.86) | (3.90) | (3.90) |
| 8 | | 0.09 | 0.03 | 0.03 | -0.11 | 0.05 | 0.05 |
| | | (1.60) | (1.58) | (1.58) | (3.84) | (3.90) | (3.89) |
| 10 | | 0.06 | 0.00 | 0.00 | -0.16 | 0.00 | -0.01 |
| | | (1.59) | (1.57) | (1.57) | (3.83) | (3.88) | (3.87) |
| 50 | | 0.06 | 0.00 | 0.00 | -0.16 | 0.01 | 0.00 |
| | | (1.59) | (1.56) | (1.56) | (3.81) | (3.85) | (3.85) |
| Paramter estimate ("exact" solution) | | | 11.14 | | | | 16.49 |

*Note:* The table presents the bias in the structural estimates induced by approximation error. Bias is measured as the difference between the *partial* ML estimates using an approximated and the "exact" solution. The "exact" solution was based on *R=5000* and n=50.

---

[15]I have tried to assess magnitude of the approximation error for various specifications of utility (linear, logarithmic, inverse hyperbolic sine and power functions). I have also evaluated performance on the basis of simulated data. Specifically, I have specified models with 2 continuous state variables, different state variable processes, correlated unobservables etc. The results are very stable: If $m \geq 6$, approximation and bias is negligible.

TABLE 3
LIKELIHOOD RATIO
ALTERNATIVE APPROXIMATIONS AGAINST "EXACT" SOLUTION

| Number of nodes, $m$ | Based on ML estimates for alternative approximations | | | | |
|:---:|:---:|:---:|:---:|:---:|:---:|
| | 10 | 20 | 50 | 100 | 5000 |
| 2 | -7.10 | -7.10 | -7.10 | -7.10 | -7.10 |
| 4 | 1.04 | 0.99 | 1.01 | 1.01 | 0.99 |
| 6 | -0.14 | -0.18 | -0.17 | -0.17 | -0.18 |
| 8 | 0.09 | 0.05 | 0.06 | 0.06 | 0.05 |
| 10 | 0.03 | -0.01 | 0.00 | 0.00 | -0.01 |
| 50 | 0.04 | 0.00 | 0.01 | 0.01 | 0.00 |
| Likelihood ("exact" solution) | | | | | -298.57 |

*Note:* The table presents the log-likelihood value under alternative levels of approximation, differenced against the "exact" solution. The "exact" solution were based on, *R=5000*, and n=50.

Table 2 reports the bias in partial ML estimates of the parameters, $c$ and $RC$ of Rust's engine replacement model with cost function $C(x) = c\sqrt{x}$. As before, the bias is measured as the difference between the estimates from the approximated model and the estimates from the "exact" model. The results point to very fast convergence in the ML estimates as the numerical approximation is improved. In fact, setting $m \geq 6$ eliminates the bias for all practical purposes.

As documented above, the approximation errors in choice probabilities and ML estimates are very small - for a reasonably small number of Chebyshev nodes. But what are the consequences for the likelihoods? This has important implications for modeling choice, since LR tests are often used to discriminate between competing models. If the likelihoods are subject to substantial approximation error, it may result in misleading inference and modeling choice.

In Table 3, I present approximation error in the log likelihood function. I.e. the approximated log-likelihood, differenced against the "exact" log-likelihood. Hence, the figures in Table 3 can be interpreted as a likelihood ratio test between the approximated and the exact model. When $m \geq 6$, the $LR$ delivers a solid answer of non-significance of approximation errors: the likelihood ratio test between the approximated model and the exact

solution is smaller than 0.2 in absolute value.[16] Compared to the findings in Fernandez-Villaverde, Rubio-Ramirez, and Santos (2006), the approximation error presented here is very small, indicating that the use of Chebyshev polynomials is a very efficient tool when solving and estimating dynamic discrete choice structural models.

One additional advantage of the polynomial approximation approach is that we are able to treat the state variables as continuous. If the model was solved using conventional discretization methods we also discretize the data, and important variation is lost. In Table 4, I have computed the mean and the variance of the approximation error that arise when discretizing state variables in Rust's bus data. From the last three columns of Table 4, it is clear that the discretization removes important variation in the data and induces substantial approximation error. Since the range of grid points must be determined by the state variable in levels, discretization has severe consequences for the differenced data; as the range of the differenced data is much lower. If the state variable is discretized with say 50 grid points, the standard deviation of the approximation error (the lost variation) is larger than the average per period change for the continuous data. For an approximation with say 90 grid points, the differenced data takes on only three values. I have tried to discretize the differenced data and then afterwards constructed the levels. As seen in the last column of Table 4, this amplifies the measurement problem.

---

[16]Vuong (1989) has derived the asymptotic distribution for this statistic when applied to non-nested models. Fernandez-Villaverde, Rubio-Ramirez, and Santos (2006) suggest using the test developed by Vuong (1989) to choose the degree of accuracy in the model. Using this test, I cannot reject the null that the log likelihood for models approximated with $m \geq 4$ is equal to the exact likelihood.

TABLE 4

BIAS AND APPROXIMATION ERROR DUE TO DISCRETIZATION

RUST'S ENGINE REPLACEMENT MODEL

COST FUNCTION: $C(x) = c\sqrt{x}$

| Number of grid points | Bias (Standard error) | | | Mean approx. error (Standard deviation of approx. error) | | |
|---|---|---|---|---|---|---|
| | c | RC | LR | $x_t(m)$ | $x_t(m)-x_{t-1}(m)$ | $dx_t(m)$ |
| 2 | -10.55 | -4.21 | -29.85 | 25.55 | -1.85 | 327.43 |
| | (1.02) | (0.36) | | (62.26) | (18.18) | (37.83) |
| 4 | -5.26 | -1.99 | -3.34 | 5.84 | -1.01 | 162.05 |
| | (2.57) | (1.14) | | (32.30) | (15.94) | (18.66) |
| 6 | -4.68 | -1.67 | -4.38 | 2.89 | -0.66 | 106.93 |
| | (2.33) | (1.03) | | (21.34) | (13.81) | (12.28) |
| 8 | -3.55 | -1.08 | -2.27 | 1.55 | -0.53 | 79.37 |
| | (2.93) | (1.33) | | (16.04) | (12.16) | (9.11) |
| 10 | -2.46 | -0.68 | -0.32 | 0.94 | -0.34 | 62.83 |
| | (2.98) | (1.32) | | (13.01) | (11.11) | (7.21) |
| 25 | -0.90 | -0.11 | -0.23 | 0.20 | -0.12 | 23.14 |
| | (3.57) | (1.54) | | (5.26) | (6.75) | (2.80) |
| 50 | -0.18 | 0.06 | 0.17 | 0.07 | -0.04 | 9.92 |
| | (3.77) | (1.59) | | (2.62) | (4.10) | (1.59) |
| 75 | 0.10 | 0.10 | 0.35 | 0.02 | -0.02 | 5.60 |
| | (3.87) | (1.61) | | (1.74) | (2.64) | (1.35) |
| 90 | 0.06 | 0.07 | 0.29 | 0.04 | -0.02 | 4.59 |
| | (3.85) | (1.59) | | (1.45) | (2.05) | (1.37) |
| 100 | 0.00 | 0.00 | 0.03 | 0.03 | -0.01 | 4.27 |
| | (3.85) | (1.59) | | (0.26) | (0.36) | (0.27) |
| | Parameter Estimates (Standard error) | | Likelihood | Mean of variable (Standard deviation of variable) | | |
| Continuous data | 16.39 | 11.14 | -298.56 | 115.91 | 3.32 | 3.32 |
| | (3.83) | (1.57) | | (84.77) | (1.42) | (1.42) |

*Note:* Bias is measured as the difference between parameter estimates, based on discretized and continuous data, respectively. The estimated standard errors of the parameter estimates are given in parenthesis. The likelihood ratio test statistic (LR) measures the difference in likelihood between models estimated on discretized and continuous data respectively. When estimating the parameters, the model was solved using 50 Chebyshev nodes and 5000 Monte Carlo draws. Approximation error is measured as the difference between the original continuous data and the discretized data. The lower part of the table presents parameter estimates and likelihood value for the model based on continuous data along with mean and standard deviation for levels and differences of the variables $x_t$ and $dx_t$. The grid points were uniformly distributed between 0 and 450. The range of the data, $x_t$, is [0;387]

Table 4 presents parameter estimates of Rust's model using discretized data, but still treating the state variables as being continuous when solving the model. Here I used the "exact" solution of the model, i.e. $R = 5000$ and $m = 50$. This allows me to isolate the effect from approximation error due to discretization of the data.

The consequences of discretizing the data are severe: If for example, $m = 6$, i.e. the data is truncated to the midpoint in 6 uniformly sized intervals, the bias in estimated replacement cost, $RC$, is approximately 2 standard errors from the estimate where the original data was used. This will indeed lead to misleading inference. For the same level of discretization, the difference in log-likelihood is $-4.38$. Suppose we want to test whether the model could be simplified to a competing (misspecified) model for which we have a closed form and therefore not would have to discretize the data. In this case, we could erroneously fail to reject the model, simply due to the presence of discretization error.

But how fine should the grid be? A least 100 grid points are needed to match the precision we obtain when solving the model with Chebyshev polynomials using $m = 6$ and $R = 20$. In terms of the computational complexity involved, the difference in using the two different approaches is moderate for Rust's model since it only has one continuous state variable. However, consider a model with four continuous state variables. If the state variable are discretized into $100^4 = 100.000.000$ points for each continuous variable, we would have to solve the model in 100 million points to obtain reasonably precise ML estimates. This involves computing a 100 million by 100 million transition matrix - and then we have to invert that matrix. This requires 1250 terabytes in storage. In contrast, only $6^4 = 1296$ points are needed when approximating with polynomials and the transition matrix $T(x) P_T$ is $1250R \times 1250$ and requires only 20.9 mega bytes of storage if $R = 100$.

This has an enormous effect on computer time for moderately sized problems. Therefore, when I adopt the suggested approach, I can estimate models with up to 5 state variables - on a three year old laptop.[17] Figure 1 presents the CPU time used to solve various discrete dynamic programming models as a function of the number of state variables. The model with 5 state variables took less than 3 hours to solve and a couple of days to estimate (using simulated data). Hence, using the suggested approach, it is for

---

[17]IBM thinkpad T41 with a 1.6 GHz Pentum M processor and 2 GB RAM.

FIGURE 1
CPU TIME USED TO SOLVE MODEL
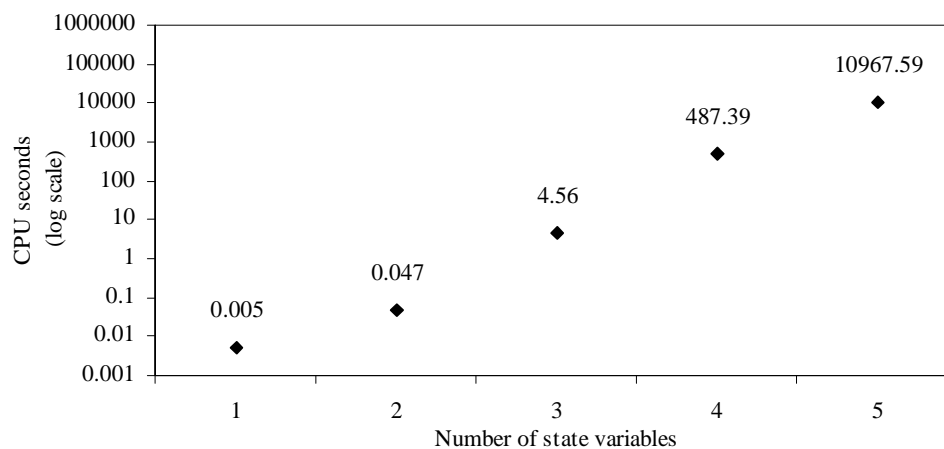


*Note:* When the models were solved, I used 100 Halton Draws to calculate integrals and 6 Chebyshev coefficients in each dimension of the state space for the models with up to 4 state variables. For the model with 5 state variables, I used only 5 Cebyshev coefficient in each dimension of the state space. The models were solved using a IMB ThinkPad T41 with a 1.6 GHz Pentium M processor and 2 GB RAM.

example possible to estimate a model with one observable state variable and three choice alternatives, where we allow for both individual specific heterogeneity with correlation across alternatives and serial correlation in the unobserved state variables.

6.2. **Unobserved heterogeneity: Monte Carlo Evidence.** In order to illustrate the model in example 1 and to evaluate the impact of potential heterogeneity bias, I carry out a Monte Carlo experiment based on a model similar to Rust's bus engine replacement model, where engine replacement costs, $RC$, are assumed to be bus specific and randomly distributed in the population of busses. Specifically, $RC$ is assumed to be normally distributed with mean $\overline{RC}$ and variance $\sigma^2_{RC}$.

If unobserved heterogeneity is introduced in a static discrete choice model where the random utility function is linear in the parameters, we will in general be able to estimate the mean of the coefficients up to a scale factor - even if we neglect the presence of heterogeneity.

I assume a linear cost function $C(x) = cx$, where the slope of the cost function, $c$, and the population mean of the replacement costs, $\overline{RC}$, are set roughly equal to the

TABLE 5
MONTE CARLO EXPERIMENT
FIXED AND RANDOM COEFFICIENTS

| | | Monte Carlo Distribution of ML and MSL estimates | | | | |
|---|---|---|---|---|---|---|
| | | Fixed Coefficients | | Random Coefficients | | |
| *Statistic* | $\sigma_{RC}^{dgp}$ | RC | c | $\mu_{RC}$ | $\sigma_{RC}$ | c |
| Mean Bias | 0 | 0.05 | 0.01 | 0.19 | 0.41 | 0.03 |
| | 1 | -0.52 | -0.07 | 0.05 | -0.02 | 0.01 |
| | 2 | -1.96 | -0.28 | 0.10 | 0.08 | 0.01 |
| | | | | | | |
| Mean Absolute E | 0 | 0.35 | 0.06 | 0.43 | 0.41 | 0.07 |
| | 1 | 0.57 | 0.08 | 0.43 | 0.29 | 0.06 |
| | 2 | 1.96 | 0.28 | 0.46 | 0.30 | 0.07 |
| | | | | | | |
| Monte Carlo | 0 | 0.44 | 0.074 | 0.51 | 0.38 | 0.082 |
| std. dev | 1 | 0.42 | 0.066 | 0.54 | 0.37 | 0.080 |
| | 2 | 0.51 | 0.075 | 0.57 | 0.38 | 0.084 |
| | | | | | | |
| Mean std. Error | 0 | 0.45 | 0.073 | 0.46 | 0.13 | 0.071 |
| | 1 | 0.41 | 0.068 | 0.52 | 0.08 | 0.079 |
| | 2 | 0.28 | 0.053 | 0.56 | 0.06 | 0.080 |

| | | Bias in estimated coefficients ratios | |
|---|---|---|---|
| | $\sigma_{RC}^{dgp}$ | RC/c | $\mu_{RC}/c$ |
| Mean Bias | 0 | -0.3% | -0.7% |
| | 1 | 2.4% | 0.2% |
| | 2 | 12.1% | 0.2% |

*Note:* The Monte Carlo experiment is based on 1000 Monte Carlo samples of sample size *N=100*, *T=250*. The model was solved using 5th order Chebyshev polynomials in each dimension of the state space. In the random coefficients model, two continuous variables were approximated (the observed state variable, *x*, and unobserved heterogeneity in RC). To evaluate integrals in the expected value function, I used *R=50* Halton draws. The number of simulations for each bus were set to *S=100*. The bias in the ratio RC/c and $\mu_{RC}/c$ is measured as the percentage discrepancy to the true value.

ML estimates from one of the linear specifications of Rust (1987).[18] For sample sizes of $N = 100$, $T = 250$, I draw 1000 Monte Carlo samples, and for each of them, I obtain partial ML estimates for models estimated with and without unobserved heterogeneity.

---

[18]Rust (1987) report ML estimates for bus groups 1,2 and 3 as $RC = 11.7270$ and $c = 0.001 * 4.8259$. However, the unit of measurement for $c$ is in units of the discretized state variable, $x_t^d = 1, 2, .., 90$. With the chosen dimension of the grid for mileage, the coefficient to the discretized variable $x_t^d$ must be divided by 5, since the interval length corresponds to 5000 miles in continuous measurement. I therefore set $\bar{c} = 1$ and $\overline{RC} = 1$

FIGURE 2: MONTE CARLO DISTRIBUTION OF ML ESTIMATES
FIXED AND RANDOM COEFFICIENTS
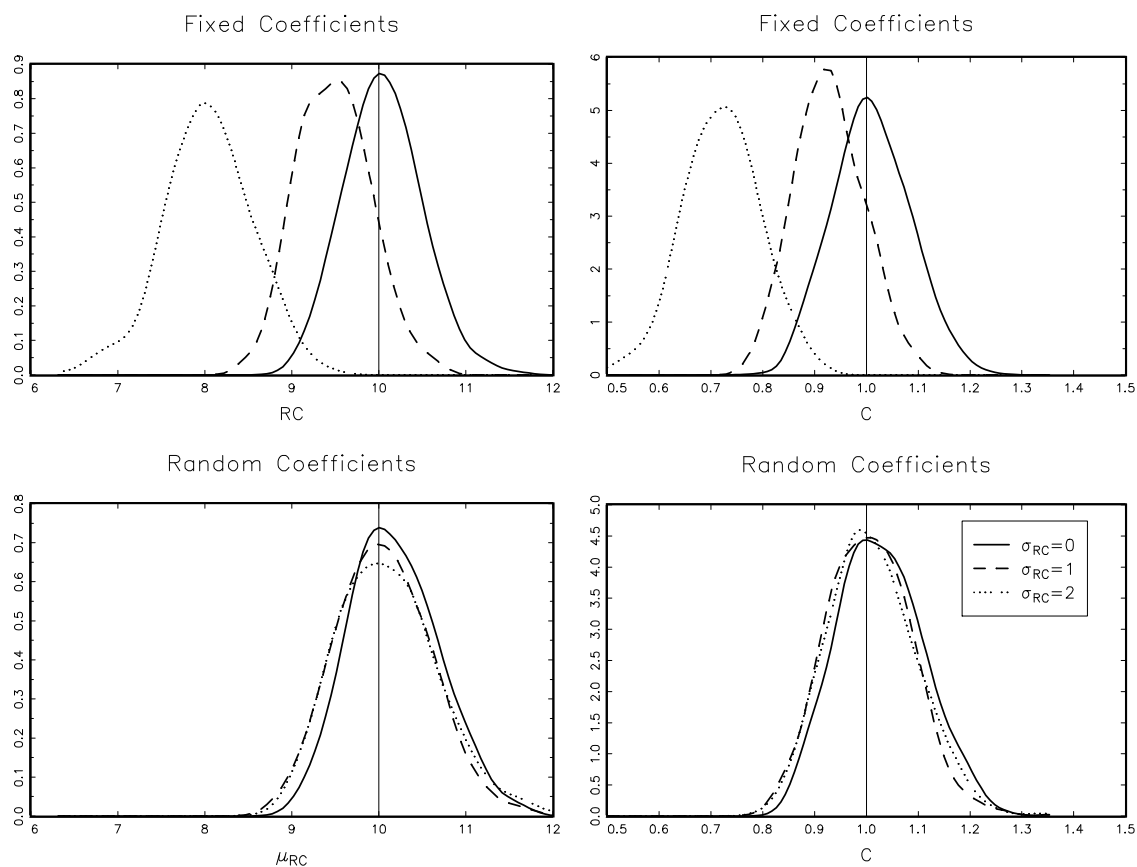


*Note:* See the note in Table 5 for a description of the experimental design.

In Figure 2 the distribution of the Monte Carlo estimates are displayed and Table 5 presents the summary statistics for this experiment. The parameters from the model that neglect unobserved heterogeneity are as expected substantially downward biased - and the bias is increasing in the heterogeneity. Moreover, due to the misspecified error structure, the estimated standard errors are substantially lower than the simulated Monte Carlo standard errors. In contrast, the model that allows for unobserved heterogeneity, provides unbiased parameter estimates and standard errors. The standard deviation of the heterogeneity can be estimated (and thus identified) when it is present, and in absence of heterogeneity the over parameterization comes at a very low efficiency cost.

As mentioned above, it is well know from static discrete choice models that neglected heterogeneity only has a scaling effect on the parameter estimates. In a dynamic model, however, it is not necessarily the case, since the coefficients of the utility function, $\theta_u$,

appear nonlinearly in the expected value function. Consequently, the threshold value of the state variables that makes the decision maker indifferent between the two alternatives will be non-linear in the coefficients of the model - even when they enter linearly in the utility function. Therefore, we will in general not estimate the mean of the structural parameters in the model. The lower panel in Table 5 shows that this is not the case in the dynamic model with fixed coefficients: The discrepancy between $RC/c$ and $\widehat{RC}/\hat{c}$ increases as the heterogeneity becomes more significant. Hence, not only the scale of the parameters are altered also their relative size. In contrast, the model with unobserved heterogeneity estimates the ratio without bias.

## 7. Conclusion

In this paper, I develop, implement and evaluate a continuous approach to Rust's Nested Fixed Point Algorithm, where orthogonal Chebyshev polynomials are used to approximate the value function in the dynamic programming problem. This strategy has several important spin-offs. First, it helps ameliorating the 'curse of dimensionality; Second, the approach permits very fast simulation of likelihood functions for mixture models, once value functions are approximated. Third, we do not need to discretize the data.

The applied approximation methods combined with ideas from the literature on mixed discrete choice models with simulation provides a useful framework for estimating very flexible models. Specifically, I assume that a part of the unobserved state variables are additively separable in utility and multivariate iid extreme value distributed. This assumption leads to the Dynamic Programming Mixed Logit (DPMXL) model.

DPMXL is a highly flexible model that obviates many of the limitations of Rust's model with additively separable, conditionally independent, extreme value distributed unobservables. DPMXL can potentially allow for i) random taste variation, ii) unrestricted substitution patterns, iii) correlation in unobserved state variables over time, iv) lots of observed and unobserved heterogeneity, and v) arbitrary shape of the distribution of unobserved factors. I present two examples of different specifications of the DPMXL

model. First, I consider the case of observed and unobserved individual specific heterogeneity, arbitrarily correlated across alternatives. Second, I consider the case of serially correlated unobservables.

The applied approximation method performs well for the DPMXL model. Since, the expected value functions are integrals over the continuous state variables and since they are choice specific, the approximated objects are very smooth. Therefore, polynomial approximations are particularly well suited for this problem. When evaluating approximation bias for alternative levels of approximation, I find that a fifth order polynomial provide sufficiently accurate approximations to the expected value function. Moreover, the approximation error, transmitted to conditional choice probabilities, likelihood functions and structural parameter estimates, are practically eliminated at this level of approximation.

The approximation method has the additional advantage that data on continuous state variables does not have to be discretized. To evaluate the impact of approximation error induced by discretizing the data, I estimate the model using discretized data, but still treating the state variables as being continuous when solving the model. This allows me to isolate the pure effect from approximation error due to discretization of the data. Discretizing data is clearly inefficient and results in substantial approximation error and variation loss in state variables. This is transmitted to significant bias in parameter estimates. When comparing to the suggested approximation approach, I find that at least 100 grid points are needed to match the precision we obtain when solving the model with a fifth order approximating Chebyshev polynomial. Hence, Chebyshev polynomials are a powerful tool in ameliorating the curse of dimensionality and mitigating the problem with approximation error associated with discretization of the data.

At the end of the paper, I present Monte Carlo evidence based on Rust's engine replacement model with unobserved heterogeneity, and analyse the relative performance of estimators with fixed and random coefficients. Even when heterogeneity enters linearly in the instant utility function, in general we do not estimate the mean of the coefficients. Parameters are heavily biased and the bias cannot entirely be attributed to the usual neglected heterogeneity problem also know from static models. Conversely, the random coefficient model performs quite well: parameter estimates and standard errors are unbiased and we are able to estimate the standard deviation of the heterogeneity

when it's present. In absence of heterogeneity, the over-parameterization comes at a very low efficiency cost. This accentuates the future role for mixture models in dealing with unobserved heterogeneity in dynamic models. Furthermore, the results highlights the importance of heterogeneity bias in dynamic programming models and adds to the debate about unobserved heterogeneity in microeconometrics.

## 8. Appendix: Simulated Likelihood and Gradients

In this appendix, I show how to compute simulated likelihood and analytical gradients for the models derived in section 5. I will carefully go through the procedure to obtain simulated likelihood and gradients for the model with individual specific heterogeneity from example 1. At the end, I briefly discuss the differences to the case of serially correlated unobservables from example 2.

For illustrative simplicity, consider the case where heterogeneity enters the utility function only. In this case, we can obtain consistent estimates for $\theta_f$, by maximizing the *partial* likelihood $l^2(\theta_f)$; we label these estimates $\hat{\theta}_f^p$. I also assume that $\beta$ is known and fixed. Hence we are interested in the parameters, $\theta_u$, that indexes fixed coefficients in the utility function, and the parameters, $\theta_g$ that index the population distribution of $\mu$. Given $\beta$ and $\hat{\theta}_f^p$, partial maximum (simulated) likelihood estimates can be obtained maximizing the partial likelihood, $l\left(\theta_u, \theta_g | \beta, \hat{\theta}_f^p\right)$. The procedure to obtain the simulated partial likelihood can be summarized as follows

**Algorithm 1.** *Simulation of the partial likelihood*

(1) *For a given set of structural parameters, $\theta$, and an appropriate choice of the smoothing parameter $\lambda$, solve the model as the fixed point $\widehat{EV}_\theta(x, \mu, d)$ for the contraction mapping defined by the functional equation (5.3).*

(2) *Given data on individual characteristics $z_n$, take $S$ individual and choice specific draws from $g(\mu | z_n; \theta_g)$ and label them $\mu_n^s = (\mu_n^s(1, z_n) \dots, \mu_n^s(J, z_n))'$, $s = 1, .., S$; $n = 1, ..., N$. To obtain identification, set $\mu_n^s(J) = 0$ for a reference alternative $J$ and normalize the variance of the first component in $\mu_n^s$.*

(3) *Given, i) the expected value function, $\widehat{EV}(x, \mu, d; \theta)$ evaluated at Chebyshev nodes,$(x, \mu)$ for each alternative, $d = 1, .., J$, ii) panel data on discrete choices $d_n = (d_{n1}, .., d_{nT_n})$ and observable state variables, $x_n = (x_{n1}, .., x_{nT_n})$ and iii) $S$ simulated values of the latent variable, $\mu_n^s$, $s = 1, ..., S$, we calculate the partial likelihood for the observed sequence of choices conditional on these data. The contribution of individual $n$ to*

*the simulated partial likelihood is*

$$\hat{l}_n^1\left(\theta\right) \;=\; \ln \hat{L}_n^1\left(\theta\right) = \ln \frac{1}{S} \sum_{s=1}^{S} \hat{L}_n^1\left(\theta|\mu_n^s\right)$$

$$=\; \ln \frac{1}{S} \sum_{s=1}^{S} \prod_{t=1}^{T_n} \hat{P}\left(d_{nt}|x_{nt}, \mu_n^s\right), \; n = 1, ..., N$$

*where*

$$\hat{P}\left(d_{nt}|x_{nt}, \mu_n^s\right) \;=\; \frac{\exp\left\{\left[\tilde{u}(x_{nt}, \mu_n^s(d_{nt}, z_n), d_{nt}) + \beta T\left(x_{nt}, \mu_n^s(d_{nt}, z_n)\right) p_T \widehat{EV}\left(x, \mu, d\right)\right]/\lambda\right\}}{\sum_{j \in D(x)} \exp\left\{\left[\tilde{u}(x_{nt}, \mu_n^s(j, z_n), j) + \beta T\left(x_{nt}, \mu_n^s(j, z_n)\right) p_T \widehat{EV}\left(x, \mu, d\right)\right]/\lambda\right\}}$$

$$s \;=\; 1, ..., S$$

To obtain maximum simulated likelihood estimates for $\theta_u$ and $\theta_g$, the BHHH algorithm is used to maximize the simulated partial likelihood function $\hat{l}^1\left(\theta\right) = \frac{1}{N} \sum_{n=1}^{N} \hat{l}_n^1\left(\theta\right)$. For each BHHH iteration we need to compute gradients in $\theta_u$ and $\theta_g$ for the log likelihood contribution of individual $n$. One immediate advantage of the smoothing approach is that it is relatively simple to derive analytical gradients of the likelihood function.

The gradient is just a sum of gradients of the simulated likelihood conditional on simulation $s$, divided by the simulated likelihood. Due to the extreme value assumption, both of these entities can be expressed analytically.

$$\nabla_\theta \hat{l}_n^1\left(\theta\right) = \frac{1}{\hat{L}_n^1\left(\theta\right)} \frac{1}{S} \sum_{s=1}^{S} \nabla_\theta \hat{L}_n^1\left(\theta|\mu_n^s\right)$$

To simplify notation, let $P\left(d\right)$ denote the conditional choice probability given $x_{nt}$ and simulation draw, $\mu_n^s$, i.e. $P\left(d\right) \equiv \hat{P}\left(d|x_{nt}, \mu_n^s\right)$. Let further $\hat{v}\left(d\right)$ denote the choice specific value function associated with alternative $j$ evaluated at the same data, i.e. $\hat{v}\left(d\right) \equiv \left[\tilde{u}(x_{nt}, \mu_n^s(d, z_n), d) + \beta T\left(x_{nt}, \mu_n^s(d, z_n)\right) p_T \widehat{EV}\left(x, \mu, d; \theta\right)\right]/\lambda$. Using Leibniz' rule for products of differentiable functions we obtain

$$\nabla_\theta \hat{l}_n^1\left(\theta, \mu_n^s\right) \;=\; \sum_{t=1}^{T_n} \frac{\nabla_\theta P\left(d_{nt}\right)}{P\left(d_{nt}\right)} \prod_{t=1}^{T_n} P\left(d_{nt}\right)$$

$$=\; \sum_{t=1}^{T_n} \left(\nabla_\theta \hat{v}\left(d_{nt}\right) - \sum_{j \in D} P\left(j\right) \nabla_\theta \hat{v}\left(j\right)\right) \hat{L}_n^1\left(\theta, \mu_n^s(z_n)\right)$$

This gradient is similar to the gradient for the conventional mixed logit used for panel data. The only thing that differs from the conventional mixed logit is the derivative of the

choice specific value function, $\nabla_\theta \hat{v}(d)$. We can write the gradient of the value function associated with alternative $d$ as the following expression

$$\nabla_\theta \hat{v}(d) = \left[ \nabla_\theta \tilde{u}(x_{nt}, \mu_n^s(d, z_n), j; \theta_u) + \beta T(x_{nt}, \mu_n^s(d, z_n)) \, p_T \nabla_\theta \widehat{EV}(x, \mu, d; \theta) \right] / \lambda$$

For most functional forms for $\tilde{u}(\cdot)$ it will be a trivial task to come up with an analytical expression for $\nabla_\theta \tilde{u}(x_{nt}, \mu_n^s(d_{nt}, z_n), d_{nt}; \theta_u)$. However, to compute the last term in the bracket, we need to differentiate the unknown value function. We can write the fixed point as $\widehat{EV} = \Gamma\left(\widehat{EV}\right)$ and use the implicit function theorem to obtain

$$\nabla_\theta \widehat{EV}(x, \mu, d; \theta) = [I - \Gamma'_\theta]^{-1} \, \partial \Gamma / \partial \theta'$$

Note that $[I - \Gamma'_\theta]^{-1}$ is a by-product of the solution algorithm. Recall that we have to compute the Fréchet derivative, $\Gamma'_\theta$, of the contraction mapping when using the Newton-Kantorowich algorithm. The last term, $\Gamma / \partial \theta'$, is a vector with the following elements

$$\partial \Gamma / \partial \theta_u = \nabla_{\theta_u} \tilde{u}(x_{nt}, \mu_n^s(d_{nt}, z_n), d_{nt}; \theta_u)$$

$$\partial \Gamma / \partial \theta_g = \nabla_\mu \tilde{u}(x_{nt}, \mu_n^s(d_{nt}, z_n), d_{nt}; \theta_u) \nabla_{\theta_g} \mu_n^s(d_{nt}, z_n)$$

The gradient of the simulation draws will in many cases be very easy to derive. For expositional simplicity, I here present a simple example where $\mu_n^s$ is $J-1$ dimensional multivariate normal distributed. Let for example, $\mu_n^s(d, z_n) = z_n \gamma_d + u_{nd}^s$ for $d = 1, , .J-1$, where $u_{nd}^s$ is a choice specific simulation draw from a $J-1$ dimensional multivariate normal with covariance matrix $\Omega_u$ and alternative specific mean $\bar{u} = (\bar{u}_1, .., \bar{u}_{J-1})'$. Let $\mu_n^s(J_{nt}, z_n) = 0$ for the reference alternative. In the case of normals, it is convenient to re-parameterize the model so that the covariance matrix $\Omega_u$ is decomposed into lower triangular Cholesky matrices $C_u$. When maximum likelihood estimates for these Cholesky factors are obtained, the covariance matrix can be derived as $\Omega_u = C_u C_u'$.

The re-parameterisation to Cholesky factors offers a simple way to simulate e.g. the multivariate normal distribution and ensures that the covariance matrix is positive definite since it is calculated as the square of the Cholesky matrix. A draw $u_n^s = \left(u_{n,1}^s, .., u_{n,J-1}^s\right)'$ from $N(\bar{u}, \Omega_u)$ is obtained by taking $S \times (J-1)$ draws from a standard normal for each individual. Label these draws $\xi_n^s = \left(\xi_{1n}^s, ..., \xi_{J-1,n}^s\right)'$ and calculate $u_n^s = \bar{u} + C_u \xi_n^r$. With this parameterization the structural parameters are $\theta_g = \left(\gamma_1, .., \gamma_{J-1}, \bar{u}_{1,...}, \bar{u}_{J-1}, c_{l,k} : \forall l, k \leq l \leq J-1\right)$.

We now express the elements of the gradient for a particular simulation draw

$$
\begin{aligned}
\nabla_{\gamma_j} \mu_n^s(d_{nt}, z_n) &= z_n I\left(j = d_{nt}\right), \ j = 1, .., J-1 \\
\nabla_{\bar{u}_j} \mu_n^s(d_{nt}, z_n) &= I\left(j = d_{nt}\right), \ j = 1, .., J-1 \\
\nabla_{c_{l,k}} \mu_n^s(d_{nt}, z_n) &= \xi_{kn}^s \qquad ,\forall l, k \leq l \leq J-1
\end{aligned}
$$

Note that at least one element of $\Omega_u$ must be normalized to ensure identification. This can be done by setting the top left element of $C_u$ equal to one , i.e. $c_{1,1} = 1$.

The simulated partial likelihood in example 2 can be derived following very similar steps. There are two important differences, however. First, when computing $\widehat{EV}\left(x, \mu, d|\theta\right)$, we must take expectations over $\mu$ as indicated in equation (5.7). Secondly, $\mu$ is drawn from $g\left(\mu'|\mu, x'\right)$ and is not constant over time. Hence each simulation $s$ consist of a $T_n$ dimensional sequence of serially correlated simulation draws rather than one permanent component. Subject to these differences, simulated likelihoods and gradients are identical to the ones derived above.

## References

AGUIRREGABIRIA, V. (1999): "The Dynamics of Markups and Inventories in Retailing Firms," *The Review of Economic Studies*, 66, 275–308.

AGUIRREGABIRIA, V., AND P. MIRA (2002): "Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models," *Econometrica*, 70, 1519–1543.

——— (2007a): "Dynamic Discrete Choice Structural Models: A Survey," *Annals of the Journal of Econometrics*, Forthcoming.

——— (2007b): "An Estimable Dynamic Model of Entry, Exit and Growth in Oligopoly Retail Markets," *American Economic Review*, Forthcoming.

——— (2007c): "Sequential Estimation of Dynamic Discrete Games," *Econometrica*, Fothcoming.

ECKSTEIN, Z., AND K. I. WOLPIN (1989): "The Specification and Estimation of Dynamic Stochastic Dicrete Choice Models: A Survey," *Journal of Human Resources*, 24, 562–598.

FERNANDEZ-VILLAVERDE, J., J. F. RUBIO-RAMIREZ, AND M. SANTOS (2006): "Convergence Properties of the Likelihood of Computed Dynamic Models," *Econometrica*, 74(1), 93–119.

GAMBA, A., AND M. TESSER (2006): "Structural Estimation of Real Options Models," Available at SSRN: http://ssrn.com/abstract=906990.

GOURIÉROUX, C., AND A. MONFORT (1996): *Simulation-Based Econometric Methods*. Oxford University Press, Oxford.

HOTZ, V. J., AND R. A. MILLER (1993): "Conditional Choice Probabilities and the Estimation of Dynamic Models," *Review of Economic Studies*, 60, 497–529.

HOTZ, V. J., R. A. MILLER, S. SANDERS, AND J. SMITH (1994): "A Simulation Estimator for Dynamic Models of Discrete Choice," *Review of Economic Studies*, 61, 265–289.

IMAI, S., N. JAIN, AND A. CHING (2006): "Bayesian Estimation of Dynamic Discrete Choice Models*," Working Paper 1118, Queen's University, Department of Economics.

IMAI, S., AND M. KEANE (2004): "Intertemporal Labor Supply and Human Capital Accumulation," *International Economic Review*, 45(2), 601–641.

JUDD, K. (1998): *Numerical Methods in Economics*. MIT press, Cambridge, Massachusetts.

KARLSTROM, A., M. PALME, AND I. SVENSSON (2004): "A Dynamic Programming Approach to Model the Retirement Behaviour of Blue-Collar Workers in Sweden," *Journal of Applied Econometrics*, 19(6), 795–807.

KEANE, M., AND K. WOLPIN (1994): "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence," *The Review of Economics and Statistics*, 74(4), 648–672.

KEANE, M. P., AND K. I. WOLPIN (1997): "The Career Decisions of Young Men," *Journal of Political Economy*, 105(3), 473–522.

KENNET, D. M. (1994): "Structural Model of Aircraft Engine Maintenance," *Journal of Applied Econometrics*, 9, 351–368.

MAGNAC, T., AND D. THESMAR (2002): "Identifying Dynamic Discrete Choice Processes," *Econometrica*, 70(2), 801–816.

MCFADDEN, D. (1989): "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration," *Econometrica*, 57(5), 995–1026.

MCFADDEN, D., AND K. TRAIN (2000): "Mixed MNL Models for Discrete Response," *Journal of Applied Econometrics*, 15(5), 447 – 470.

MILLER, R. A. (1997): "Estimating Models of Dynamic Optimization with Microeconomic Data," in *Handbook of Applied Econometrics*, ed. by H. Pesaran, and P. Smidth, pp. 246–299, Oxford, U.K. Blackwell.

PAKES, A. (1994): "Dynamic Structural Models: Problems and Prospects. Mixed Continuous Discrete Controls and Market Interactions," in *Advances in Econometrics, Proceedings of the Sixth World Congress of the Econometric Society*, ed. by C. Simms, pp. 171–259, Cambridge. Cambridge University Press.

RUST, J. (1987): "Optimal Replacement of GMC Bus Engines: An Emprirical Model of Harold Zurcher," *Econometrica*, 55, 999–1033.

——— (1988): "Maximum Likelihood Estimation of Discrete Control Processes," *SIAM Journal of Control and Optimization*, 26, 1006–1024.

———— (1994): "Structural Estimation of Markov Decision Processes," in *Handbook of Econometrics*, ed. by R. Engle, and D. McFadden. Elsevier, Norh-Holland.

———— (1996): "Numerical Dynamic Programming in Economics," in *Handbook of Computational Economics*, ed. by D. K. H. Amman, and J. Rust. Elsevier, North Holland.

RUST, J., G. HALL, H. BENÍTEZ-SILVA, G. J. HITSCH, AND G. PAULETTO (2005): "A Comparison of Discrete and Parametric Approximation Methods for Solving Dynamic Programming Problems in Economics," Unpublished Manuscript. Department of Economics. University of Maryland.

RUST, J., AND C. PHELAN (1997): "How Social Security and Medicare Affect Retirement Behavior In a World of Incomplete Markets," *Econometrica*, 65, 781–831.

SANCHEZ-MANGAS, R. (2001): "Pseudo-Maximum Likelihood Estimation of a Dynamic Structural Investment Model," Working Paper 01-56, Universidad Carlos III de Madrid.

SANTOS, M. S. (1999): "Numerical Solution of Dynamic Economic Models," *Handbook of Macroeconomics*, 1a, Chapter 5.

STINEBRICKNER, T. R. (2000): "Serially Correlated Variables in Dynamic, Discrete Choice Models," *Journal of Applied Econometrics*, 15(6), 595–624.

TRAIN, K. (2003): *Discrete Choice Models with Simulation.* Cambridge University Press, Cambridge.

VUONG, Q. H. (1989): "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses," *Econometrica*, 57(2), 307–333.

CENTRE FOR ECONOMIC AND BUSINESS RESEARCH (CEBR)

*E-mail address*: bsc.cebr@cebr.dk

*URL*: www.bschjerning.com

# LICENTIATAFHANDLINGER/PH.D.-AFHANDLINGER I ØKONOMI
## Københavns Universitet
_____

1. Hans Keiding: Some Topics in the Theory of Large Economies with Production. 1974. (Upubl.).

2. Rolf Norstrand: Statens økonomiske styring af kommunerne. 1974.
Publ.: *Gult memo* nr. 28, Københavns Universitets Økonomiske Institut 1974.

3. Uffe Bundgaard-Jørgensen & Iben Spliid: Skitse til en procedure for kommunalplanlægning.
1975. Publ.: *PhD theses* no. 21/22, IMSOR, Lyngby 1974.

4. Chr. Hjorth-Andersen: Forureningsøkonomi. 1975.
Publ. af Akademisk Forlag, København 1975.

5. Niels Kærgård: Efterspørgsel og varighed for varige goder. 1975. (Upubl.).

6. Peter Toft-Nielsen: MUCO ANALYSE om multicollinearitet i lineære økonomiske modeller. 1976.
(Upubl.).

7. Palle Geleff: Makromodeller for integreret fysisk økonomisk planlægning på regional niveau.
1978. (Upubl.).

8. Flemming Dalby: 1978
        a) Statsfinansernes indenlandske likviditetsvirkning I.
        Publ.: *Blåt memo* nr. 46, Københavns Universitets Økonomiske Institut 1976.
        b) Statsfinansernes indenlandske likviditetsvirkning II.
        Publ.: *Blåt memo* nr. 51, Københavns Universitets Økonomiske Institut 1976.
        c) Statsgældsspolitik. (Upubl.).

9. Peter Erling Nielsen: Reale og monetære renteteorier siden Wicksell. 1978. (Upubl.).

10. Niels Lihn Jørgensen: 1978.
        a) En perspektivplanlægningsmodel.
        Publ.: *Gult memo* nr. 44. Københavns Universitets Økonomiske Institut 1977.
        b) Dynamisk input-output teori.
        Publ.: *Gult memo* nr. 47, Københavns Universitets Økonomiske Institut 1977.
        c) En dynamisk input-output model for Danmark.
        Publ.: *Gult memo* nr. 56, Københavns Universitets Økonomiske Institut 1978.

11. Thorbjørn Waagstein: Monopol, teknologisk udvikling og investering i barrierer.
Publ.: *Blåt memo* nr. 94, Københavns Universitets Økonomiske Institut, 1980.

12. Michael Fuchs: Aspects of Monetary Policy as an Element of Stabilization Policy in open Economies. 1980.
Publ.: *Rød serie* nr. 1, Københavns Universitets Økonomiske Institut, 1981.

13. Lars Thorlund-Petersen: Independence and Convexity of Preference Relations. 1981.
Publ.: *Rød serie* nr. 2, Københavns Universitets Økonomiske Institut, 1981.

14. Niels Holger Skou: Substitution mellem arbejdskraft og kapital i udviklingslandene. 1982.
Publ. som følgende *Gule memoer* fra Københavns Universitets Økonomiske Institut:
> Nr. 54: Teknologioverførsel til udviklingslandene. 1978.
> Nr. 61: En model til undersøgelse af faktorsubstitutionsmulighederne i udviklingslande: Chenery-Raduchel modellen. 1978.
> Nr. 62: Substitution mellem arbejdskraft og kapital i udviklingslande. En empirisk undersøgelse for Indien og Mexico. 1978.

15. Christian Groth: Konjunkturteorier set i lyset af moderne uligevægtsteori. 1982.
Publ. som *Blåt memo* nr. 101a og nr. 101b (Matematisk apendiks): Cyklisk kapitalakkumulation. En diskussion og nogle udvidelser af Marx-Goodwin-modellen. Københavns Universitets Økonomiske Institut, 1981.

16. Niels Westergård-Nielsen: The Labor Market for a Group of Young Professionals. 1982.
Licentiatafhandlingen består af følgende 8 dele:
> 1. A Study of a Professional Labor Market - Introduction and Data. *Working Paper* 1981-2, Studies in Labor Market Dynamics, Århus Universitets Økonomiske Institut.
> 2. Job Search in a Professional Labor Market, 1980. *Discussion Paper* 606-80, Institute for Research on Poverty, University of Wisconsin-Madison.
> 3. Determinants for Search Prior to Unemployment, 1981. Cykelafdelingen, Københavns Universitets Økonomiske Institut.
> 3a. Appendix: A Note on Estimation of Probability Functions, 1981. Cykelafdelingen, Københavns Universitets Økonomiske Institut.
> 4. Do the Jobs Differ for Groups with Various Durations of Search? 1981. (Upubl.)
> 5. The Composition of Salary and Other Types of Compensation, 1981. *Gult memo* nr. 84, Cykelafdelingen, Københavns Universitets Økonomiske Institut.
> 6. Three Salary Models, 1981. *Working Paper* 1981-3. Studies in Labor Market Dynamics, Århus Universitets Økonomiske Institut
> 7. Estimation of the Reservation Wage, 1981. *Working Paper* 1981-4, Studies in Labor Market Dynamics, Århus Universitets Økonomiske Institut.
> 8. The Grades of Law Graduates, 1981. *Gult memo* nr. 85, Cykelafdelingen, Københavns Universitets Økonomiske Institut.

17. Troels Østergaard Sørensen: Koordinations- og allokeringsproblemer i en decentral markedsøkonomi med særligt henblik på arbejdsmarkedet. 1983.
Publ.: *Rød serie* nr. 3, Københavns Universitets Økonomiske Institut, 1983.

18. Esben Dalgaard: Moderne teorier om indkomstdannelsen . 1984.
Publ.: *Rød serie* nr. 4, Københavns Universitets Økonomiske Institut, 1984.

19. Flemming Møller : Social cost-benefit analyse - teori og praksis. 1984.
Publ.: *Rød serie* nr. 5, Københavns Universitets Økonomiske Institut, 1984.

20. Gert Aage Nielsen: En empirisk analyse af dansk eksport. 1984. Del 1+2.
Publ.: *Rød serie* nr. 6, Københavns Universitets Økonomiske Institut, 1984.

21. Michael Møller: Det danske boligmarked. 1984. Udgivet af Institut for Finansiering, Handelshøjskolen i København.

22. Søren Bo Nielsen: Bidrag til analysen af kortsigtede makroøkonomiske effekter af energipris chok i åbne økonomier. 1984 . Del 1+2.
Publ.: *Rød serie* nr. 7, Københavns Universitets Økonomiske Institut, 1984.

23. Tom Hefting: Energiprischok og disses kortsigtede makroøkonomiske effekter i en lille åben økonomi. 1984.
Publ.: *Rød serie* nr. 8, Københavns Universitets Økonomiske Institut, 1985.

24. Bodil Olai Hansen: Efficiensproblemer under ubegrænset tidshorisont. Del 1+2. 1985.
Publ.: *Rød serie* nr. 9, Københavns Universitets Økonomiske Institut, 1985.

25. Peter Birch Sørensen: Stabilization Policy in an Inflationary Economy. Essays in Monetary and Fiscal Policy. Del 1+2. 1985
Publ.: *Rød serie* nr. 10, Københavns Universitets Økonomiske Institut, 1985.

26. Lars Otto: En analyse af produktionsmulighederne i fremstillingserhverv. 1986.
Publ.: *Rød serie* nr. 11, Københavns Universitets Økonomiske Institut, 1986.

27. Fredrik Pitzner-Jørgensen: COMECON-samarbejdet - modeller for praktisk planlægning for et system af økonomier. 1986.
Publ.: *Rød serie* nr. 12, Københavns Universitets Økonomiske Institut, 1986.

28. Leif Hasager: National Economic Planning. 1987.
Publ.: *Rød serie* nr. 13, Københavns Universitets Økonomiske Institut, 1987.

29. Paul Bruniche-Olsen: Mobilitet og tilpasninger på arbejdsmarkedet. 1987.
Publ.: *Rød serie* nr. 14, Københavns Universitets Økonomiske Institut, 1987.

30. Jesper Munksgaard Pedersen: LINRAT - en energirationeringsmodel for Danmark. 1987.
Publ.: *Rød serie* nr. 15, Københavns Universitets Økonomiske Institut, 1987.

31. Tage Vosbein: Udvalgte områder af økonomisk reform i Kina 1978 til 1985. 1987. (Upubl.).

32. Hans Aage: Eksempler på anvendt makroteori. December 1987.
Publ.: *Blå memoer* nr. 50 (1976), 59 (1977), 123 (1984), 142 (1986 ) Københavns Universitets Økonomiske Institut. Desuden indgår et antal upublicerede manuskripter og konferenceindlæg i afhandlingen.

33. Karsten Albæk: En økonomisk analyse af lære- og praktik-pladsproblemet.
Publ.: *Rød serie* nr. 16, Københavns Universitets Økonomiske Institut, 1988.

34. Klavs Lindeneg: Planlægning og decentralisering af økonomisk aktivitet, 1988.
Publ.: *Rød serie* nr. 17, Københavns Universitets Økonomiske Institut, 1988.

35. Agnete Raaschou-Nielsen: Institutionel ændring og økonomisk teori, 1988.
Publ.: *Rød serie* nr. 18, Københavns Universitets Økonomiske Institut, 1988.

36. Michael Teit Nielsen: Expectations and Learning in an Overlapping Generations Economy, 1989.
Publ.: *Rød serie* nr. 19, Københavns Universitets Økonomiske Institut, 1989.

37. Hans Jørgen Jacobsen og Christian Schultz: Arbejdsløshed i markedsøkonomier.
Licentiatafhandlingen består af:
      1. Indledning. (Upubl).
      2. "Inefficiency of Unemployment Allocations", en revideret udgave heraf "Decreasing Unemployment Increases Welfare", publ. i *Scandinavian Journal of Economics*, vol. 94(3), 1992.
      3. "A General Equilibrium View of Unemployment". I R. Lund and P. Pedersen eds.: "*Unemployment: Theory, Policy and Structure*", Walter de Gruyter and Co., Berlin, New York, 1987.
      4. "A Macro Model of Conjectural Equilibrium". *Scandinavian Journal of Economics*, vol. 88(3), 1986.
      5. "On the Existence of Conjectural Equilibria with a priori Chosen Rationing Patterns". *European Economic Review* 31, 1987.
      6. Jacobsen, H.J. and C. Schultz: "A General Equilibrium Macro Model with Wage Bargaining". Discussion Paper 87-05, Institute of Economics, University of Copenhagen, en revideret udgave er kommet med samme titel i *Scandinavian Journal of Economics*, vol. 92(3), 1990.

7. Jacobsen , H.J. and C. Schultz: "Unemployment and Economic Policy in a General Equilibrium Model with Decentralized Wage Bargaining". Upub., en revideret udg. heraf er: "Wage Bargaining and Unemployment in a General Equilibrium Model", *Discussion Paper* 89-01, Institute of Economics, University of Copenhagen.

38. Bent E. Sørensen: 4 Essays on Econometric Estimation of Economic and Financial Models. 1990. Publ.: *Rød serie* nr. 20, Københavns Universitets Økonomiske Institut, 1990.

39. Lars Haagen Pedersen: Intertemporal Optimization & Investment. June 1991. Publ.: *Rød serie* nr. 21, Københavns Universitets Økonomiske Institut, 1991.

40. Flemming Christensen: Topics in Cooperative Game Theory. December 1991. Publ.: *Rød serie* nr. 22, Københavns Universitets Økonomiske Institut, 1991.

41. Sven Wunder: Dutch Disease theory and the case of Colombia. Februar 1992. Publ.: *Rød serie* nr. 23, Københavns Universitets Økonomiske Institut, 1992.

42. Niels Buus Kristensen: Lønstrukturens udvikling på funktionærområdet i Danmark. Februar 1992. Publ.: *Rød serie* nr. 24, Københavns Universitets Økonomiske Institut, 1992.

43. Jørgen Ravn Elkjær: Entreprenører, markeder og hierarkier. Februar 1992. Publ.: *Rød serie* nr. 25, Københavns Universitets Økonomiske Institut, 1992.

44. Eskil Heinesen: Makroøkonomiske rationeringsmodeller baseret på aggregering af mikro-markeder: Teori og estimation. Juni 1992. Publ.: Rød serie nr. 26, Københavns Universitets Økonomiske Institut, 1992.

45. Birgitte Sloth: Topics in Non-cooperative Game Theory: Fictitious Play and Majority Voting. November 1992. Publ.: *Rød serie* nr. 27, Københavns Universitets Økonomiske Institut, 1992.

46. Finn Tarp: Stabilization and Structural Adjustment: Macroeconomic Frameworks for Analysing the Crisis in Sub-Saharan Africa. (Udkommet på forlaget Routledge, 1993).

47. Henrik Dan Lando: Six Essays on Unemployment Theory, Flexibility Preference and Financial Contracts. Januar 1993. Publ.: *Rød serie* nr. 28, Københavns Universitets Økonomiske Institut, 1993.

48. Lisbeth Funding la Cour: Divisia Monetary Aggregates for Denmark: Theory, Construction and Empirical Performance. Februar 1993. Publ.: *Rød serie* nr. 29, Københavns Universitets Økonomiske Institut, 1993.

49. Sverrir Sverrisson: The Economic Interdependence between the Industrial and the Developing Countries - the Channels of Transmission. Juni 1993.
Publ.: *Rød serie* nr. 30, Københavns Universitets Økonomiske Institut, 1993.

50. Christian Hansen: Uddannelse som produktionsfaktor - formelle uddannelsers betydning for økonomien i 1980'erne. (Udkommet på Institut for Grænseregionsforskning). Oktober 1993.
51. Lasse Møller: Rent Seeking. Juni 1994.
Publ.: *Rød serie* nr. 31, Københavns Universitets Økonomiske Institut, 1994.

52. Ulrika Enemark: An Economic Analysis of Increased Patient Autonomy in Primary Health Care. August 1994.
Publ.: *Rød serie* nr. 32, Københavns Universitets Økonomiske Institut, 1994.

53. Jens Leth Hougaard: Six Essays on Performance Evaluation. August 1994.
Publ.: *Rød serie* nr. 33, Københavns Universitets Økonomiske Institut, 1994.

54. Lisbeth Pedersen: Løndannelse og løntransmission - en analyse af arbejderlønninger i Danmark. (Udkommet på Socialforskningsinstituttet, København). December 1994.

55. Karsten Nimb Pedersen: Empiriske makromodeller for det mellemlange sigt. September 1994.
Publ.: *Rød serie* nr. 34, Københavns Universitets Økonomiske Institut, 1994.

56. Martin Rasmussen: Labor unions and international trade. December 1994.
Publ.: *Rød serie* nr. 35, Københavns Universitets Økonomiske Institut, 1994.

57. Jens Bonke: Faktotum - husholdningernes produktion. 1994.
Publ.: Udkommet på Socialforskningsinstituttet, København, 1995.

58. Ebbe Hendon: Fictitious Play in Games and Lower Probabilities in Decision Theory. 1994.
Publ.: *Rød serie* nr. 37, Københavns Universitets Økonomiske Institut, december 1995.

59. Torben Tranæs: Essays on Games and Decisions. 1994.
Publ.: *Rød serie* nr. 40, Københavns Universitets Økonomiske Institut, januar 1996.

60. Peter Fristrup: Social Choice - Implementering og Manipulering. Maj 1995.

61. Anders Holm: Unemployment and Education. Micro and Macro Empirical Evidence. August 1995.
Publ.: *Rød serie* nr. 38, Københavns Universitets Økonomiske Institut, 1996.

62. Peter Stephensen: General Equilibrium and Growth. September 1995.

63. Henrik Hansen: Cointegrated Vector Autoregressive models: Theory, Applications and Software. November 1995.
Publ.: *Rød serie* nr. 36, Københavns Universitets Økonomiske Institut, 1995.

64. Hans Christian Kongsted: Dynamic Models of Foreign Trade under Fluctuating Exchange Rates: Theoretical and Empirical Applications. December 1995.
Publ.: *Rød serie* nr. 39, Københavns Universitets Økonomiske Institut, 1996.

65. Henning Bjerregård Bach: Lønmodtageres mobilitet og mobilitet på arbejdsmarkedet. Januar 1996.

66. Trine Bille Hansen: Kulturøkonomiske studier - Økonomiske metoder vedrørende ressourceallokering til kultur- og fritidsaktiviteter. Juli 996.
Publ.: *Rød serie* nr. 41, Københavns Universitets Økonomiske Institut, 1996.

67. Lars Gårn Hansen : Aspects of Designing Institutions for Environmental Regulation. April 1997.
Publ.: *Rød serie* nr. 42, Københavns Universitet. Økonomisk Institut, 1997.

68. Peter Hjertholm: An Inquiry Into the Fiscal Dimension of External Debt: The Case of Sub-Saharan Africa. Juni 1997.
Publ.: *Rød serie* nr. 43, Københavns Universitet. Økonomisk Institut, 1997.

69. Karsten Stæhr: Six Essays on Wages, Price Inflation, and Stabilization Policies. Juni 1998.
Publ.: *Rød serie* nr. 44, Københavns Universitet. Økonomisk Institut, 1998.

70. Jonathan Rubin: Econometric Studies of a European Economic and Monetary Union. September 1998.
Publ.: *Rød serie* nr. 45, Københavns Universitet. Økonomisk Institut, 1998.

71. Morten Binder: Aktiv arbejdsmarkedspolitik og økonomisk marginalisering - en teoretisk analyse af den aktive arbejdsmarkedspolitiks langsigtede virkninger på ledighedens størrelse og fordeling. Oktober 1997.
Publ.: *Rød serie* nr. 55, Københavns Universitet. Økonomisk Institut, 1999.

72. Mette Wier: Vækst, struktur og miljø - 6 bidrag til makroøkonomisk miljømodeludvikling. December 1997.
Udgivet af Danmarks Miljøundersøgelser 1998.

73. Kim Martin Lind: Sectoral Models For Producers' Long and Short Run Behaviour Applied to Danish Pig Production: Construction, Estimation and Testing. August 1998.
Publ.: *Rød serie* nr. 46, Københavns Universitet. Økonomisk Institut, 1998.

74. Torsten Sløk: Essays in Policy-oriented Macroeconomics. September 1998.
Publ.: *Rød serie* nr. 47, Københavns Universitet. Økonomisk Institut, 1998.

75. Claus Thustrup Hansen: Imperfections in Commodity and Labour Markets – a Macroeconomic Approach. Oktober 1998.
Publ.: *Rød serie* nr. 48, Københavns Universitet. Økonomisk Institut. 1998.

76. Peter Lyk-Jensen: Essays in Collusive and Co-operative Auction Procedures. Oktober 1998.
Publ.: *Rød serie* nr. 49, Københavns Universitet. Økonomisk Institut. 1998.

77. Niels Vestergaard: Property Rights Based Regulation of Fishery: Applications and Theory. Oktober 1998.
Publ.: *Rød serie* nr. 50, Københavns Universitet. Økonomisk Institut, 1998.

78. Lill Thanning Hansen: Endogenous Growth Theory: Essays on Optimal Policy, Market Imperfections and Fertility. November 1998.
Publ.: *Rød serie* nr. 51, Københavns Universitet. Økonomisk Institut, 1998.

79. Tove Christensen: Topics in Environmental and Natural Resource Economics with Emphasis on Uncertainty and Irreversibility. December 1998.
Publ.: *Rød serie* nr. 52, Københavns Universitet. Økonomisk Institut, 1998.

80. Trine Filges: Labour Unions and Unemployment. Januar 1999.
Publ.: *Rød serie* nr. 53, Københavns Universitet. Økonomisk Institut, 1999.

81. Thomas Bue Bjørner: Transport Demand and Environmental Regulation in the Transport Sector. Januar 1999.
Publ.: *Rød serie* nr. 54, Københavns Universitet. Økonomisk Institut, 1999.

82. Birthe Larsen: Unemployment in Search Equilibrium. Maj 1999.
Publ.: *Rød serie* nr. 60, Københavns Universitet. Økonomisk Institut, 1999.

83. Lise Nielsen: Forhandlinger i åbne og lukkede økonomier. Juni 1999. (Upubl.).

84. Poul Schou: Endogenous Growth, Nonrenewable Resources and Environmental Problems. August 1999.
Publ.: *Rød serie* nr. 56, Københavns Universitet. Økonomisk Institut, 1999.

85. Anne Kaag Andersen: Location and Commuting. November 1999.
Publ.: *Rød serie* nr. 57, Københavns Universitet. Økonomisk Institut, 1999.

86. Jacob Vastrup: Essays on Policy Related Issues in Macroeconomics. December 1999.
Publ.: *Rød serie* nr. 61, Københavns Universitet. Økonomisk Institut, 2000.

87. Jens Hauch: Environmental Regulation and Nordic Energy Markets: An Economic Simulation Approach. Februar 2000.
Publ.: *Rød serie* nr. 58, Københavns Universitet. Økonomisk Institut, 2000.

88. Rasmus Heltberg: Institutions in Rural Development: Property Rights and Natural Resources. Februar 2000.
Publ.: *Rød serie* nr. 59, Københavns Universitet. Økonomisk Institut, 2000.

89. Clara M.D. Jørgensen: On the I(2) Cointegration Model. Oktober 1999.
Publ. *Rød serie* nr. 66, Københavns Universitet. Økonomisk Institut, 2000.

90. Jacob Gyntelberg: Essays in Corporate Finance and Contract Theory. Marts 2000.
Publ.: *Rød serie* nr. 62, Københavns Universitet. Økonomisk Institut, 2000.

91. Lars Jebjerg: Regulation. Agency Problems in Regulation. Marts 2000.
Publ.: *Rød serie* nr. 63, Københavns Universitet. Økonomisk Institut, 2000.

92. Lars Even Rasmussen: Local Housing Markets with Partial Rent Control. Marts 2000. (Upubl.).

93. Henrik Klinge Jacobsen: Structural Change of the Economy, Technological Progress and Long-Term Energy Demand. April 2000.
Publ.: *Rød serie* nr. 64, Københavns Universitet. Økonomisk Institut, 2000.

94. Jens Anders Kovsted: Topics in Development Economics. Financial Sector, Child Health, and Reconstruction from War. April 2000.
Publ.: *Rød serie* nr. 65, København. Universitet. Økonomisk Institut, 2000.

95. Frank M. Rasmussen: Essays in International Trade and Labor Market Flexibility. Juni 2000.
Publ.: *Rød serie* nr. 67, Københavns Universitet. Økonomisk Institut, 2000.

96. Mogens Jensen: Essays in Non-cooperative Game Theory, Evolutionary Learning and Signalling Games. Maj 2000.
Publ.: *Rød serie* nr. 69, Københavns Universitet. Økonomisk Institut, 2000.

97. Mette Ejrnæs: Topics in Microeconometrics: Panel Data and Repeated Cross Sections. Juni 2000.
Publ.: *Rød serie* nr. 68, Københavns Universitet. Økonomisk Institut, 2000.

98. Jesper Jensen: Economic Analyses of Carbon Abatement Policies. September 2000.
Publ.: *Rød serie* nr. 70, Københavns Universitet. Økonomisk Institut, 2000.

99. Søren Kyhl: Five Essays on Imperfections in Economics. September 2000.
Publ.: *Rød serie* nr. 71, Københavns Universitet. Økonomisk Institut, 2000.

100. Anders S. Hoffmann: Imperfect Competition, Multinational Firms and Economic Policy. Maj 2001.
Publ.: *Rød serie* nr. 72, Københavns Universitet. Økonomisk Institut, 2001.

101. Klaus Skytte: Topics on Electricity Trade. Juni 2001.
Publ.: *Rød serie* nr. 73, Københavns Universitet. Økonomisk Institut, 2001.

102. Martin Kjeldsen-Kragh: Vintage Capital and Replacement Echoes. Februar  2001. (Upubl.)

103. Morten Igel Lau: Choice over Time: Individual Discount Rates and Dynamic Life Cycle Models. September 2001.
Publ.: *Rød serie* nr. 74, Københavns Universitet. Økonomisk Institut, 2001.

104. Claes Bengtsson: Voters, Politicians, and Game Theory: Some Theoretical and Empirical Results. November 2001.
Publ.: Rød serie nr. 75, Københavns Universitet. Økonomisk Institut, 2001.

105. Jacob Lester: Essays in Capital Structure Theory and Financial Markets. November 2001.
Publ.: *Rød serie* nr. 76, Københavns Universitet. Økonomisk Institut, 2001.

106. Anders Tauber Lassen : Essays in Finance: An International Term Structure Model and Optimal Choice and Asset Prices in a Monetary Economy. December 2001.
Publ.: *Rød serie* nr. 77, Københavns Universitet. Økonomisk Institut, 2002.

107. Claus Chr. Pörtner: Children and Household Behaviour in Developing Countries. December 2001. Publ.: *Rød serie* nr. 78, Københavns Universitet. Økonomisk Institut, 2002.

108. Toke Ward Petersen: Uncertainty, Indivisibility and Hyperbolic Preferences in Computable General Equilibrium Models. Januar 2002.
Publ.: *Rød serie* nr. 79, Københavns Universitet. Økonomisk Institut, 2002.

109. Uffe Nielsen: Essays in Environment and Development Economics. April 2002.
Publ.: *Rød serie* nr. 80, Københavns Universitet. Økonomisk Institut, 2002.

110. David Dreyer Lassen: Essays in the Political Economy of Public Finance. Juni 2002.
Publ.: *Rød serie* nr. 81, Københavns Universitet. Økonomisk Institut, 2002.

111. Jacob Nielsen Arendt: Essays in Applied Micro-Econometrics: Socio-Economic Differences in Health. Juni 2002.
Publ.: *Rød serie* nr. 82, Københavns Universitet. Økonomisk Institut, 2002.

112. Signe Krarup: Voluntary Agreements in Energy Regulation. September 2002.
Publ.: *Rød serie* nr. 83, Københavns Universitet. Økonomisk Institut, 2002.

113. Torben Mark Pedersen: Business Cycles and Economic Growth. Februar 2002. (Upubl.)

114. Martin Junker Nielsen: Five Essays on Commodity Pricing, Irreversible Investments, Convertible Bonds, and the Liquidation of Bad Loans. Oktober 2002.
Publ.: *Rød serie* nr. 84, Københavns Universitet. Økonomisk Institut, 2002.

115. Søren Arnberg: Empirical Studies of Agricultural Land Allocation. Oktober 2002.
Publ.: *Rød serie* nr. 85, Københavns Universitet. Økonomisk Institut, 2002.

116. Chantal Pohl Nielsen: Supply-Side Issues in Developing Country Agriculture: Constraints and Opportunities. December 2002.
Publ.: *Rød serie* nr. 86, Københavns Universitet. Økonomisk Institut, 2003.

117. Henrik Jacobsen Kleven: Taxation, Time Allocation, and Economic Efficiency. Januar 2003.
Publ.: *Rød serie* nr. 87, Københavns Universitet. Økonomisk Institut, 2003.

118. Carl-Johan Dalgaard: Growth and Inequality of Nations. September 2002.
Publ.: *Rød serie* nr. 88, Københavns Universitet. Økonomisk Institut, 2003.

119. Søren Leth-Petersen: Empirical Studies of Micro Data on Residential Energy Demand. Februar 2003.
Publ.: *Rød serie* nr. 89, Københavns Universitet. Økonomisk Institut, 2003.

120. Tim V. Bedsted: A Dynamic Analysis of the Costs of Dementia in Denmark. Februar 2003.
Publ.: *Rød serie* nr. 90, Københavns Universitet. Økonomisk Institut, 2003.

121. Ole Kveiborg: Regulating Road Transport Using Mechanism Design. April 2003.
Publ.: *Rød serie* nr. 91, Københavns Universitet. Økonomisk Institut, 2003.

122. Martin Kaae Jensen: Balanced Growth, Dynamic Stability, and the Turnpike. September 2002.
Publ.: *Rød serie* nr. 92, Københavns Universitet. Økonomisk Institut, 2003.

123. Lina Wøhlk Olsen: Essays on Georg Rasch and his contributions to statistics. Maj 2003.
Publ.: *Rød serie* nr. 93, Københavns Universitet. Økonomisk Institut, 2003.

124. Ninette Pilegaard: Essays in Transport and the Economy. Juni 2003.
Publ.: *Rød serie* nr. 94, Københavns Universitet. Økonomisk Institut, 2003.

125. Kasper Pasternak Jørgensen: Aspects of Innovation and Innovative Industries. September 2003. Publ.: *Rød serie* nr. 95, Københavns Universitet. Økonomisk Institut, 2003.

126. Anne Møller Danø: Empirical Studies of Individual Labour Market Behaviour and Health. November 2003.
Publ.: *Rød serie* nr. 96, Københavns Universitet. Økonomisk Institut, 2003.

127. Lars Pico Geerdsen: Marginalisation Processes in the Danish Labour Market. Januar 2003.
Publ.: *Rapport* nr. 03:24, Socialforskningsinstituttet, 2003.

128. Edith Madsen: Unit Roots and Cointegration in Panel Data Models. Februar 2004.
Publ.: *Rød serie* nr. 97, Københavns Universitet Økonomisk Institut, 2004.

129. Heino Bohn Nielsen: I(1) and I(2) Cointegration Analysis: Theory and Applications. Februar 2004.
Publ.: *Rød serie* nr. 98, Københavns Universitet. Økonomisk Institut, 2004.

130. Henning Tarp Jensen: CGE Modeling, Income Distribution and Agricultural Development with a Focus on Mozambique and Vietnam. Marts 2004.
Publ.: *Rød serie* nr. 99, Københavns Universitet. Økonomisk Institut, 2004.

131. Lars Peter Østerdal: Topics in Game Theory and Health Economics. April 2004.
Publ.: *Rød serie* nr. 100, Københavns Universitet. Økonomisk Institut, 2004.

132. Thomas Barnebeck Andersen: Essays on Financial Factors in Economic Development. Marts 2004. Publ.: *Rød serie* nr. 101, Københavns Universitet. Økonomisk Institut, 2004.

133. Stine Grenaa Jensen: Promoting Renewable Energy Technologies. Juni 2004.
Publ.: *Rød serie* nr. 102, Københavns Universitet. Økonomisk Institut, 2004.

134. Thomas Harr: Essays in Banking Regulation and Financial Crises. April 2004.
Publ.: *Rød serie* nr. 103, Københavns Universitet. Økonomisk Institut, 2004.

135. Jacob L. Weisdorf: On the Road to Riches: Aspects of Economic and Demographic Growth from the Stone Age and Beyond. Juli 2004.
Publ.: *Rød serie* nr. 104, Københavns Universitet. Økonomisk Institut, 2004.

136. Andreas Koch: Five Essays on Economic Theory: Complexity and Coordination. November 2004. Publ.: *Rød serie* nr. 105, Københavns Universitet. Økonomisk Institut, 2004.

137. Paul Lassenius Kramp: Topics in Asymmetric Information and Cascades. Januar 2005.
Publ.: *Rød serie* nr. 106, Københavns Universitet. Økonomisk Institut, 2005.

138. Kristian Schultz Hansen: Using Disability-Adjusted Life Years and Cost-Effectiveness Analysis to Define Priorities for the Public Health Care Sector in Zimbabwe. Februar 2005.
Publ.: *Rød serie* nr. 107, Københavns Universitet. Økonomisk Institut, 2005.

139. Morten Marott Larsen: Essays in Regional and Transport Economics. April 2005.
Publ.: *Rød serie* nr. 108, Københavns Universitet. Økonomisk Institut, 2005.

140. Jesper J. Kühl: Microeconomic Studies on Risk, Poverty and Climate Change. Maj 2005.
Publ.: *Rød serie* nr. 109, Københavns Universitet. Økonomisk Institut, 2005.

141. John Rand: Studies in Economic Development: Business Cycles, Economic Structure and Enterprise Dynamics. Juni 2005.
Publ.: *Rød serie* nr. 110, Københavns Universitet. Økonomisk Institut, 2005.

142. Mette Lunde Christensen: Essays in Empirical Demand Analysis: Evidence from Panel Data. Juni 2005.
Publ.: *Rød serie* nr. 111, Københavns Universitet. Økonomisk Institut, 2005.

143. Jesper Gregers Linaa: Business Cycles and Monetary Policy. September 2005.
Publ.: *Rød serie* nr. 112, Københavns Universitet. Økonomisk Institut, 2005.

144. Søren Gaard: Labor Market Dynamics in Macroeconomics. Juni 2005.
Publ.: *Rød serie* nr. 113, Københavns Universitet. Økonomisk Institut, 2005.

145. Anne Dyrberg Rommer: Accounting-based Credit-scoring Models: Econometric Investigations. November 2005.
Publ.: *Rød serie* nr. 114, Københavns Universitet. Økonomisk Institut, 2005.

146. Steen Winther Blindum: Strict Exogeneity and Nonlinear Panel Data Models with Unobserved Heterogeneity. November 2005.
Publ.: *Rød serie* nr. 115, Københavns Universitet. Økonomisk Institut, 2005.

147. Kasper Meisner Nielsen: Corporate Governance and Performance in Firms with Concentrated Ownership. December 2005.
Publ.: *Rød serie* nr. 116, Københavns Universitet. Økonomisk Institut, 2006.

148. Mette Knudsen: Essays on the Distribution of Incomes: Effects of Monetary, Financial and International Trade Policies. Januar 2006.
Publ.: *Rød serie* nr. 117, Københavns Universitet. Økonomisk Institut, 2006.

149. Mette Gørtz: Leisure, Household Production, Consumption and Economic Well-being. Maj 2006. Publ.: *Rød serie* nr. 118, Københavns Universitet. Økonomisk Institut, 2006.

150. Jens Erik Nielsen: Essays in the Economics of Transport. Marts 2007.
Publ.: *Rød serie* nr. 119, Københavns Universitet. Økonomisk Institut, 2007.

151. Elisabeth Hermann Frederiksen: Inquiries into Economic Growth, Natural Resources, and Labor Allocation. Maj 2007.
Publ.: *Rød serie* nr. 120, Københavns Universitet. Økonomisk Institut, 2007.

152. Mikkel Barslund: Microeconometric Applications in Development Economics. June 2007.
Publ.: *PhD Series* no. 121, Københavns Universitet. Økonomisk Institut, 2008.

153. Martin Junge: Three Essays in Applied Economics: Social Insurance. September 2007.
Publ.: *Rød serie* nr. 122, Københavns Universitet. Økonomisk Institut, 2007.

154.  Birgitte Gersfelt: Developing Country Agriculture and International Trade: Impact and Future Challenges.  April 2007.
Publ.: *Rød serie* nr. 123, Københavns Universitet. Økonomisk Institut, 2007.

155. Thomas Jensen: Topics in Political Economy: Voting, Elections, and Terrorism. November 2007.
Publ.: *PhD Series* no. 124, Department of Economics, University of Copenhagen, 2008.

156. Niels Arne Dam: Essays on Price Formation, Business Cycles, and Monetary Policy. April 2008.
Publ.: *PhD Series* no. 125, Department of Economics, University of Copenhagen, 2008.

157. Steffen Andersen: Individual Preferences and Choice Behavior: From the Laboratory to the Field. August 2007. (Upubl.).

158. Cecilie Dohlmann Weatherall: Education at Workplaces: Long-Term Unemployment, Wages and Enrolment. September 2007.
Publ.: *PhD Series* no. 126, Department of Economics, University of Copenhagen, 2008.

159. Bertel Schjerning: Dynamic Aspects of Entrepreneurial Behavior. April 2008.
Publ.: *PhD Series* no. 127, Department of Economics, University of Copenhagen, 2008.

## PhD Series (formerly Rød serie)
Department of Economics, University of Copenhagen

A list of previously published issues is found in earlier volumes, or at:
http://www.econ.ku.dk/Research/Publications/red/red.default.asp

---

124. Thomas Jensen: Topics in Political Economy: Voting, Elections, and Terrorism. (November  2007). 2008.

125. Niels Arne Dam: Essays on Price Formation, Business Cycles, and Monetary Policy. (April 2008). 2008.

126. Cecilie Dohlmann Weatherall: Education at Workplaces: Long-Term Unemployment, Wages and Enrolment. (September 2007). 2008.

127. Bertel Schjerning: Dynamic Aspects of Entrepreneurial Behavior. (April 2008). 2008.