Essays on Entrepreneurship in Developing Economies

by

Tenzin Yindok

M.A., Simon Fraser University, 2010
B.A., Simon Fraser University, 2008

Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in the Department of Economics Faculty of Arts and Social Sciences

© Tenzin Yindok 2016
SIMON FRASER UNIVERSITY
Summer 2016

All rights reserved.
However, in accordance with the Copyright Act of Canada, this work may be reproduced without authorization under the conditions for “Fair Dealing.” Therefore, limited reproduction of this work for the purposes of private study, research, education, satire, parody, criticism, review and news reporting is likely to be in accordance with the law, particularly if cited appropriately.
## Approval

<table>
<thead>
<tr>
<th>Name</th>
<th>Tenzin Yindok</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degree</td>
<td>Doctor of Philosophy (Economics)</td>
</tr>
<tr>
<td>Title</td>
<td>Essays on Entrepreneurship in Developing Economies</td>
</tr>
</tbody>
</table>

**Examiner Committee:**

- **Chair:** Simon Woodcock  
  Associate Professor
- **Alexander Karaivanov:**  
  Senior Supervisor  
  Professor
- **Brian Krauth:**  
  Supervisor  
  Associate Professor
- **Christian Muris:**  
  Supervisor  
  Assistant Professor
- **Fernando Aragon:**  
  Internal Examiner  
  Assistant Professor  
  Department of Economics
- **Markus Poschke:**  
  External Examiner  
  Associate Professor  
  Department of Economics  
  McGill University

**Date Defended:** June 2 2016
Abstract

I explore the theoretical foundations and the empirical relevance of the idea that entrepreneurship in developing countries could arise out of poor or non-existent alternatives. In the first chapter, I use an occupational choice model to show that the observed large increase in the rate of business ownership in rural Thailand during the 1997 Asian crisis can be explained by a negative shock to the labor market. According to my GMM estimates, a 47% fall in the outside option of entrepreneurship is required to explain the observed increase in business ownership from 17% to 37% between 1997 and 1998. I find that endogenously starting a business enabled households to offset about 40% of the income loss during the crisis, but also that low entrepreneurial productivity limits the extent to which pro-business policies can stand in as unemployment insurance for the average household.

In the second chapter, joint with A. Karaivanov, we explicitly model and distinguish between voluntary and so-called involuntary entrepreneurship, which arises for those who prefer the non-business occupation (e.g., wage work) but cannot obtain it (with some probability that we estimate), due to labor market frictions. We also allow for credit constraints and analyze their interaction with the labor market constraint. We estimate the model via GMM using data from semi-urban Thailand from 2005, and find that 11% of all households in our sample (approximately 17% of all households running a business) are classified as involuntary entrepreneurs. While there are large potential income gains, especially for poorer households, from relaxing either the labor market or credit constraints, involuntary entrepreneurship can only be significantly reduced by addressing the labor market constraint.

In the final chapter, I structurally estimate a model in which risk neutral agents maximize total income by optimally allocating capital and labor into entrepreneurship, subject to credit and time constraints. I estimate the model via GMM using 2005 Thai urban data, where about 20% of business owners report a second occupation. I find that while most entrepreneurs that hold two jobs are skill-constrained (the first-best scale of the business does not exhaust the time-constraint), there is a small fraction that are credit-constrained. These two groups within the multiple occupation group are also predicted to be considerably heterogeneous in terms of initial wealth, schooling and entrepreneurial talent.

**Keywords:** Entrepreneurship, Involuntary entrepreneurship, Multiple occupations
Dedication

I would like to dedicate this to my family.
Acknowledgements

I am grateful to Alex Karaivanov above all.

I am also grateful to Chris Muris and Brian Krauth for their invaluable guidance and constant encouragement, Simon Woodcock for being an amazing graduate chair, and Gwen Wild and Kathleen Vieira-Ribeiro for their assistance throughout the program.

Finally, I am grateful to Raged for proofreading my work every few weeks or months, among other things.
# Table of Contents

Approval ii  
Abstract iii  
Dedication iv  
Acknowledgements v  
Table of Contents vi  
List of Tables ix  
List of Figures xi  
Introduction 1  

1 Entrepreneurship During an Economic Crisis - Evidence from Rural Thailand 3  
1.1 Introduction .................................................. 3  
1.2 Model ....................................................... 8  
1.2.1 Occupational choice ..................................... 8  
1.2.2 Exogenous shock to the outside option of entrepreneurship .................. 10  
1.2.3 Occupational choice probabilities ........................ 12  
1.3 Data ......................................................... 12  
1.3.1 The Townsend Thai Project ................................ 12  
1.3.2 Evidence of labor market shock ......................... 14  
1.3.3 Household borrowing and loan characteristics during the crisis ............ 17  
1.4 Structural estimation ........................................ 19  
1.4.1 GMM - matched moments and computation ................. 19  
1.4.2 Estimates and model fit .................................. 21  
1.4.3 Analysis .................................................. 24  
1.4.4 Robustness checks ...................................... 34  
1.5 Conclusions .................................................. 36
List of Tables

Table 1.1  Summary statistics ......... 13
Table 1.2  Fixed effects regressions of business ownership .......... 16
Table 1.3  Borrowing and loan characteristics .................. 18
Table 1.4  List of matched moments .......................... 21
Table 1.5  Structural estimates and model fit .................. 23
Table 1.6  Predicted change in income during the crisis .......... 26
Table 1.7  Robustness checks ......................... 35

Table 2.1  Voluntary and involuntary entrepreneurship ........... 48
Table 2.2  Occupation and source of income ..................... 50
Table 2.3  Summary statistics ............................... 51
Table 2.4  Probit regression of business ownership ............... 52
Table 2.5  List of matched moments .......................... 55
Table 2.6  Structural estimates and model fit .................. 56
Table 2.7  Model predictions at the GMM estimates .............. 56
Table 2.8  Model predictions, distribution of entrepreneurs by type 57
Table 2.9  Model fit: matched moments at GMM estimates .......... 58
Table 2.10 Model fit: non-matched moments at GMM estimates ...... 59
Table 2.11 Relaxing the labor or credit constraints .............. 60
Table 2.12 Policy counterfactual - microfinance ................. 64
Table 2.13 Robustness checks ............................... 65

Table 3.1  Summary statistics ............................... 79
Table 3.2  Reduced-form regressions of multiple occupations ....... 80
Table 3.3  Types of businesses among multiple occupation holders 81
Table 3.4  Types of non-business occupations among multiple occupation holders 81
Table 3.5  List of matched moments .......................... 83
Table 3.6  Structural estimates and model fit .................. 85
Table 3.7  Model predictions at GMM estimates .................. 89
Table 3.8  Counterfactual analysis ........................... 90

Table A.1 Business ownership over the years in Townsend Thai Data .... 108
Table A.2 Sources of household income .......................... 108
Table A.3 Loan characteristics
### List of Figures

<table>
<thead>
<tr>
<th>Figure Reference</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>Occupational choice in $z, \theta$ plane (for a fixed $x$)</td>
<td>11</td>
</tr>
<tr>
<td>Figure 1.2</td>
<td>Impact of a decrease in $\mu$ between 1997 and 1998</td>
<td>11</td>
</tr>
<tr>
<td>Figure 1.3</td>
<td>Occupational proportions as function of wealth</td>
<td>25</td>
</tr>
<tr>
<td>Figure 1.4</td>
<td>Percentage income-loss during the crisis</td>
<td>27</td>
</tr>
<tr>
<td>Figure 1.5</td>
<td>Effect on entrepreneurship in 1998 - increase in $\lambda$</td>
<td>29</td>
</tr>
<tr>
<td>Figure 1.6</td>
<td>Percentage gain in income in 1998 - increase in $\lambda$</td>
<td>29</td>
</tr>
<tr>
<td>Figure 1.7</td>
<td>Percentage gain in income in 1998 as function of wealth - $2\lambda$</td>
<td>29</td>
</tr>
<tr>
<td>Figure 1.8</td>
<td>Effect on entrepreneurship in 1998 - increase in $\theta$</td>
<td>30</td>
</tr>
<tr>
<td>Figure 1.9</td>
<td>Percentage gain in income in 1998 - increase in $\theta$</td>
<td>30</td>
</tr>
<tr>
<td>Figure 1.10</td>
<td>Percentage gain in income in 1998 as function of wealth - 10% increase in $\theta$</td>
<td>30</td>
</tr>
<tr>
<td>Figure 1.11</td>
<td>Effect on entrepreneurship in 1998 - transfers</td>
<td>32</td>
</tr>
<tr>
<td>Figure 1.12</td>
<td>Percentage gain in median income in 1998 - transfers</td>
<td>32</td>
</tr>
<tr>
<td>Figure 1.13</td>
<td>Percentage gain in income in 1998 as function of wealth - transfers</td>
<td>32</td>
</tr>
<tr>
<td>Figure 2.1</td>
<td>Probability of entrepreneurship as function of wealth</td>
<td>58</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>Probability of entrepreneurship - model vs. data</td>
<td>59</td>
</tr>
<tr>
<td>Figure 2.3</td>
<td>Expected income gains - relaxing the labor or credit constraints</td>
<td>63</td>
</tr>
<tr>
<td>Figure 2.4</td>
<td>Income gains by wealth and schooling - microfinance</td>
<td>66</td>
</tr>
<tr>
<td>Figure 3.1</td>
<td>Occupational map in $(z, \theta)$ plane for a given level of $x$</td>
<td>75</td>
</tr>
<tr>
<td>Figure 3.2</td>
<td>Model fit - unmatched moments</td>
<td>86</td>
</tr>
<tr>
<td>Figure 3.3</td>
<td>Predicted probability of multiple occupation as function of wealth</td>
<td>87</td>
</tr>
<tr>
<td>Figure 3.4</td>
<td>Percentage gain in income - eliminating the credit constraint</td>
<td>91</td>
</tr>
<tr>
<td>Figure 3.5</td>
<td>Percentage gain in income - increasing entrepreneurial productivity (↑ $\theta$ by 10%)</td>
<td>92</td>
</tr>
</tbody>
</table>
Introduction

Why do people become entrepreneurs? This question is especially important in developing countries where one-third to one-half of the non-agricultural work-force is estimated to be self-employed (Gindling and Newhouse, 2012; de Mel et al., 2010). In the following three chapters, I analyze the causes and implications of entrepreneurship in developing countries, with a particular focus on the role of the outside options. A key motivation of my research is addressing the somewhat simplistic assertion that more entrepreneurial activity necessarily indicates better economic conditions for business ownership.

A fundamental assumption in economics is that of revealed preferences. The theory of occupational choice is predicated on the observed choice being better than the next best alternative, and changes in the alternatives faced by individuals and households are counterfactuals. Therefore a key requirement for my research is to find meaningful ways to highlight or measure the lack of alternatives.

In the first chapter, I analyze entrepreneurship during a period of weak employment demand in the labor market. Using an occupational choice model, I allow entrepreneurship to arise endogenously due to lower outside options and evaluate the effectiveness of entrepreneurship as a coping strategy. To identify the income mitigation that starting a business provides, we would generally need to know the size of the outside-option shock, and additionally, the counterfactual incomes of those who start a business. I rely on the structural estimation of a two-period extension of the occupational choice model of Evans and Jovanovic (1989), to get an estimate of this outside option shock.

In the second chapter, joint with A. Karaivanov, we extend the baseline Evans and Jovanovic model to allow for the possibility of so-called involuntary entrepreneurship. We explicitly model and estimate a probability that an agent does not have access to employment elsewhere, and define involuntary entrepreneurs as entrepreneurs who could earn a higher income in a wage job, but cannot obtain it due to a labor market constraint. We parameterize the model in such a way that different values of a parameter correspond to different levels of involuntary entrepreneurship, and in particular, allows the data to reveal whether the choice constraint is negligible or significant.

In the third chapter, I analyze entrepreneurship in a framework that allows skill-constrained individuals to take on a second occupation. Again, in an extension of Evans and Jovanovic,
I relax the assumption that occupational choice is a binary and a mutually exclusive choice between entrepreneurship and alternative occupations. I define skill-constrained multiple occupation holders as those who take on a second occupation despite being able to operate their business at the first-best scale. While such individuals are not credit-constrained, their presence could be indicative of underdeveloped labor markets that are unable to employ low-talent entrepreneurs in full-time occupations.

In a nutshell, I identify an issue related to entrepreneurship that the literature has not explored sufficiently in each chapter, use an appropriate extension of an overarching theoretical framework for occupational choice, and estimate the parameters of each model via GMM. The data for my chapters come from the rural and urban surveys of the Townsend Thai Data (Townsend, 1997; Townsend, 1998; NORC, 2008). I then use the results to analyze the causes and consequences of entrepreneurship, and finally conduct policy counterfactuals.
Chapter 1

Entrepreneurship During an Economic Crisis - Evidence from Rural Thailand

1.1 Introduction

In a survey of rural households in Thailand (Townsend, 1997; Townsend, 1998), household business ownership increased dramatically from 16.5% to 37% during the Asian financial crisis in 1997. At the peak of the crisis in 1998, the rate of increase in entrepreneurship was more than six times the rate in 1997 and 1999-2001. This paper explores and empirically evaluates the idea that a significant share of business ownership in developing countries can be driven by the low value of alternative occupations available to households.

I use a structural model of occupational choice to estimate the responsiveness of entrepreneurship to a change in the outside option, using household survey data gathered during the 1997 Asian crisis. In addition, I quantify the extent to which starting a business mitigates an income shock of this nature. To the best of my knowledge, this is the first paper to evaluate the impact of the Asian crisis on household business ownership using a structural estimation approach.

Specifically, I propose a two-period extension of the occupational choice model in Evans and Jovanovic (1989). Households have the option of earning business income using capital and entrepreneurial talent, or can work elsewhere for an income that is increasing in labor market credentials (for example, education). Each household chooses the occupation that maximizes its income. There is a borrowing limit for business investment that is assumed to be a fixed multiple of the household’s initial wealth, so that those with a higher starting wealth are able to invest more. I extend this framework in two ways in order to analyze the effect of the Asian crisis on entrepreneurship. First, I consider occupational choice over two periods which allows me to study the response of households to the crisis. Second, I allow
for an exogenous aggregate shock to the outside option (non-business income) to model the impact of the crisis. The resulting model yields a joint probability of occupational choices, before and after the crisis, conditional on observing education (assumed unchanged) and initial wealth in each year, and model parameters. I use the generalized method of moments to estimate the parameters of the model using data from the Townsend Thai Project’s annual rural surveys.

I find that the rate of entrepreneurship is very responsive to the outside option. At the GMM estimates, a 47% fall in the outside option of entrepreneurship is required to explain the observed 124% increase in business ownership in the data, implying an elasticity of 2.6. Consequently, average income in the sample is predicted to fall by 17% due to the crisis, with close to a quarter of the sample endogenously switching into business ownership between the pre-crisis year, 1997, and the post-crisis year, 1998. Conditional on endogenously starting a business during the crisis, the decrease in new entrepreneurs’ average income is approximately 13.4%, about 40% lower than the fall in their average income in the counterfactual where occupations are fixed at their pre-crisis assignments.

I also apply my structural model to compare quantitatively the effectiveness of several counterfactual policies that could be used to mitigate the income shock in the post-crisis period. Since households maximize their income in my model, I look at the effect of these policies on household income to assess their welfare implications. First, I find that relaxing the credit constraint by increasing the borrowing limit in the post-crisis economy does not affect occupational choice. In the counterfactual where all households can afford the optimal level of business investment, the predicted increase in the 10th percentile of income is about 10% (with a close to zero change in median and average income), illustrating the limited scope of policies that relax the credit constraint. This finding is consistent with the recent literature on the effectiveness of microcredit. For example, Kaboski and Townsend (2012) find that a credit injection in rural Thailand had statistically insignificant effects on business start-ups and investment. A large number of randomized control trials similarly find modest impacts of microcredit on business (and consumption) outcomes.

Second, I find that in contrast to the credit policies, increases in entrepreneurial skill (for example, through business training) cause a steady increase in business start-ups and income, implying that lack of productivity is the constraint that binds faster than the credit constraint. Banerjee and coauthors (2015) similarly state that the average business run by those who start a business due to microcredit interventions might not be profitable, “given the skill sets of the entrepreneurs and their life situations” (p52). While a substantial increase in income is predicted if entrepreneurs could become more productive, business

---

1 As I will argue later, a worsening labor market during the crisis is consistent with reduced-form findings and the prior literature on the Asian crisis.

training or similar programs are likely to leave out those who are the least talented (who do not make enough as business owners even with some business training), or the poorest (who are unable to invest the higher level of desired capital following an increase in productivity). For households that are poor and less talented, both entrepreneurial skill and credit are binding constraints, necessitating a combination and coordination of policies to help them.

I also consider the effect of three kinds of monetary transfers to households - wage guarantees to non-business households, grants for business investment, and unconditional income transfers - taking into account the endogenous occupational choices. Households on average gain the most from unconditional transfers, followed by wage guarantees. I find that providing grants solely for business investment is a particularly wasteful way to offset income losses during the crisis – the marginal return from investing in business for non-business households in 1998 is predicted to be less than the opportunity cost of funds.

The key hypothesis in my model is that during the 1997 Asian crisis, households in rural Thailand used business startups to respond to the shock to their job alternatives. I report evidence in support of this hypothesis from the prior literature pertaining to the crisis, and directly among the households I study in this paper. It is well-documented that the Asian crisis severely affected the corporate sector, real-estate and the financial sector in urban Thailand, and spread to rural areas through the channels of migrant-work, remittance and rural-urban labor market linkages (Bresciani et al., 2002; Chandoewit, 2010; Siamwalla, 2011). For example, the number of temporary migrants returning home from Bangkok increased by 37% in the first quarter of 1998 (World Bank, 2000). The Northeastern region in particular suffered the most from crisis-induced unemployment, as the region was the largest source of rural-urban migration (Chandoewit, 2010; Kittiprapas, 2002). Accordingly, Subhadhira et al. (2002) reports that almost two million migrant workers may have returned to the Northeast, and that the region had the highest dry-season unemployment rate (8.6%) in the country in February 1998. None of these papers however quantifies and systematically analyzes the effect of this implied decrease in the value of the outside option of entrepreneurship (in the form of unemployment or wage reductions) on business ownership and startups in rural areas.

Consistent with the findings in the papers above, and in further support of the mechanism I model, I find two key results using fixed effects regressions. First, in the central provinces closer to Bangkok, households that earned wage-income before the onset of the crisis (that is, those which are affected by the hypothesized fall in the outside option) were 14.8 percentage points more likely to start a business during the crisis. Relative to the pre-crisis rate of business ownership in this region, this effect represents about a 70% increase in business ownership. Second, in the Northeast, households that experienced an increase in business ownership and startups in rural areas.

---

3A real world counterpart of policy that guarantees non-business income is the National Rural Employment Guarantee Act of India, that guarantees a hundred days of wage-employment in a year in unskilled manual work to an eligible rural household.
in male household members (presumably due to reverse migration from the urban centers) were 11.7 percentage points more likely to start a business during the crisis. Relative to the pre-crisis rate of business ownership in the region, this effect represents an increase of 91% in the rate of business ownership.

Finally, I provide evidence that the increase in business ownership is unlikely to be caused by a change in the credit market for business investment. Paulson and Townsend (2006) argue that in explaining the rise in business ownership in rural and semi-urban areas, it is “difficult to imagine that imperfections in these financial markets were somehow alleviated during the crisis period” (p35).\(^4\) Taking advantage of the relatively detailed account of household borrowing in the Townsend data set, I find direct evidence that there was no change on average in formal borrowing (from the Bank for Agriculture and Agriculture Cooperatives, commercial banks and village funds). While a higher proportion of households borrowed from relative and neighbors, these loans were mostly for the purpose of consumption. Excluding loans from friends and neighbors, there is also no significant change in the average loan size, interest rate and loan duration during the period. All this suggests that the observed increase in business ownership is very unlikely to be due to a relaxation in financial constraints.

Related literature

Small-scale entrepreneurship in developing countries has received considerable attention as a means to raise incomes and reduce poverty. When entrepreneurship is accessible only to those with sufficient skill and starting capital, a high level of enterprise could indicate an abundance of skill or well-performing credit markets. For country-specific evidence on higher incomes of entrepreneurs, see House et al. (1993) on Kenya, Mohapatra et al. (2007) on China, and Bradford (2003) and Quadrini (2000) on the United States. Alternatively, some authors have argued that a growing rate of entrepreneurship is not necessarily a sign of development (Pietrobelli et al., 2004), and could instead be symptomatic of necessity (Banerjee and Duflo, 2007; Poshke, 2012) or the national unemployment cycle (Koellinger and Thurik, 2012). Classical occupational choice models are able to reconcile these competing views by allowing heterogeneity along the lines of skill, wealth and outside options (Evans and Jovanovic, 1989; Banerjee and Newman, 1993; Lloyd-Ellis and Bernhardt, 2000).\(^5\) Among papers that structurally estimate such models in the context of developing countries, Karaivanov (2012) and Paulson et al. (2006) use the Townsend Thai data but

\(^4\)At the other end, despite the documented fact that the corporate sector and urban SMEs (small and medium-sized enterprises) faced liquidity constraints during the crisis, none of the studies suggest that a contagion to rural credit markets was a possible channel of transmission. In particular, Bresciani et al. (2002) conclude that there is no evidence of a rural credit crunch and that for rural areas, “small-scale business activities ... were less affected by the recession” (p12).

focus on uncovering the underlying cause of credit market failures. I apply similar methodology (with a simpler structure for the credit constraint), but instead allow the outside option to change over time according to macroeconomic conditions, unlike in the papers mentioned above, where the outside option is held fixed. Additionally, in contrast to these papers, I match the joint probability of occupational choices for a given household in two periods (pre- and post-crisis), rather than the probability of business ownership for a single year.\(^6\)

The only other paper which studies business ownership during the Asian crisis in Thailand is Paulson and Townsend (2005). It uses the same data source (the rural surveys from Townsend Thai Data) and performs reduced-form regressions of business ownership separately for pre-crisis, crisis and post-crisis business owners. The explanatory variables include past wealth, household head’s age and schooling, household composition and association with various lending institutions. Their main finding is that the correlation between past-wealth and business start-ups falls and becomes statistically insignificant during the crisis. In contrast, I build and estimate an explicit structural model that can generate this phenomenon. Specifically, I assume that the marginal product of business capital is higher everywhere for more talented entrepreneurs, and therefore, everything else equal, more talented business owners would want to invest more. However, due to the borrowing constraint that makes maximum investment a function of initial wealth, richer households are more likely to afford their desired level of investment and earn higher entrepreneurial income. When alternative occupations are more attractive (i.e., in the pre-crisis period), the reservation entrepreneurial income is higher. This also implies that the required level of entrepreneurial talent which makes entrepreneurship the preferred occupation is higher. The correlation between past-wealth and business ownership is consequently stronger among observed business owners because (a) more talented households require more investment, and (b) only wealthier households will be able to afford it. Crisis-induced unemployment and a fall in wages, which I model as a fall in the outside option of business ownership, therefore lowers the reservation entrepreneurial income and the correlation between past-wealth and entrepreneurship.

In related work, Adhvaryu et al. (2014) uses fluctuations in international coffee prices to estimate the causal effect of a decrease in household income on entrepreneurship among coffee farmers in Tanzania. Similarly, I use the financial crisis as an exogenous aggregate shock to the outside option of entrepreneurship. In addition to finding a negative effect of international coffee prices on business ownership, Adhvaryu and coauthors also find that businesses that are started to weather shocks earn lower profits. They suggest that those who start a business purely to mitigate a fall in income could be fundamentally different in terms of skills and access to capital. My model can generate this scenario, where lower

\(^6\) Another difference is that these papers use maximum likelihood estimation while I use the generalized method of moments.
outside options cause households who are less skilled and/or poorer (and therefore with less access to capital) to start a business. The structural estimates further allow me to quantitatively evaluate the various suggested policy prescriptions as I described previously.

The paper is organized as follows. Section 1.2 sets up the model from which the probabilities of occupational choices are derived. In Section 1.3, I describe the main features of occupational choice and credit conditions in the Townsend Thai data and present results from fixed-effects regressions of business ownership to provide evidence for the hypothesis that households started a business following a fall in the outside option of entrepreneurship during the crisis. Structural estimation of the model parameters and an analysis of the results, including counterfactual analysis is reported in Section 1.4, followed by robustness exercises in Section 1.4.4, before concluding the paper in Section 1.5.

1.2 Model

1.2.1 Occupational choice

Consider a two-period extension of the static occupational choice model in Evans and Jovanovic (1989), hereafter denoted by \textit{EJ}. Assume that a risk-neutral household maximizes expected income, is endowed with initial wealth \(z_t\), labor market characteristic \(x_t\), and entrepreneurial talent \(\theta_t\), with \(x\) and \(\theta\) are assumed to be time-invariant. Output from entrepreneurship is a function of talent \(\theta\) and capital investment \(k_t\), and a time-invariant stochastic term \(\vartheta\) with expected value equal to one:

\[
q^E(\theta, k_t) = \theta k_t^\alpha \vartheta.
\]

The household can borrow and save at a risk-free gross interest rate \(r\), and only a positive multiple of the beginning-of-period wealth \(z_t\) can be invested in the business in period \(t\), such that\(^7\)

\[
k_t \leq \lambda z_t. \tag{1.1}
\]

The expected end-of-period wealth from choosing entrepreneurship in period \(t\) is

\[
y^E(\theta, z_t) = \arg \max_k \{\theta k_t^\alpha + r(z_t - k_t)\}. \tag{1.2}
\]

Assume that the capital stock depreciates completely each period and the household reinvests the next period, subject to the credit constraint that capital stock is less than \(\lambda\) times the new beginning-of-period wealth.\(^8\) Income from entrepreneurship will depend on

\(^7\)This credit-constraint is derived as an endogenous outcome in a limited enforcement model in Appendix A.1. I do not restrict \(\lambda\) to be bigger than one because depending on the liquidity of the asset, the investment limit could be lower than the value of the asset.

\(^8\)Although the model contains two periods, it is static in the sense that there is no explicit connection between this period’s endogenous occupational choice (and income) and next period’s capital stock. In
whether the household is able to invest $k_u(\theta)$, the optimal amount of capital that maximizes expected profit from entrepreneurship:

$$k_u(\theta) = \arg \max_k \{\theta k^\alpha - rk\} = \left(\frac{\theta \alpha}{r}\right)^{\frac{1}{1-\alpha}}.$$ (1.3)

Since the household can only invest up to $\lambda z_t$ (due to the credit constraint from equation 3.1), $k_u(\theta)$ is only feasible if $k_u(\theta) = \left(\frac{\theta \alpha}{r}\right)^{\frac{1}{1-\alpha}} \leq \lambda z_t$. Define $B(z_t)$ as the level of $\theta$ at which $k_u(\theta) = \lambda z_t$:

$$B(z_t) \equiv \frac{r}{\alpha} (\lambda z_t)^{1-\alpha}.$$ (1.4)

This is the level of talent at which the household can just about finance its first-best level of capital. Since $k_u(\theta)$ is increasing in $\theta$, the more talented the household, the more it would want to invest in capital. Therefore, for a given level of $z_t$, the credit constraint is more likely to bind if the household is more talented. The marginal product of capital is larger than its marginal cost in this case and the best the household can do is to invest $\lambda z_t$.

Therefore

$$y^E(\theta, z_t) = \begin{cases} \theta [k_u(\theta)]^\alpha + r(z_t - k_u(\theta)) & \text{if } \theta \leq B(z_t), \\ \theta (\lambda z_t)^\alpha + rz_t (1 - \lambda) & \text{if } \theta > B(z_t). \end{cases}$$

$$\Rightarrow y^E(\theta, z_t) - rz_t = \begin{cases} (1 - \alpha) \theta^{\frac{1}{1-\alpha}} \left(\frac{\alpha}{r}\right)^{\frac{\alpha}{1-\alpha}} & \text{if } \theta \leq B(z_t), \\ \theta (\lambda z_t)^\alpha - \lambda rz_t & \text{if } \theta > B(z_t). \end{cases}$$

Households earn $q^A(x) = \mu (1 + x)^\gamma \xi$ in the alternative occupation, where $x$ indicates labor market characteristics, $\mu$ is a parameter that captures the expected alternative income if $x$ were equal to zero, and $\xi$ is a stochastic term with $E(\xi|x) = 1$. The alternative output does not depend on entrepreneurial talent and capital, and encapsulates the outside option to entrepreneurship, which could be wage-work, subsistence farming or non-subsistence farming. The total expected end-of-period income of the household from choosing this occupation is

$$y^A(z_t, x) = \mu (1 + x)^\gamma + rz_t.$$ (1.5)

The household runs a business in period $t$ if the expected net income from doing so exceeds the expected alternative income in period $t$. That is,

$$E_t = 1 \quad \text{if } y^E(\theta, z_t) \geq y^A(z_t, x).$$

Define $\pi(z_t, \theta, x) = y^E(\theta, z_t) - y^A(z_t)$ as the expected income differential between entrepreneurship and its alternative. Noting that $rz_t$ cancels out, the household chooses to particular, because of the absence of consumption in the model, the end-of-period wealth in time $t$ is not connected to the beginning-of-period wealth in $t + 1$. When I estimate the model in Section 1.4, I use the one-year lag of household wealth as a measure of $z_t$ in each year.
run a business only if the additional income from business ownership is non-negative.

\[
\pi(z_t, \theta, x) \geq 0 \iff \begin{cases} (1 - \alpha) \theta \left( \frac{1}{\alpha} \right) (\frac{r}{\alpha}) \gamma \geq 0 & \text{if } \theta \leq B(z_t), \\ \theta (\lambda z_t)^{\alpha} - r \lambda z_t - \mu (1 + x)^{\gamma} \geq 0 & \text{if } \theta > B(z_t). \end{cases}
\]

\[= \begin{cases} \theta \geq \left( \frac{\mu}{1 - \alpha} \right) (1 + x)^{\gamma} \left( \frac{r}{\alpha} \right)^{\alpha} & \text{if } \theta \leq B(z_t), \\ \theta \geq (\lambda z_t)^{-\alpha} [\mu (1 + x)^{\gamma} + r \lambda z_t] & \text{if } \theta > B(z_t). \end{cases}
\]

Define the following threshold values of \( \theta \) for a given \((x, z_t)\):

\[A(x) = (\frac{\mu}{1 - \alpha}) (1 + x)^{\gamma} \left( \frac{r}{\alpha} \right)^{\alpha}, \quad (1.7)\]

\[C(z_t, x) = (\lambda z_t)^{-\alpha} [\mu (1 + x)^{\gamma} + r \lambda z_t]. \quad (1.8)\]

Occupational choice at time \( t \) for a household with characteristics \((z_t, x, \theta)\) can be summarized as:

\[E_t = 1 \iff \pi(z_t, \theta, x) \geq 0 \iff \begin{cases} \theta \geq A(x) & \text{if } \theta \leq B(z_t), \\ \theta \geq C(x, z_t) & \text{if } \theta > B(z_t). \end{cases} \quad (1.9)\]

Figure 1.1 shows the partition of the \((z, \theta)\)-plane into constrained entrepreneurs, unconstrained entrepreneurs, and non-entrepreneurs for a fixed level of \( x \). The line \( B(z) \) separates the plane into constrained and unconstrained (potential) entrepreneurs. The occupational indifference curve is given by \( A(x) \) when \( \theta \leq B(z) \), and by \( C(z, x) \) when \( \theta > B(z) \).

### 1.2.2 Exogenous shock to the outside option of entrepreneurship

I extend the basic framework developed in the model in \( EJ \) by allowing the parameter \( \mu \) to be different between periods \( t \) and \( t + 1 \) and derive the implications of this change on two-period occupational choice of the household. As an illustration, assume that household characteristics \((\theta, z, x)\) remain constant over time.\(^9\) Additionally, assume that the outside option parameter \( \mu \) decreases between the two periods. Figure 1.2 shows that a fall in \( \mu \) causes a downward shift in the occupational indifference curve. In other words, the minimum level of talent required to find entrepreneurship profitable is lower when the value of the outside option is lower. A household with \((\theta, z, x)\) is predicted by the model to be a non-business owner in period \( t \) and a business owner in \( t + 1 \) if

\[
\begin{align*}
\theta &< B(z) \text{ and } \theta < A(x; \mu_t) \text{ and } \theta > A(x; \mu_{t+1}) \text{ or } \\
\theta &> B(z) \text{ and } \theta < C(z, x; \mu_t) \text{ and } \theta > C(z, x; \mu_{t+1})
\end{align*}
\]

where I make the dependency of \( A(x) \) and \( C(z, x) \) on \( \mu \) explicit. In Figure 1.2, this translates to the area between the two occupational indifference curves for \( t = 1997 \) and \( t + 1 = 1998 \).

\(^9\)For the actual structural estimation in Section 1.4, beginning-of-period wealth \( z \) is allowed to be different each year.
Figure 1.1: Occupational choice in $z, \theta$ plane (for a fixed $x$)

Figure 1.2: Impact of a decrease in $\mu$ between 1997 and 1998
The area above the 1997 occupational indifference curve represents those that are in business in both years, and the area below the 1998 occupational indifference curve represents those that are not in business in both years.

1.2.3 Occupational choice probabilities

For the model to predict the probability of starting a business during the crisis, a distribution for talent needs to be specified (although households in the model know their level of talent \( \theta \), it is unobserved by the econometrician). Following the literature, I assume a log normal distribution for \( \theta \), and allow it to be correlated with initial wealth \( z \) in 1997, and schooling \( x \).\(^{10}\) Defining \( \bar{\theta} \) as the conditional mean of \( \ln \theta \) and \( \varepsilon \) as a normally distributed variable with mean zero and variance \( \sigma^2 \),

\[
\ln \theta = \bar{\theta} + \varepsilon, \quad \bar{\theta} \equiv \delta_0 + \delta_1 \ln(z) + \delta_2 \ln(1 + x), \quad \varepsilon|z, x \sim N(0, \sigma^2). \tag{1.10}
\]

With the distributional assumption on \( \theta \) and conditional on \( z_t \) and \( x \), in the illustrative example described above, the probability that \( \theta \) is less than \( B(z_t) \) and greater than \( A(x; \mu_{t+1}) \) at time \( t \) is given by

\[
P(A(x; \mu_{t+1}) < \theta < B(z_t)) = P\left(\frac{\ln A(x; \mu_{t+1}) - \bar{\theta}}{\sigma} < \frac{\varepsilon}{\sigma} < \frac{\ln B(z_t) - \bar{\theta}}{\sigma}\right).
\]

The probability that \( \theta \) falls between any two thresholds of \( \theta \) can be similarly specified. I derive the predicted probabilities of occupational choice of a household in two periods, conditional on observing household characteristics \( (z_t, z_s, x) \) for two different time periods \( t \neq s \), and on the assumption that \( \mu_t \neq \mu_s \). See Appendix A.2 for a detailed derivation of these probabilities and additional model moments, that will be used in Section 1.4 for structural estimation.

1.3 Data

1.3.1 The Townsend Thai Project

I use data from a longitudinal survey of Thai households started in July of 1997 (two months before the onset of the Asian crisis) by the Townsend Thai Project (Townsend, 1997; Townsend, 1998). The surveys cover households in four provinces located in two distinct regions in Thailand - the richer and more developed central provinces of Chachoengsao and Lopburi closer to Bangkok, and the poorer northeastern provinces of Buriram and Sisaket.\(^{10}\)In some papers, talent is also allowed to be correlated with initial wealth. I explore this issue in the Section 1.4.4.
Table 1.1: Summary statistics

<table>
<thead>
<tr>
<th></th>
<th>1997</th>
<th></th>
<th>1998</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Business</td>
<td>Non-business</td>
<td>Business</td>
<td>Non-business</td>
</tr>
<tr>
<td>one-year lag of wealth $z_t$</td>
<td>1283.0*</td>
<td>678.0*</td>
<td>1083.6*</td>
<td>647.7*</td>
</tr>
<tr>
<td></td>
<td>(2041.8)</td>
<td>(1286.8)</td>
<td>(1760.9)</td>
<td>(1219.1)</td>
</tr>
<tr>
<td></td>
<td>[504.4]</td>
<td>[233.8]</td>
<td>[439.8]</td>
<td>[234.3]</td>
</tr>
<tr>
<td>years of schooling (head) $x$</td>
<td>4.58*</td>
<td>3.85*</td>
<td>4.25</td>
<td>3.80</td>
</tr>
<tr>
<td></td>
<td>(2.60)</td>
<td>(2.77)</td>
<td>(2.71)</td>
<td>(2.77)</td>
</tr>
<tr>
<td>gross business income $q_t^E$</td>
<td>212.3</td>
<td>-</td>
<td>133.7</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(457.0)</td>
<td>-</td>
<td>(378.8)</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>[49.0]</td>
<td>-</td>
<td>[26.9]</td>
<td>-</td>
</tr>
<tr>
<td>gross total income</td>
<td>309.2*</td>
<td>74.7*</td>
<td>221.9*</td>
<td>70.0*</td>
</tr>
<tr>
<td></td>
<td>(507.5)</td>
<td>(102.8)</td>
<td>(492.7)</td>
<td>(96.8)</td>
</tr>
<tr>
<td></td>
<td>[144.2]</td>
<td>[46.0]</td>
<td>[100.8]</td>
<td>[40.6]</td>
</tr>
<tr>
<td>percent of sample</td>
<td>14.4%*</td>
<td>85.6%*</td>
<td>33.7%*</td>
<td>66.3%*</td>
</tr>
<tr>
<td>sample size</td>
<td>120</td>
<td>713</td>
<td>281</td>
<td>520</td>
</tr>
</tbody>
</table>

Mean, standard deviation (parentheses) and median (brackets) are reported respectively. Monetary values are reported in 000s of 1997 Thai baht. (*) indicates that the difference-in-means test between business and non-business is significant at the 10% level in a given year. Households in the top percentile of the wealth distribution in 1997, with missing data on household head’s schooling, and business households with zero business income are dropped from the sample.

The initial survey in 1997 covered 2880 households. A subset of these households was included in the longitudinal survey.\(^{11}\)

I use data on household business ownership, annual gross income, wealth and years of schooling for structural estimation in Section 1.4. I use additional data on household composition and demography, and various indicators of occupational status for reduced-form regressions in Section 1.3.2. Here, I define how I measure (or proxy for) each of the key variables from the model that are directly used to estimate the structural parameters, and leave the additional variable definitions for reduced-form analyses to be explained in Section 1.3.2.

A household is defined to be a business owner in year $t$ if any member of the household reports owning a business at the time of the survey. Initial wealth in each time period, $z_t$, is measured as the one-year lag of the monetary value of total assets owned in the form of land, household durables and agricultural assets.\(^{12}\) Since the 1997 survey was fielded right before

---

\(^{11}\) About 910 households are observed in both 1997 and 1998. The median wealth of households in the resurvey is 3.5% lower than the median wealth in the initial survey. Pre-crisis business ownership is also lower among the resurvey; 17.7% compared to 21.9% in the initial survey.

\(^{12}\) Household durables include items such as television, VCR, telephone, refrigerator, washing machine, stereo and other items that cost at least 1000 baht. Agricultural assets include tractors, machines, storage and livestock buildings and other items that cost at least 1000 baht. Since the survey asks for the purchase price of household durables and agricultural assets, the values are depreciated annually at 10% from the
the onset of the crisis, this does not capture any decrease in wealth that could have occurred right before the occupational choice is made in 1998. About 88% of the sample households did not lose any land, and 97% did not acquire any additional land between the two years. The assumption that the financial crisis did not cause a significant decrease in the stock of household wealth in the sample therefore seems reasonable. Labor market characteristics \( x \) is measured as the years of schooling of the head of the household in 1997. In addition to occupational proportions in 1997 and 1998, I use moments related to business income \( q_t^E \) to help identify the structural parameters of the model. As a proxy for this variable, I use gross business income of business owners.\(^{13}\)

The sample used for structural estimation is restricted to households that are observed both before and after the crisis, respectively in the 1997 and 1998 waves of the survey. In addition, I exclude households with zero or negative wealth, zero income, households in the top percentile of the wealth distribution in 1997, and in the top percentile of the gross-income distribution in both years. I also drop households with missing information on years of schooling of the head of the household. Table 1.1 summarizes the variables I use in structural estimation - the one-year lag of wealth, years of schooling of head of household, gross business income, and gross total income - separately for business and non-business owners for 1997 and 1998.

### 1.3.2 Evidence of labor market shock

In Section 1.2, I predict the probability of occupational choices in 1997 and 1998 assuming that the outside option parameter \( \mu \) in the model changes during the crisis, and based on this assumption, I structurally estimate the model parameters in Section 1.4. To motivate this assumption (that the observed increase in household business ownership was due to a labor market shock to the alternatives available to business ownership), I compare the characteristics of households that started a business during the crisis to those that did not do so. Specifically, I run fixed-effects regressions to examine if the increase in business ownership during the crisis can be associated with an adverse shock in the labor market.

Consider the following regression equation:

\[
E_{it} = \beta_0 + \beta_1 O_i D_t + x_{it} \beta_2 + \beta_3 D_t + c_i + u_{it}.
\]

The dependent variable, \( E_{it} \), is an indicator that takes value 1 if household \( i \) is in business in year \( t \), and assigns a household to business ownership if it reports that one or more

---

\(^{13}\)Since business income is observed only for business households, I predict average business income conditional on business ownership to match average business income in the data. As a robustness check, I also use annual household gross income, exclusive of transfers and interest income, as a measure of output.
members from the household owned a business during the survey year. According to this definition, business ownership increases from 17% to 37% between 1997 and 1998 among households observed in both years.\textsuperscript{14}

The variable $O_i$ is an indicator for a household’s wage status in 1997, and is intended to measure a household’s exposure to labor market shocks through lower wages during the Asian crisis. I assign wage status to a household if it had any wage income in 1997.\textsuperscript{15} Defining $D_t$ as the dummy for year 1998, the interaction term $O_iD_t$ is equal to one if $t = 1998$ and $O_i = 1$. In the Townsend Thai data, households typically have multiple sources of income. Of the roughly 67% of households that had some wage income in 1997, about 11% also owned a business. The coefficient $\beta_1$ on this variable measures the difference-in-difference in the proportion of business owners before and after the crisis, between pre-crisis wage- and non-wage households. A positive and significant estimate of $\beta_1$ indicates that wage-earning households in 1997 entered business ownership at a faster rate than their non-wage counterparts, after accounting for common aggregate factors, household fixed effects (such as talent and preferences, and characteristics that are fixed in the short-run) and a vector of household characteristics $x_{it}$. The vector $x_{it}$ consists of the one-year lag of household wealth\textsuperscript{16}, household size\textsuperscript{17}, number of male household members, age of the head of the household, years of schooling of the head of the household, and household occupation indicators.\textsuperscript{18} Finally, the regressions include the year dummy ($D_t$) and household fixed effects ($c_i$).

I report the regression results separately by region in Table 1.2. The central region refers to the provinces of Chachoengsao and Lopburi, which are closer to the capital city of Bangkok than Buriram and Sisaket in the Northeast. The estimate of $\beta_1$ in for the central region indicates that wage-earners from 1997 entered business ownership by an additional 15 percentage points compared to households without any wage-income in 1997. Relative to the pre-crisis rate of business ownership in the region, this effect represent about 70%.

\textsuperscript{14}An alternative definition for business ownership could be based on source of income. The proportion of business households, with the majority of their income from business sources, increased from 7.6% to 12.1% between 1997 and 1998.

\textsuperscript{15}Analogous to the majority income definition for business ownership, I can assign wage status to households that derive the majority of their annual gross income from wages in 1997.

\textsuperscript{16}The wealth measure is as defined previously in Section 1.3. Including current wealth instead of lagged wealth does not changes the results qualitatively, and using the current value of land instead also does not change the estimates.

\textsuperscript{17}A household consists of all the people who lived and ate in the house for at least six months out of the last twelve months, and children who are studying away from home who are supported by members of the household.

\textsuperscript{18}Occupation variables include the interaction terms of being a rice-farmer in 1997 and $D_t$, and of being involved in any other kind of farming in 1997 and $D_t$.

\textsuperscript{19}In supplementary regressions, I find that households with wage-income as the majority source of income in 1997 were 19.1 percentage points more likely to start a business in 1998, relative to households where wage-income was not the source of the majority of income in 1998. I also find that both definitions of wage status in 1997 are statistically significant determinants of business ownership even when it is defined based on whether business is the majority source of income for the household.
Table 1.2: Fixed effects regressions of business ownership

<table>
<thead>
<tr>
<th>Dependent var: business ownership</th>
<th>Central coefficient (s.e)</th>
<th>Northeast coefficient (s.e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>year=1998</td>
<td>0.148** (0.061)</td>
<td>0.046 (0.065)</td>
</tr>
<tr>
<td>lag of wealth (millions)</td>
<td>0.176 (0.148)</td>
<td>0.175 (0.146)</td>
</tr>
<tr>
<td>wage in 1997 × year=1998</td>
<td>0.147*** (0.054)</td>
<td>0.015 (0.044)</td>
</tr>
<tr>
<td>rice in 1997 × year =1998</td>
<td>-0.006 (0.057)</td>
<td>0.006 (0.067)</td>
</tr>
<tr>
<td>farm in 1997 × year=1998</td>
<td>-0.019 (0.061)</td>
<td>0.099 (0.084)</td>
</tr>
<tr>
<td>household size</td>
<td>0.005 (0.034)</td>
<td>-0.048* (0.029)</td>
</tr>
<tr>
<td>no. of male in household</td>
<td>0.041 (0.051)</td>
<td>0.112** (0.045)</td>
</tr>
<tr>
<td>age of head</td>
<td>-0.006 (0.006)</td>
<td>0.009 (0.007)</td>
</tr>
<tr>
<td>schooling of head</td>
<td>0.006 (0.078)</td>
<td>-0.008 (0.016)</td>
</tr>
<tr>
<td>intercept</td>
<td>0.078 (0.425)</td>
<td>-0.408 (0.367)</td>
</tr>
<tr>
<td>household fixed effects</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>no. of observations</td>
<td>864</td>
<td>889</td>
</tr>
<tr>
<td>no. of households</td>
<td>433</td>
<td>446</td>
</tr>
</tbody>
</table>

Notes: * p<0.10, ** p< 0.05, ***p<0.001. Regressions exclude top and bottom 1% of changes in household size and number of men between the two years, the current wealth. “wage in 1997” is an indicator for having wage income pre-crisis in 1997; “rice in 1997” is an indicator for being involved in rice-farming in 1997, “farm in 1997” is an indicator for being involved in non-rice farm activities in 1997.

In the central region, I find that households which received an additional male member are 11 percentage points more likely to start a business in 1998. Relative to the pre-crisis rate of business ownership in the region, this effect represents an increase of about 91% in the rate of business ownership. In the Northeast, although pre-crisis wage-status does not have a similar effect on business ownership, I find that households which received an additional male member are 11 percentage points more likely to start a business in 1998. Relative to the pre-crisis rate of business ownership in the region, this effect represents an increase of about 91% in the rate of business ownership.

This result lines up with prior research that singles out the Northeast for receiving a significant amount of reverse migration during the crisis as a result of unemployment. Subhadhira et al. (2002) state that almost two million migrant workers may have returned to the region, and that it had the highest dry-season unemployment rate in the country of 8.6% in February 1998.

The central provinces are closer to Bangkok, and more reliant on wage-income compared to the Northeast. Accordingly, I find that wage-earners in the pre-crisis period were disproportionately more likely to become a business owner during the crisis. In the Northeast, wage-labor is less significant, but the crisis still affected their outside option in the form of remittances.

---

20 In supplementary regressions, I find that the level of remittance income in 1997 is a significant predictor of entry into business ownership during the crisis in the Northeast.

21 For example, Table A.2 in Appendix A.3 shows that 74.4% and 60.4% of households earned wage income respectively in the central region and the Northeast in 1997.
unemployment among migrant workers that originated from the villages there. Although
the most relevant proxy for the crisis varies by region, I argue that pre-crisis wage-status
and an increase in the number of men represent the same underlying shock for the respective
regions - a decrease in the value of the outside option to entrepreneurship.\textsuperscript{22} The argument
that the key change in these villages during the crisis was a decrease in the outside option
of entrepreneurship is further supported by the fact that there is no evidence that the in-
crease in business ownership is explained by changes in household wealth right before or
during the crisis. I interpret these findings (a positive estimate of $\beta_1$ for the central region
and an association between increase in number of men and household business ownership
for the Northeast) as supportive evidence for my main hypothesis that a negative shock to
the (wage) labor market pushed households in business ownership, and use it to motivate
the key assumption that a change in the outside-option parameter $\mu$ triggers the business
start-ups during the crisis.

1.3.3 Household borrowing and loan characteristics during the crisis

A natural competing theory for the observed increase in business ownership is that house-
holds were able to borrow more freely in 1998 relative to 1997. On the other hand, a concern
could be that borrowing constraints would actually worsen during a macroeconomic finan-
cial crisis. The credit constraint parameter ($\lambda$) in the model is not observed directly in the
data. However, I use the fact that the Townsend data has a detailed section on household
borrowing and loan characteristics.

Table 1.3 reports summary statistics on households borrowing and loan characteristics in
1997 and 1998. Between the two years, median annual borrowing increases from 16.2 to 20
thousand baht, with a median borrowing-to-wealth ratio of 0.05 and 0.07 respectively in 1997
and 1998. The proportion of households with a loan increases from 64\% to 70\%, however
the proportion of household with loans from the Bank for Agriculture and Agricultural
Cooperatives (BAAC), commercial banks and village funds do not change significantly. Data
on loan-use indicate that while more households borrowed for the purpose of consumption
during the crisis, the proportion of households borrowing for business and farm-related
investment did not increase significantly between the two years. Table 1.3 also shows that
loan characteristics (as captured in the annual interest rate, loan duration and loan size) did
not change significantly between the two years. The decrease in median loan-size goes away
completely when I exclude loans from friends and neighbors. The most common increase
in borrowing came from smaller informal loans, extended by friends and neighbors at much
lower interest rates.\textsuperscript{23}

\textsuperscript{22}In Section 1.4, I assume that both interpretations of earning lower wages or losing a job are consistent
with an aggregate shock to the alternative income.

\textsuperscript{23}Table A.3 in Appendix A.3 summarizes the data on loan characteristics by loan source for 1997 and
1998.
Table 1.3: Borrowing and loan characteristics

<table>
<thead>
<tr>
<th></th>
<th>1997</th>
<th>1998</th>
</tr>
</thead>
<tbody>
<tr>
<td>% with loans</td>
<td>64.1</td>
<td>70.4</td>
</tr>
<tr>
<td>annual borrowing (000s)</td>
<td>46.2</td>
<td>54.7</td>
</tr>
<tr>
<td></td>
<td>(100.1)</td>
<td>(104.1)</td>
</tr>
<tr>
<td></td>
<td>[16.2]</td>
<td>[20]</td>
</tr>
<tr>
<td>borrowing to wealth ratio</td>
<td>0.53</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>(4.78)</td>
<td>(2.07)</td>
</tr>
<tr>
<td></td>
<td>[0.05]</td>
<td>[0.07]</td>
</tr>
<tr>
<td>% households with loans from:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>relatives and neighbors*</td>
<td>19.4</td>
<td>33.3</td>
</tr>
<tr>
<td>BAAC</td>
<td>32.3</td>
<td>32.5</td>
</tr>
<tr>
<td>village funds</td>
<td>12.8</td>
<td>15.1</td>
</tr>
<tr>
<td>moneylender*</td>
<td>17.9</td>
<td>26.5</td>
</tr>
<tr>
<td>bank</td>
<td>3.5</td>
<td>3.1</td>
</tr>
<tr>
<td>other*</td>
<td>8.2</td>
<td>10.9</td>
</tr>
<tr>
<td>% households with loans for:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>farm-equipment and inputs</td>
<td>55.2</td>
<td>56.9</td>
</tr>
<tr>
<td>business investment</td>
<td>7.6</td>
<td>9.1</td>
</tr>
<tr>
<td>consumption</td>
<td>32.6</td>
<td>57.8</td>
</tr>
<tr>
<td>loan characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>size (000s)</td>
<td>37.0</td>
<td>32.8</td>
</tr>
<tr>
<td></td>
<td>(76.3)</td>
<td>(66.6)</td>
</tr>
<tr>
<td></td>
<td>[20]</td>
<td>[15]</td>
</tr>
<tr>
<td>duration (months)</td>
<td>21.9</td>
<td>28.8</td>
</tr>
<tr>
<td></td>
<td>(27.8)</td>
<td>(28.8)</td>
</tr>
<tr>
<td></td>
<td>[12]</td>
<td>[12]</td>
</tr>
<tr>
<td>annual interest rate</td>
<td>0.30</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>(1.84)</td>
<td>(1.62)</td>
</tr>
<tr>
<td></td>
<td>0.10</td>
<td>0.10</td>
</tr>
</tbody>
</table>

Mean, standard deviation (parentheses) and median (brackets) are reported respectively.
* indicates that difference-in-means test between years is significant at 5%. Loan characteristics are calculated at the loan level.
Overall, the evidence from household borrowing and loan characteristics do not point to any noticeable easing or a tightening in the rural credit market. In particular, the loan terms from the formal financial institutions do not indicate a structural change in the credit market that made it easier for business owners to borrow for business investment. While estimating the structural parameters of the model in the following section, I use these findings to motivate the assumption that the credit constraint parameter $\lambda$ in the model remains constant during the crisis.

1.4 Structural estimation

1.4.1 GMM - matched moments and computation

In this section, I structurally estimate the parameters of the model in Section 1.2 using the generalized method of moments. The model can be used to compute the joint probability of occupational choices in 1997 and 1998 conditional on initial wealth in each period $z$, labor market characteristics $x$ and conditional on the assumption that except for the outside-option parameter $\mu$, model parameters stay fixed between the two periods. Additionally, the model predicts the expected gross business income for year $t$ conditional on occupation in year $t$. I match these model predicted moments to their sample analogs, described in detail later in this section.

As specified in Section 1.3, the sample analogs of the occupational proportions are measured using self-reported business ownership at the household level, before the crisis (1997) and after the crisis (1998). I use the one-year lag of household assets as the measure of initial wealth $z_t$ at time $t$, and I use years of schooling of the head of the household as a measure of labor market attribute $x$. I measure gross business income in a given year, $q^E_t$, as the annual gross business income.

I make the following two key assumptions in the estimation of the model outlined previously in Section 1.2.

(i) The outside-option parameter $\mu$, which represents the minimum alternative income a household can earn, changes between 1997 and 1998, and consequently $\mu_{97} \neq \mu_{98}$.

(ii) The credit constraint parameter $\lambda$ does not change between 1997 and 1998.

Maximum likelihood estimation is an alternative method of estimation. In this case, the likelihood of joint probabilities of occupational choice in 1997 and 1998, and business incomes in 1997 and 1998 are analytically complicated. The optimization routine for MLE based only on occupational choice does not converge, with the likelihood function flat over a range of parameters. I found GMM as a suitable compromise; it is analytically and computationally easier and allows me to use data on both occupational choice and income.

Business incomes are only observed for business households. Therefore, I will match expected business income, conditional on choosing entrepreneurship, as described in detail in Table 1.4.
the crisis and business ownership during the crisis in Section 1.3.2. The second assumption is consistent with evidence from the literature that states that the rural credit markets were not affected by the financial crisis, and with Section 1.3 where I show that summary statistics on household borrowing and loan characteristics do not indicate any improvement or deterioration in the rural credit market. In the short-run, the remaining structural parameters of the model can reasonably be expected to remain constant. The elasticity of entrepreneurial return with respect to capital ($\alpha$) and the elasticity of the alternative income with respect to schooling ($\gamma$) can be thought of as slow-moving technology parameters. Finally, the implicit assumption throughout the model is that of the fixed nature of entrepreneurial talent $\theta$ between the two periods, which includes the parameters of the talent specification ($\delta_0, \delta_1, \delta_2$ and $\sigma$) being fixed as well between the two periods. Denote by $\phi$ the set of ten parameters in the model:

$$\phi = \{ \alpha, \lambda, \gamma, \mu_{t=97}, \mu_{t=98}, \delta_0, \delta_1, \delta_2, \sigma^2, r \}$$

I fix $r$ at the median annual gross interest rate on loans in the data, equal to 1.1 for both years.\(^{26}\) I estimate the remaining nine parameters of the model by matching ten moments reported in Table 1.4. Let the $j$th model-predicted moment be $h_j(z, x, \phi)$ for $j = 1, \ldots, 10$, the sample analog of the moment be $h^d_j$ for $j = 1, \ldots, 10$ and the percentage deviation of the predicted moment from its sample analog be

$$q_j(z, x, \phi) \equiv \frac{h_j(z, x, \phi) - h^d_j}{h^d_j}, \quad j = 1, \ldots, 10.$$

The first three moments in Table 1.4 are joint probabilities of occupational choice in 1997 and 1998 - the proportion of being a non-business owner in 1997 and a business owner in 1998 ($E_{01} = 1$), of being a business owner in both years ($E_{11} = 1$), and of being a non-business owner in both years ($E_{00} = 1$). Only about 3% of households in the sample transition out of business ownership between the two years. These households do not experience a fall in wealth, and the model cannot predict their exit in the case where the outside option of entrepreneurship falls as well. Since it constitutes a small number of households, I do not match this probability. The overall proportions, listed as items 1 to 3 in the table, only match the probabilities of occupational choice averaged over all values of initial wealth ($z$) and years of schooling ($x$). The next three moments relate to the probability of each category ($E_{01}, E_{11}, E_{00}$) for schooling less than the median, as listed in items 4 to 6 in the table.\(^{27}\)

\(^{26}\)Interest rates were computed using loan term and required repayment in the data for loans taken by households in the sample.

\(^{27}\)Note that matching the occupational probabilities for schooling higher than the median is redundant, as it is covered by matching the overall proportions and the proportions below median schooling. While I do not use the moments related to the occupational choices conditional on initial wealth to estimate the parameters, I will use them to look at the ability of the model to match untargeted moments.
of these income moments. The GMM estimates are produced by minimizing the following
I therefore expect the identification of the parameters to be improved with the inclusion
of business output each year, and on the selection into entrepreneurship implied by them.
ing choices at the extensive margin, the estimated parameters are also based on the level
these moments along with the occupational proportions ensures that in addition to match-
andan business ownership
additional on business ownership in each year, and expected gross business income conditional
99 bootstrap samples.
entrepreneurial income with respect to capital, \( \alpha \), is estimated to be 0.44, implying that a

<table>
<thead>
<tr>
<th>Description</th>
<th>( h_j(z, x, \phi) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. proportion of switchers</td>
<td>( \sum_{i=1}^{N} P(E_{01,i} = 1</td>
</tr>
<tr>
<td>2. proportion of business in both years</td>
<td>( \sum_{i=1}^{N} P(E_{11,i} = 1</td>
</tr>
<tr>
<td>3. proportion of non-business in both years</td>
<td>( \sum_{i=1}^{N} P(E_{00,i} = 1</td>
</tr>
<tr>
<td>4. proportion of ( E_{01} ) if ( x \leq x_m )</td>
<td>( \sum_{i=1}^{N} 1_{x_i &lt; x_m} P(E_{01,i} = 1</td>
</tr>
<tr>
<td>5. proportion of ( E_{11} ) if ( x \leq x_m )</td>
<td>( \sum_{i=1}^{N} 1_{x_i &lt; x_m} P(E_{11,i} = 1</td>
</tr>
<tr>
<td>6. proportion of ( E_{00} ) if ( x \leq x_m )</td>
<td>( \sum_{i=1}^{N} 1_{x_i &lt; x_m} P(E_{00,i} = 1</td>
</tr>
<tr>
<td>7. average gross business income in 1997</td>
<td>( \frac{1}{N} \sum_{i=1}^{N} E(q_{97}^E</td>
</tr>
<tr>
<td>8. average gross business income in 1998</td>
<td>( \frac{1}{N} \sum_{i=1}^{N} E(q_{98}^E</td>
</tr>
<tr>
<td>9. average ( q_{97}^E ) if ( x \leq x_m )</td>
<td>( \sum_{i=1}^{N} 1_{x_i &lt; x_m} E(q_{97}^E</td>
</tr>
<tr>
<td>10. average ( q_{98}^E ) if ( x \leq x_m )</td>
<td>( \sum_{i=1}^{N} 1_{x_i &lt; x_m} E(q_{98}^E</td>
</tr>
</tbody>
</table>

See Appendix A.2 for detailed derivations of \( h_j(z, x, \phi) \)’s.

\( x_m \) indicates the median value of \( x \).

The last four moments listed are, respectively, the expected gross business income conditional on business ownership in each year, and expected gross business income conditional on business ownership and on schooling being less than the median, in each year. Matching these moments along with the occupational proportions ensures that in addition to matching choices at the extensive margin, the estimated parameters are also based on the level of business output each year, and on the selection into entrepreneurship implied by them. I therefore expect the identification of the parameters to be improved with the inclusion of these income moments. The GMM estimates are produced by minimizing the following criterion function:

\[
Q_N(\phi) = \left[ \frac{1}{N} \sum_{i=1}^{N} q_i(z_i, x_i, \phi) \right] \left[ \frac{1}{N} \sum_{i=1}^{N} q_i(z_i, x_i, \phi) \right]^{-1} \]

where \( q_i(z_i, x_i, \phi) \) is a 10 \( \times \) 1 vector of percentage deviations for household \( i \). Since all the moments are converted into percentage deviations, I use the identity matrix as the weighting matrix.

### 1.4.2 Estimates and model fit

GMM estimates of the model parameters are reported in Table 1.5.\(^{28}\) The elasticity of entrepreneurial income with respect to capital, \( \alpha \), is estimated to be 0.44, implying that a

\(^{28}\) use the Matlab minimizer particleswarm, and I compute standard errors by estimating the model on 99 bootstrap samples.
10% increase in capital leads to a 4% increase in entrepreneurial income for unconstrained entrepreneurs.

The credit constraint parameter $\lambda$ is estimated to be 0.69, implying that households can invest up to 69% of their measured wealth as capital. On average, households in the sample have 74% of their wealth in 1997 in the form of land, followed by household durables (20%) and agricultural assets (6%). Therefore, it is plausible that households cannot liquidate all of their wealth. In the data, average business assets are equal to about 80 thousand baht in 1997 and 47.5 thousand baht in 1998, representing approximately 10% and 7% of average household wealth each year, and well below the estimated investment limit on average.

About 18% of entrepreneurs in 1997 and 13.4% of entrepreneurs in 1998 are predicted to be credit-constrained at the estimated credit constraint.

The constant term in the log talent equation, $\delta_0$, is estimated to be 2.1, with a conditional variance $\sigma$ of 0.46, an elasticity with respect to initial wealth in 1997 of 0.03, and an elasticity with respect to schooling of 0.24. Since schooling is positively correlated with entrepreneurial talent, higher education is associated with both higher entrepreneurial income and the alternative non-business income.\footnote{The estimates of $\delta_1$ and $\delta_2$ are not statistically significant at the 10% level. From conducting several robustness checks (reported in detail in Section 1.4.4), I find that there is indeed a lot of variability in the estimates of $\delta_1$ and $\delta_2$ across different subsamples. Setting $\delta_1$ and $\delta_2$ equal to zero produces much greater model errors, and therefore, I continue to use the estimated values in the counterfactual analyses.}

The remainder of the parameters describe the outside option of business ownership. In 1997, I estimate that households expect to earn a minimum alternative income ($\mu$) of 57 thousand Thai baht. To explain the increase in business ownership during the crisis, $\mu$ is estimated to fall by 47% to 30 thousand Thai baht. The estimate of $\gamma$ implies that a 10% increase in schooling will increase the alternative income by about 2%.

Before discussing the implications of the structural estimates, I evaluate how well the model does in matching the ten targeted moments. At the estimated parameters, Table 1.5 shows that the model predicts occupational choice proportions that are closely comparable to the observed counterparts. The six occupational proportions are predicted within 4% of the observed data.\footnote{An extension of the model with an exogenous probability of exit will further reduce these gaps, as the model’s inability to predict exit from business ownership during the crisis ($E_{10}$) contributes to the gaps.} The four income moments are even more closely matched by the model. The average business income of business households, overall and conditional on below median schooling, are matched within 2% of the observed averages.

Still, one can reject the overidentifying restrictions in the model at conventional confidence intervals, as the J-statistic is equal to 19.01. This is not surprising given the highly stylized model. The magnitude of the J-statistic is impacted most heavily by the fifth moment (proportion of households that are business owners in both 1997 and 1998 among those whose schooling is below or equal to the median), followed by the sixth moment (proportion
### Table 1.5: Structural estimates and model fit

<table>
<thead>
<tr>
<th>parameters</th>
<th>estimate</th>
<th>(s.e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>return to capital in business income</td>
<td>α</td>
<td>0.4376*** (0.0285)</td>
</tr>
<tr>
<td>credit constraint parameter</td>
<td>λ</td>
<td>0.6896*** (0.2270)</td>
</tr>
<tr>
<td>talent - standard deviation</td>
<td>σ</td>
<td>0.4651** (0.1930)</td>
</tr>
<tr>
<td>talent - constant</td>
<td>δ₀</td>
<td>2.0676*** (0.3231)</td>
</tr>
<tr>
<td>talent - elasticity w.r.t initial wealth</td>
<td>δ₁</td>
<td>0.0341 (0.0560)</td>
</tr>
<tr>
<td>talent - elasticity w.r.t schooling</td>
<td>δ₂</td>
<td>0.2399 (0.1551)</td>
</tr>
<tr>
<td>non-business income parameter – 1997</td>
<td>µ₉₇</td>
<td>57.11*** (5.9366)</td>
</tr>
<tr>
<td>non-business income parameter – 1998</td>
<td>µ₉₈</td>
<td>30.07*** (5.3842)</td>
</tr>
<tr>
<td>return to schooling in non-business income</td>
<td>γ</td>
<td>0.1683** (0.0810)</td>
</tr>
</tbody>
</table>

Bootstrap standard errors are reported. (* p<0.10, ** p< 0.05, ***p<0.001).

### Model fit for targeted moments

<table>
<thead>
<tr>
<th>moment</th>
<th>predicted</th>
<th>observed</th>
<th>% deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>proportion of switchers (E₀₁)</td>
<td>0.2201</td>
<td>0.2241</td>
<td>2.43%</td>
</tr>
<tr>
<td>proportion of business in both years (E₁₁)</td>
<td>0.1219</td>
<td>0.1224</td>
<td>-0.43%</td>
</tr>
<tr>
<td>proportion of non-business in both years (E₀₀)</td>
<td>0.6580</td>
<td>0.6411</td>
<td>2.64%</td>
</tr>
<tr>
<td>proportion of E₀₁ if x ≤ median(x)</td>
<td>0.2135</td>
<td>0.2144</td>
<td>-0.40%</td>
</tr>
<tr>
<td>proportion of E₁₁ if x ≤ median(x)</td>
<td>0.1117</td>
<td>0.1100</td>
<td>1.52%</td>
</tr>
<tr>
<td>proportion of E₀₀ if x ≤ median(x)</td>
<td>0.6748</td>
<td>0.6516</td>
<td>3.55%</td>
</tr>
<tr>
<td>expected gross business income in 1997 (q₉₇)</td>
<td>212.19</td>
<td>212.30</td>
<td>-0.05%</td>
</tr>
<tr>
<td>expected average gross business income in 1998 (q₉₈)</td>
<td>135.58</td>
<td>133.66</td>
<td>1.43%</td>
</tr>
<tr>
<td>expected q₉₇ if x ≤ median(x)</td>
<td>201.34</td>
<td>201.28</td>
<td>0.03%</td>
</tr>
<tr>
<td>expected q₉₈ if x ≤ median(x)</td>
<td>128.57</td>
<td>130.59</td>
<td>-1.55%</td>
</tr>
<tr>
<td>sum of squared % deviations (criterion function)</td>
<td></td>
<td></td>
<td>0.0033</td>
</tr>
<tr>
<td>J-statistic</td>
<td></td>
<td></td>
<td>19.01</td>
</tr>
</tbody>
</table>
of households that are not in business in both 1997 and 1998 among those whose schooling is below or equal to the median).\textsuperscript{31}

Nevertheless, Figure 1.3 shows that the GMM estimates are also able to replicate the relationship between wealth in 1997 and the occupational proportions relatively well, although I do not use this information to estimate the parameters. Local polynomial fit of the observed and predicted proportions of entrepreneurship in 1997 and 1998 are plotted in the first two graphs. The model predictions lie within the 95% confidence interval of the observed proportions. Consequently, the observed and predicted relationship in the remaining two graphs that combine occupational proportions over two years are also closely matched.

1.4.3 Analysis

\textbf{Predicted fall in income due to the crisis and income-mitigation through occupational choice}

In this section, I analyze the implications of the estimated model, particularly in quantifying how effective starting a business is in mitigating the income shock during the crisis. To do so, I use the model to calculate the following statistics for income in two counterfactual scenarios:\textsuperscript{32}

(i) Average (and median) income of the sample in 1998, assuming that the outside option parameter $\mu$ had remained at the pre-crisis level. That is, I predict income in 1998 by setting $\mu = \mu_{97}$, instead of using the estimated value of $\mu_{98}$, and use it as the counterfactual income in the absence of the crisis.\textsuperscript{33}

(ii) Average (and median) income of the sample in 1998, where households are assigned their pre-crisis occupation, resulting in some households being unable to mitigate the decrease in their income by starting a business during the crisis. This counterfactual is referred to as \textit{crisis+fixed occupation}.

The predicted fall in income during the crisis is the difference between a household’s without-crisis income and the predicted income for 1998 using the baseline estimates. The top panel of Table 1.6 shows that for the whole sample, average income is estimated to fall by 17%. The fall in average income could have been higher at 19% had some of the households not switched into business ownership following the fall in their outside options.\textsuperscript{34}

\textsuperscript{31}The J-statistic is equal to the sample size times the sum of squared deviations evaluated at the GMM estimates obtained in two stages. Evaluated at these estimates, the fifth and sixth moments have percentage deviations of 5.5% and 2.8% respectively. The remaining moments have deviations of less than 2.5%.

\textsuperscript{32}I define income as a household’s end-of-period wealth and initial wealth in each period. I calculate predicted income for each household in the sample by integrating over 100 draws of talent $\theta$ from the estimated distribution.

\textsuperscript{33}Alternatively, I can use average or median income in 1997. Even if $\mu$ had stayed constant, the model predicts a 0.12 percentage points increase in entrepreneurship in 1998 due to changes in initial wealth, and a 2% (3%) increase in average (median) income. Using income in 1997 as the “no-crisis” counterfactual would not account for these changes.
Figure 1.3: Occupational proportions as function of wealth

Proportion of business owners, 1997

Proportion of business owners, 1998

Proportion with $E_{01} = 1$

Proportion with $E_{00} = 1$

Proportion with $E_{11} = 1$
Table 1.6: Predicted change in income during the crisis

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>median</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All households</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>without crisis</td>
<td>158.4</td>
<td>108.1</td>
</tr>
<tr>
<td>crisis (baseline)</td>
<td>-17.1%</td>
<td>-24.8%</td>
</tr>
<tr>
<td>crisis + fixed occupation</td>
<td>-19.1%</td>
<td>-27.9%</td>
</tr>
<tr>
<td>income mitigation from starting a business</td>
<td>10.5%</td>
<td>11.1%</td>
</tr>
<tr>
<td><strong>Conditional on switching into business ownership</strong> ($E_{01} = 1$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>without crisis</td>
<td>155.4</td>
<td>103.4</td>
</tr>
<tr>
<td>crisis (baseline)</td>
<td>-13.4%</td>
<td>-19.3%</td>
</tr>
<tr>
<td>crisis + fixed occupation</td>
<td>-22.4%</td>
<td>-33.3%</td>
</tr>
<tr>
<td>income mitigation from starting a business</td>
<td>40.2%</td>
<td>42.0%</td>
</tr>
<tr>
<td><strong>Conditional on remaining in non-business</strong> ($E_{00} = 1$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>without crisis</td>
<td>152.5</td>
<td>102.1</td>
</tr>
<tr>
<td>crisis (baseline)</td>
<td>-22.7%</td>
<td>-34.1%</td>
</tr>
<tr>
<td>crisis + fixed occupation</td>
<td>-22.7%</td>
<td>-34.1%</td>
</tr>
<tr>
<td>income mitigation from starting a business</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>

Therefore, the model predicts that about 11% of the fall in average income is mitigated due to households starting a business as a response to the labor market shock. Conditional on starting a business ($E_{01} = 1$), income mitigation due to the endogenous occupational switches is estimated to be much higher at about 40%.\(^{34}\)

Figure 1.4 plots the predicted percentage loss in income during the crisis for households against initial wealth and years of schooling. First, it can be seen that the predicted loss in income due to the crisis is higher the poorer the household. To begin with, wealthier households are more likely to already be in business in 1997 because they are able to invest the desired level of capital, and are therefore less likely to be affected by a decrease in the outside option. Second, among the non-business owners from before the crisis, wealthier households are more likely to be able to transition into business ownership during the crisis, again through the channel of the credit-constraint. Finally, among those who are able to switch into business ownership as a result of a fall in $\mu$, the wealthier households are more likely to be able to invest the first-best level of capital, and earn higher income from business ownership. The graph also shows predicted income loss during the crisis from the counterfactual exercise where households are assigned the same occupation from the pre-crisis period. Loss mitigation through entrepreneurship is most relevant for those in

\(^{34}\)This is an obvious artifact of the model assumptions. Note that the statistics in the first panel (all households) is the weighted sum of the corresponding numbers in the second ($E_{01} = 1$) and the third panel ($E_{00} = 1$), where the weights are the respective predicted proportions of each category. Also note that for households predicted to be entrepreneurs in both years ($E_{11} = 1$), the model predicts zero loss in income due to the crisis.
the middle of the wealth distribution. All else equal, the wealthiest households are more likely to already be business owners (making the crisis irrelevant in terms of income loss), while the poorest households are less likely to be able to transition into entrepreneurship to smooth their incomes. Percentage loss in income during the crisis is also higher the lower the household head’s schooling. This reflects the fact that households with higher years of schooling are more likely to be entrepreneurs. For a given level of wealth, schooling affects the probability of entrepreneurship positively through its effect on $\theta$, and negatively through its effect on the alternative income. Overall, it seems that households with higher years of schooling have a comparative advantage in being entrepreneurs.

In the following subsections, I simulate the effect of relaxing the credit constraint for business investment and improving entrepreneurial skills $\theta$ - both of which make entrepreneurship feasible and more profitable, and improve the ability of households to weather shocks to the alternative income. Within these policies, the effectiveness of one over the other would depend on the structural parameters and the observed distribution of wealth and education. For example, in an economy where households are sufficiently talented, but are relatively more constrained by the borrowing limit, further increases in talent would be ineffective in increasing business start-ups and in raising incomes. I also simulate the effect of three types of direct transfers to households, and evaluate their effectiveness on the basis of gain in income for a given level of spending. An economy with a large number of potentially profitable but credit-constrained entrepreneurs will be able to generate more income per dollar through business grants, than through one-to-one income transfers.

**Relaxing the credit constraint**

Conditional on starting a business, the structural estimates suggest that entrepreneurship helped reduce about 40% of the loss in average income during the crisis. I evaluate the
scope for further income mitigation by relaxing the credit constraint in the presence of the crisis, particularly among those that do not start a business during the crisis in the baseline. About 8.9% of non-entrepreneurs, and 13.7% of entrepreneurs in 1998 are predicted to be credit-constrained. That is, the first-best level of capital, $k_u(\theta)$, is higher than the limit $\lambda z$. A priori, these households could benefit from policies that increase the parameter $\lambda$, allowing them to borrow a higher fraction of their wealth. Relaxing the credit-constraint could be achieved mechanically, by allowing households to borrow a higher fraction of their wealth while holding all the other parameters fixed at the GMM estimates. It could also be targeted through improvements in contract enforcement.

First, Figure 1.5 shows that increasing $\lambda$ has a less than one percentage point effect on entrepreneurship in the post-crisis economy. Even though the proportion of credit-constrained households reduces from 10.6% to 2.4% when $\lambda$ is increased by ten fold, the rate of business ownership increases by less than a percentage point, from 34.3% to 34.9%. Figure 1.6 shows that the 10th percentile of income does increase steadily as $\lambda$ increases, but there is a no discernible effect on median income.

To look more closely at the effect of the policy, Figure 1.7 plots the percentage gain in income in 1998 from increasing $\lambda$ to $2\lambda$ against initial wealth and schooling. Poorer households are more likely to be credit-constrained because the borrowing limit is a function of wealth.\(^{35}\) Initially, there is a small part where income gain from the policy is increasing in wealth, reflecting the fact that doubling $\lambda$ does not sufficiently increase entrepreneurial income for the poorest households to change their occupation or operate their businesses at the first-best level. However, for the most part, easing the borrowing limit leads to higher gains in income the poorer the household.\(^{36}\) While the model predicts that relaxing the credit constraint would help poorer households with income smoothing, it does so in a limited way. Virtually eliminating the credit constraint (in the case where $\lambda$ is increased by twenty fold) does not offset the fall in income created during the crisis, although it reduces it somewhat for poorer households.

**Increasing entrepreneurial productivity**

An increase in $\theta$ can be interpreted as the potential outcome of a business training initiative that increases entrepreneurial productivity. The results of this program do not rely on forcing every household to participate in the training program, and can be interpreted as a voluntary opt-in business training program. Since talent and capital are complementary in the entrepreneurial production function, the rate of increase in income from business training would depend on whether the credit constraint binds for business owners as $\theta$ increases.

\(^{35}\)Since wealth and entrepreneurial talent are positively correlated (through $\delta_1$), this effect is dampened a little because poorer households have lower talent on average.

\(^{36}\)Income gain is less sensitive to years of schooling. However, income gain is noticeably lower for households with highest years of schooling, possibly because these households are predicted to have the highest outside options.
Figure 1.5: Effect on entrepreneurship in 1998 - increase in $\lambda$

Figure 1.6: Percentage gain in income in 1998 - increase in $\lambda$

Figure 1.7: Percentage gain in income in 1998 as function of wealth - $2\lambda$
Figure 1.8: Effect on entrepreneurship in 1998 - increase in $\theta$

Figure 1.9: Percentage gain in income in 1998 - increase in $\theta$

Figure 1.10: Percentage gain in income in 1998 as function of wealth - 10% increase in $\theta$
When the baseline level of talent is relatively high, a further increase in productivity has a smaller effect on entrepreneurial income, households enter business ownership at a slower rate, and therefore equilibrium income rises at a slower rate as well.

Figure 1.8 shows that the rate of increase in the proportion of credit-constrained households is slower than the rate of increase in business ownership. Although it is hard to compare the scale of this policy to that of relaxing the credit constraint analyzed in the previous section, the steady increase in entrepreneurship and median income (Figure 1.9) indicate that households are more constrained by the lack of entrepreneurial talent compared to the credit-constraint. To look more closely at the distribution of income gains from this policy, Figure 1.10 plots the predicted gain in income from a 10% increase in $\theta$. Households in the middle of the wealth distribution experiences the highest percentage increase in income. Unlike those in the lower quartiles of wealth, they are likely to be able to invest the required amount of capital following the increase in productivity. Unlike those in the higher quartiles of wealth, the middle-wealth households are also more likely to undergo a switch from non-business to business ownership, and therefore experience additional gain in income. A drawback of using policies that increase entrepreneurial productivity to cope with this particular income shock is the lack of their effect on the poorest households. Both credit and talent are important constraints for such households, necessitating a combination of policies that relax their credit constraint and improve their entrepreneurial productivity.

**Transfers**

I consider three kinds of direct transfers to households as an antidote to the crisis, each of which incentivizes occupational choice in different ways - wage guarantees to non-business households, grants for business investment (where the transfer is contingent on being wholly invested in business), and unconditional income transfers. Figures 1.11 and 1.12 compare the rate of entrepreneurship and the median income of the sample under these programs for the same level of spending, measured in terms of *average transfer per household.*

Suppose households receive an unconditional transfer of size $T$. The model predicts that they will re-optimize in terms of how much to invest if they were to start a business. The opportunity cost of investing a dollar in business is $r$, and therefore unconditional transfers do not distort the incentive to become a business owner. The resulting increase in end-of-period income is at least $rT$; it is higher if the policy induces a business household to expand the scale of its business, or a non-business household to switch into business ownership. Figure 1.11 shows that there is no increase in business ownership when transfers

---

37 Note that business owners do not receive any transfer under the wage guarantee program, while non-business owners do not receive any transfer under the business-grant program. As a result, for the same level of spending, business owners would receive more than the average transfer as business grants, and non-business owners would receive more than the average transfer as outside-option guarantees.

38 Households that remain non-business receives an increase of $rT$. The income of unconstrained business owners also increase by $rT$. If a business becomes unconstrained because of the program, income increases by...
Figure 1.11: Effect on entrepreneurship in 1998 - transfers

Figure 1.12: Percentage gain in median income in 1998 - transfers

Figure 1.13: Percentage gain in income in 1998 as function of wealth - transfers
are distributed unconditionally, indicating that non-business households in the baseline are relatively inefficient entrepreneurs.

With business grants of size $T$, households now have two sources of credit, each priced differently. Denote funds borrowed from the traditional source as $k_a$, priced at the market interest rate $r$. The opportunity cost of investing a dollar from $T$ in business is zero. Since the marginal product of capital is always positive, households will take $T$ dollars regardless of how large the transfer is. The household solves the following capital investment problem:

$$\max_{k_a} \theta(k_a + T)^\alpha - rk_a, \quad k_a \leq \lambda z$$

The solution for $k_a$ is as follows:

$$k_a^* = \begin{cases} \max\{k_u(\theta) - T, 0\} & \text{if } k_u(\theta) - T \leq \lambda z \\ \lambda z & \text{if } k_u(\theta) - T > \lambda z \end{cases}$$

With a business grant of size $T$, a household will generate at least $rT$ in additional income if

$$\theta(k_a^* + T)^\alpha - rk_a^* + rz > \theta k_b^ {\alpha} - rk_b^* + r(z + T) \iff \theta > \frac{r(k_a^* - k_b^*) + T}{(k_a^* + T)^\alpha - k_b^ {\alpha}}$$

where $k_a^*$ and $k_b^*$ are respectively the equilibrium levels of capital when under the business grant policy and the unconditional transfer policy of size $T$. Therefore, business grants could still be efficient for households that are sufficiently talented. In an economy with severely credit-constrained talented households, additional funds in the form of business grants could generate much higher incomes than in-kind transfers (such as food stamps). However, since the opportunity cost of investing a dollar in business is zero, households for whom the return to investment is less than $r$ might take up the grant, leading to a scenario where a dollar in transfer creates than less $r$ dollars in income.

Figure 1.12 shows that for the same level of spending, the increase in median income is the highest under unconditional transfers.\textsuperscript{39} The impact of business grants on median income is equally high initially, and is then subject to diminishing returns. This reflects the fact that households who choose business ownership even at smaller amounts of business grants are the most suited to be entrepreneurs and are therefore able to generate more additional income.\textsuperscript{40}

---

\textsuperscript{39} This is true when I plot the same graph for the 10 percentile of income, and for mean income.

\textsuperscript{40} For the same average transfer per household, business owners under the business grant program would receive a bigger transfer compared to that under the unconditional transfer.
With wage-guarantees, households will compare their total alternative income (including the handout) to their income from business ownership. Households with low prospects as entrepreneurs (those with low talent and/or low wealth) are therefore more likely to take the option. The increase in income for households that would have chosen the alternative occupation even in the absence of the program is mechanically equal to the size of the handout. For households that would have chosen business ownership in the absence of the program, the increase in their income is equal to the size of the handout minus the income differential between the two occupations. By discouraging entrepreneurship among some households, an outside-option guarantee reduces additional income generation. However, the fact that wage-guarantees increase median income by more than business grants (in Figure 1.12) reiterates the point about how unsuitable entrepreneurship is for most of the sample.

Figure 1.13 also illustrates how each of the transfers affects household income. It plots percentage gain in income from the baseline in 1998 when approximately 6% of median wealth is spent in transfers per household. It is not surprising that unconditional transfers of the same amount to every household produces the biggest percentage gain for the poorest households, because they are also predicted have lower baseline incomes. Following similar logic, wage-guarantees and business grants also lead to the biggest percentage gain in income the poorer the household, because the transfers (that are equal for all households given that they opt in) make up a bigger percentage of their baseline incomes. More importantly, for a given percentile of wealth, the predicted gain in income is the highest under unconditional transfers, followed by wage guarantees, and then by business grants. Spending on business grants is therefore a particularly wasteful way to mitigate the income shock.

1.4.4 Robustness checks

For the sake of parsimony, I do not account for several plausible heterogeneities in the model, particularly in how the crisis affects different households. In order to account for some of the heterogeneity, I estimate the model on different subsamples. The second and third columns of structural estimates in Table 1.7 are from estimating the model separately for the central region and the Northeast. Compared to the Northeast, the outside option of business is estimated to be larger in the central region (in both years), which is consistent with the generally higher income levels in the region. Nonetheless, the fall in the outside option during the crisis is estimated to be about 50% in both regions.

The next two columns report estimates for households below and above the median wealth respectively. The fall in the outside option of business ownership with below median wealth is estimated to be about 45%, while it is estimated to be about 60% for households with above median wealth.\textsuperscript{41} I look at different subsamples based on households’

\textsuperscript{41}As 61% of heads of household have four years of schooling, I do not conduct a similar robustness check for below- and above-median schooling.
### Table 1.7: Robustness checks

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Base</th>
<th>CE</th>
<th>NE</th>
<th>zL</th>
<th>zH</th>
<th>Wage</th>
<th>Agr</th>
<th>Inc</th>
</tr>
</thead>
<tbody>
<tr>
<td>return to capital, $q^E$</td>
<td>$\alpha$</td>
<td>0.44</td>
<td>0.43</td>
<td>0.25</td>
<td>0.44</td>
<td>0.40</td>
<td>0.45</td>
<td>0.43</td>
</tr>
<tr>
<td>credit constraint</td>
<td>$\lambda$</td>
<td>0.69</td>
<td>0.70</td>
<td>0.90</td>
<td>0.38</td>
<td>0.89</td>
<td>0.28</td>
<td>0.86</td>
</tr>
<tr>
<td>talent - stdev</td>
<td>$\sigma$</td>
<td>0.46</td>
<td>0.50</td>
<td>0.57</td>
<td>0.51</td>
<td>0.75</td>
<td>0.56</td>
<td>0.36</td>
</tr>
<tr>
<td>talent - constant</td>
<td>$\delta_0$</td>
<td>2.07</td>
<td>2.36</td>
<td>2.52</td>
<td>2.88</td>
<td>1.86</td>
<td>2.86</td>
<td>2.16</td>
</tr>
<tr>
<td>talent - initial wealth</td>
<td>$\delta_1$</td>
<td>0.03</td>
<td>0.05</td>
<td>-0.17</td>
<td>-0.25</td>
<td>0.03</td>
<td>-0.28</td>
<td>0.04</td>
</tr>
<tr>
<td>talent - schooling</td>
<td>$\delta_2$</td>
<td>0.24</td>
<td>0.13</td>
<td>0.72</td>
<td>0.40</td>
<td>0.24</td>
<td>0.62</td>
<td>0.22</td>
</tr>
<tr>
<td>non-business –1997</td>
<td>$\mu_{97}$</td>
<td>57.1</td>
<td>71.1</td>
<td>42.1</td>
<td>55.8</td>
<td>57.8</td>
<td>42.9</td>
<td>59.0</td>
</tr>
<tr>
<td>non-business –1998</td>
<td>$\mu_{98}$</td>
<td>30.1</td>
<td>35.4</td>
<td>21.9</td>
<td>30.6</td>
<td>22.5</td>
<td>16.6</td>
<td>34.2</td>
</tr>
<tr>
<td>return to edu. in $q^A$</td>
<td>$\gamma$</td>
<td>0.17</td>
<td>0.20</td>
<td>0.15</td>
<td>0.01</td>
<td>0.10</td>
<td>0.30</td>
<td>0.15</td>
</tr>
<tr>
<td>sample size</td>
<td></td>
<td>833</td>
<td>388</td>
<td>455</td>
<td>417</td>
<td>416</td>
<td>555</td>
<td>664</td>
</tr>
<tr>
<td>criterion func. value</td>
<td></td>
<td>0.003</td>
<td>0.003</td>
<td>0.010</td>
<td>0.008</td>
<td>0.007</td>
<td>0.030</td>
<td>0.011</td>
</tr>
</tbody>
</table>

“CE” and “NE” refer to the central and the Northeast region respectively. $zL$ and $zH$ refer to subsamples of below- and above-median wealth. “Wage” refers to wage-earning households in 1997; “Agri.” to those that report being involved in agriculture in 1997; “Inc” refers to estimates where total income is used.

non-business occupation in the next two columns. Under “wage”, I consider wage-earning households in the pre-crisis period of 1997 (67% of the sample). The fall in the outside option among them is indeed estimated to be higher at about 61%, suggesting that crisis led to a higher income shock for these households. Under “agri”, I report estimates for households that were occupied in agriculture (farming of rice and other crops, and livestock raising) in 1997, and find that the fall in the outside option is estimated to be about 42%.

Estimates in the last column of Table 1.7 uses total gross household income in each year as a proxy for entrepreneurial income $q^E_t$. Total income and business income for business households are different because households in the Townsend data set typically have multiple sources of income. One could reinterpret the quantity $q^E_t$ as measuring household income for a business household, rather than income from business. The outside option parameter, $\mu$, is estimated to fall by 24% in this version of the model, lower than the 50% fall estimated when only business income is used to measure $q^E_t$. This might be an indication that starting a business is just one of many ways in which households smooth their incomes.

---

42It could be argued that primarily agricultural households (in 1997) were not exposed to the labor market shock, and therefore including them biases the estimates downward. About 36% of households derive the majority of their gross income from farming in 1997. However, about 55% of these households earn some wage-income. I estimated the model a subsample that excludes these households, and found that the decrease in the outside option parameter is about 43%, compared to the baseline of 47%.

43I also estimate a version of the model where I set the outside option parameter $\mu$ to be the same in both years, and instead allow an increase in the credit-constraint parameter $\lambda$ to account for the trends in occupation and business income during the crisis. The fit of the model deteriorates to a sum of squared deviations of 0.0087, compared to the baseline of 0.0033, although the number of free parameters in the baseline and in this exercise is the same. More importantly, the estimates become implausible; $\lambda$ in 1997 goes to the lower bound $(10^{-3})$, and $\lambda$ in 1998 is estimated to be 0.03 or 300 times higher. If I restrict the factor by which $\lambda$ increases between the two years to be less than 10, the fit of the model deteriorates substantially to 0.2371.
1.5 Conclusions

This paper is motivated by the fact that small business ownership occupies a large proportion of the population in developing countries, substantially higher than that observed in developed countries. I study an episode of a remarkable increase in the rate of business ownership in rural Thailand during the 1997 Asian crisis, and use it explore the idea that much of the entrepreneurship in developing countries could be due to the lack of better alternatives, as opposed to being driven by entrepreneurial opportunities and improving market conditions.

I argue that the crisis affected business ownership in Thai villages primarily through a negative shock to the outside option of rural households. To support this hypothesis, I present reduced-form evidence that in villages close to the capital city of Bangkok, wage earning households from the pre-crisis period were disproportionately more likely to start a business during the crisis. Additionally, I show that in the Northeast region of the country, an additional male member during the crisis is associated with starting a business during the crisis, consistent with accounts of massive reverse-migration from the urban centers to the Northeast. In further support of my claim that the increase in business ownership during the crisis is driven by the negative shock to paid employment, I show that it is highly unlikely that the crisis relaxed rural credit markets - the median annual interest rate, loan size and loan duration from formal lending sources remained stable during the crisis.

Motivated by these findings, I model the 1997 economic crisis as an exogenous shock to the outside option of entrepreneurship in an extended version of the occupational choice model of Evans and Jovanovic (1989), which maps a household’s initial wealth and level of schooling to its occupational choice before and after the crisis. I structurally estimate the model to quantify the fall in the outside option of entrepreneurship required to explain the observed increase in business ownership between 1997 and 1998, along with other parameters, such as the parameters of the entrepreneurial talent distribution and a parameter that signifies the credit constraint. My results indicate that the increase in business ownership during the crisis was induced by an estimated 47% fall in the outside option. I also find that conditional on (endogenously) starting a business during the crisis, households are able to self-insure against such income shocks to a significant extent by starting a business - they are able to reduce approximately 40% of the fall in average income as a result of starting a business. Business ownership therefore helps smooth income to a large extent for some households. I evaluate various policy counterfactuals and estimate that the effectiveness of easing the credit constraint for business investment is restrained by the lack of productivity among potential entrepreneurs in the post-crisis economy. For the same total expenditure in transfers, the percentage gain in income from business grants is lower than that achieved under unconditional transfers and wage guarantees. For the poorest and the least entrepreneurially-skilled households, a coordinated effort to raise both entrepreneurial
productivity and a relaxation of credit constraints is required. Otherwise, encouraging entrepreneurship as an alternative to instituting unemployment insurance is likely to leave out the poorest and those with low entrepreneurial talent.

This paper can be extended in several ways. Although income maximization is a reasonable approximation for how households make occupational choices, it is possible that there are non-monetary aspects of occupational choice that individuals and households care about. In the case where self-employment has undesirable non-pecuniary features, a model that assumes only income maximization as the criterion for occupational choice will underestimate the required fall in the outside option to explain a given increase in entrepreneurship. Second, the impact of the crisis on the outside option of business ownership is assumed to be proportionately the same for all households. This assumption can be justified by treating the labor shock (or at least part of it) as an economy-wide increase in the probability of unemployment. In addition, when I estimate the model on various subsamples, I find that the crisis affected the outside option of most households by roughly the same proportion. However, it should be noted that if the incidence of the labor market shock varied by characteristics other than region, wealth and pre-crisis sources of income (resulting in some households not having to start a business), the model wrongly attributes a fall in their income during the crisis. Finally, households might respond to an income shock in ways other than starting a business, for example by working longer hours in non-business occupations. A model that explicitly endogenizes labor supply and also allows for a utility cost of labor would be more appropriate for exploring this possibility.
Chapter 2

Involuntary Entrepreneurship -
Evidence from Urban Thai Data

This chapter is co-authored with Alexander Karaivanov.¹

2.1 Introduction

Ever since the classic writings by Smith, Knight and Schumpeter, entrepreneurship, or running one’s own business, has been viewed by most economists as an important engine of economic growth and innovation. Taxes and other government policies are often explicitly designed to help ‘small businesses’ grow and prosper. On the other hand, self-employment is particularly widespread in developing countries – for example, the World Bank Development Indicators data shows that more than 80% of total employment is in the form of self-employment in the poorest countries. How to reconcile the notion of entrepreneurship as a driver of growth and innovation with the fact that it is predominant in very poor countries, often with low or negative GDP growth?

As Banerjee and Duflo (2007) put it, “...it is important not to romanticize the idea of these penniless entrepreneurs”. They add “...Are there really a billion barefoot entrepreneurs, as the leaders of microfinance institutions and the socially minded business gurus seem to believe? Or is it just an optical illusion, stemming from a confusion about what we call an entrepreneur?” (Banerjee and Duflo, 2011).

Obviously, the way to solve the apparent contradiction about the role of entrepreneurship in the economy is to acknowledge that entrepreneurs are not all alike. Some people start own businesses purely on their own volition, sometimes quitting a wage job to do so. Others, however, become self-employed involuntarily or by necessity, as their only option to survive. Clearly the potential policy implications differ for these two categories – while some may need tax rebates, others may need social insurance or marketable skills. While this point

¹Department of Economics, Simon Fraser University, akaraiva@sfu.ca.
is easy to make, it is much harder to distinguish in the data which entrepreneurs fall in which category. Most of the existing empirical literature (reviewed in detail below) adopts a reduced form approach and uses an ad-hoc criterion to distinguish between the two categories of entrepreneurs.\(^2\) Naturally, self-identified data on involuntary entrepreneurs is rare, with one exception being the Global Entrepreneurship Monitor (GEM) survey which finds that, on average, 21% of the respondents in OECD countries and about 46% in non-OECD countries choose the second option in the question: “Are you involved in this start-up/firm to take advantage of a business opportunity or because you have no better choices for work?” (Poschke, 2012)

Along with the empirical literature, there is a large literature on occupational choice between wage work and starting a business (Banerjee and Newman, 1993; Piketty, 1997; Aghion and Bolton, 1997; Evans and Jovanovic, 1989; Lloyd-Ellis and Bernhard, 2000; Paulson et al., 2006; Karaivanov, 2012; Buera, 2009; Nguimkeu, 2014 among others). In all these papers, the key assumption is that economic agents freely choose, out of all available options, the occupation they prefer the most. Typically this is taken to mean maximizing expected income. Many of the models allow for credit or labor market imperfections which shape the optimal choice by affecting the expected payoffs of the different occupations, but all occupations are always considered and can be chosen by all individuals, for example, depending on an unobserved “entrepreneurial talent” variable. This modeling assumption is hard to reconcile with the data presented earlier, which seems to imply that some individuals would ideally choose a different occupation (e.g., wage work instead of running a business) if it were available to them.

We build and estimate, using urban data from Thailand, a structural occupational choice model that explicitly allows for the possibility that some individuals may have a restricted choice set of occupations. In particular, in our model some agents engage in self-employment due to lack of access to wage work.\(^3\) This can be motivated either by education, qualifications, and other similar barriers to finding paid work, or as an outcome of queuing for a restricted pool of wage jobs in an efficiency wage setting.

Specifically, we extend the classic occupational choice model of Evans and Jovanovic (1989). In that model, individuals who differ in their initial wealth and ‘entrepreneurial ability’ choose between starting a business and wage work. They can invest up to a fixed fraction of their initial wealth, \(\lambda z\) in the business, representing a credit market constraint. Entrepreneurship is chosen over wage-work if the net income from running a business is larger than the income from wage work. We extend this basic framework by adding a prob-

\(^2\)For example, one could compare individuals who left a paid job to start a business vs. all others (Block and Wagner, 2010) or those who run an own-account business vs. those who employ other people (de Mel et al., 2012).

\(^3\)Unemployment is ruled out as a viable choice, e.g., due to lack of social safety nets while the other typical option from the literature, subsistence agriculture, is not applicable to the urban environments we have in mind.
ability with which an agent, based on his observable characteristics, does not have access to wage work. This gives rise to involuntary entrepreneurship if, in the absence of this choice constraint, the agent would have maximized his income as a wage worker. We parameterize the occupational choice (labor market) constraint in such a way that different values of the parameter correspond to different levels of the tightness of the constraint. Upon estimation, this allows the data to reveal whether the choice constraint is negligible or significant, and therefore whether our extension to the basic income-maximization model does or does not help explain the observed occupational choices in the data. Additionally, our approach allows to quantify the number of involuntary entrepreneurs and their distribution based on observables such as initial wealth and years of schooling.

We use the Townsend Thai Project Initial Household Survey (Urban Area) of 2005 (NORC, 2008). It covers six different Thai provinces (Chachoengsao, Lopburi, Srisaket, Buriram, Phrae, and Satun), and surveys households in the municipal areas considered urban and semi-urban. The data include retrospective information on wealth and assets, income, household business, lending and borrowing, and individual level demographic and occupation variables. In our sample 66.1% of all households are classified as entrepreneurs or ‘in business’, based on answering “yes” to the question whether any household member has an own business. The rest of the households are classified as “non-business”, e.g., working for a wage.

We estimate the structural parameters of the model via the generalized method of moments (GMM) by matching observed and model-predicted occupations and income levels conditional on households’ initial wealth and education. Entrepreneurial ability is modeled as a source of unobserved heterogeneity. We match 11 moments (seven occupational choice based and six income based) and estimate nine structural parameters.

The baseline estimates indicate that nearly 11% of all households in our sample, or 17% of all households who report running a business, are classified by the model as involuntary entrepreneurs. The predicted probability (propensity) of involuntary entrepreneurship at the GMM estimates varies across the agents between as high as .59 to 0 and it is decreasing in the years of schooling and in initial wealth. That is, poorer and less educated agents are more likely to be involuntary entrepreneurs – almost half of the latter are among the households with both initial wealth and schooling below the median.

At the GMM parameter estimates, we find that the credit constraint is more likely to bind for voluntary entrepreneurs (it binds for 57% of them) than for involuntary entrepreneurs (23%). The reason is that voluntary entrepreneurs have higher entrepreneurial ability on average and hence are more likely to be credit constrained for a given wealth level. Voluntary entrepreneurs are estimated to earn significantly higher income on average (554 thousand Baht) compared to involuntary entrepreneurs (83 thousand Baht) or households not running a business (195 thousand Baht).
We also study three counterfactuals using simulated data from the model at the GMM estimates. First, we consider the elimination of the labor market constraint – that is, when each agent is always able to choose their income-maximizing occupation, as assumed in the literature (Evans and Jovanovic, 1989). Naturally, this reduces the rate of entrepreneurship in the economy since only the voluntary entrepreneurs remain. The average net income in the economy goes up by 1.8% but the income gains are unevenly spread over the income distribution, with the 10th income percentile receiving a 6% increase compared to the baseline with a binding labor market constraint, versus only 1% increase for the 90th income percentile. While all households weakly gain from the counterfactual, eliminating the labor constraint has important composition effects: lowering the average income of ex-post non-business households because of the entry of the relatively less-skilled former involuntary entrepreneurs, and raising the mean income of entrepreneurs.

The next counterfactual we consider is relaxing the credit constraint, which we perform by doubling the credit tightness parameter $\lambda$ (that determines the maximum capital level that can be borrowed and invested in the business). As with the labor constraint, this policy is Pareto improving by construction. At the GMM estimates, we find that relaxing the credit constraint has only a minor effect on the rate of involuntary entrepreneurship among those running a business (it falls from 16.6% to 16.2%). This suggests that the labor market constraint is relatively more important. However, relaxing the credit constraint has significant impact on income in the economy via enabling entrepreneurs to invest more. Mean income goes up by almost 5% accompanied with gains across the income distribution. The income gains are largest again among poorer households (9% at the 10th income percentile).

The third counterfactual we consider introduces the option for agents in the model to take a microfinance loan of size up to 10% of the median gross income in the data, $M$ (20 thousand Baht). The loan has the same interest rate as in the baseline economy, so all it does is effectively raise the credit upper bound from $\lambda z$ to $\lambda z + M$. Once again, we find that, at our GMM estimates, the microfinance policy does not change the rate of involuntary entrepreneurship by much (it falls from 16.6% to 15.8% of all entrepreneurs) but it does raise the overall rate of entrepreneurship in the model economy from 65.2% in the baseline to 66.3%. The effect of the microfinance policy is more significant on household incomes. Average income goes up by 3% but households at the bottom of the income distribution benefit more from the ability to expand their businesses (or select into a higher-income occupation) – the income gain is 14% at the 10th income percentile. We also analyze the incidence of the policy by household initial wealth and schooling and find that its effect is very uneven with the largest gains (up to 75% net income increase) occurring for the households with both very low wealth and schooling.
Related literature

Much of the existing empirical work on the topic looks at ‘voluntary’ vs. ‘involuntary’ entrepreneurs by using an ad-hoc definition based on available data. For example, Block and Wagner (2010) finds a 16% earnings premium in Germany for individuals who start a business after voluntarily leaving their previous job, compared to those who start a business after losing their previous job. Using data from six ex-USSR countries, Earle and Sakova (2000) find that own-account workers would earn more as employees, and conclude that at least some of them are occupational choice constrained. In Sri Lanka, de Mel et al (2010) find that, along a wide range of dimensions (parental and childhood background, labor history, measures of ability and risk-attitude), the majority of own-account workers resemble wage-workers more than larger firm owners.

A few papers analyze entrepreneurship in the framework of income maximization while allowing for a wage-market friction, as we do here. For example, Falco and Haywood (2013) estimate the returns to observable characteristics in self-employment vs. wage work in Ghana. They assume that ‘job queueing’ may exist in the wage market, which is modeled as an entry cost possibly depending on unobservable worker characteristics. The focus of the paper is on obtaining consistent estimates of the returns to observables and unobservables in each sector, and therefore its results are not directly comparable to ours.

Another related paper is Gunther and Launov (2012) who model observed income as a finite mixture of incomes from a segmented labor market. Accounting for selection, they model earnings in each segment as a linear function of demographic variables (sex, age, education and training, religion, etc). Using a 1998 Ivorian household survey, they conclude that the informal sector is made up of at least two latent segments. The model allows to determine the segment in which a person’s income would be maximized. They show that 44% of informal sector workers are predicted to maximize their earnings in a different labor market segment than the one they are engaged in, which is interpreted to imply that involuntary employment is a significant phenomenon in the urban labor market. Our approach differs in that, instead of a statistical model, we propose a structural economic model of involuntary entrepreneurship based on maximizing behavior subject to constraints. We are also able to distinguish between labor market-constrained and credit-constrained households.

Our paper also differs from two recent working papers on entrepreneurship in a structural setting, respectively by Banerjee et al. (2015) and Donovan (2015). Banerjee et al. use data from a microfinance randomized trial in India and define two types of entrepreneurs: “gung-ho entrepreneurs” (GE) defined as those who already owned a business before the intervention, and “reluctant entrepreneurs” (RE), defined as those without a business prior

4In their sample, 52.6% of those between the age of 15 and 65 years are inactive. In contrast, we use household level business ownership and all our households are occupied in at least one income-earning activity.
to the intervention. Their definition thus differs from our endogenous determination of voluntary vs. involuntary entrepreneurship, within the structural model. The authors estimate a model of technology choice in which REs only have access to a decreasing returns to scale technology, while GEs can, in addition, access another technology with large fixed costs but higher return. Using data on various outcome variables separately for the GEs and REs in the treatment and control neighborhoods, they find that most of the impact from the treatment is driven by the GEs who expand their businesses as opposed to REs for whom most effects are insignificant. Unlike in our paper, the author’s focus is not on determining who and how many the ‘involuntary’ entrepreneurs are (a particular ex-ante definition is used instead) but on quantifying the heterogeneity in policy outcomes.

Donovan (2015) defines “subsistence entrepreneurs” similarly to us here, as business owners who would accept a salaried job if offered. He focuses on the role of unemployment and search frictions. In his model, subsistence entrepreneurship arises as a result of low unemployment benefits and financial market imperfections. He studies the impact of the resulting misallocation between entrepreneurs and salaried workers on the firm size distribution and the impact of lending to poor entrepreneurs, calibrating his model with macroeconomic statistics from the United States. The model is tested empirically (but not estimated structurally) with microeconomic data from Chile and Mexico, finding that misallocated business owners earn lower profit conditional on observables and are more likely to have left their previous job involuntarily.

The remainder of the paper is organized in the following manner: Section 2.2 outlines the model and formally defines involuntary entrepreneurship, we provide a brief description of the data in Section 2.3, followed by the details of structural estimation, results and policy analysis in Section 2.4, and then conclude with Section 2.5.

2.2 Model

2.2.1 Preferences, endowments, and technology

Consider a large number of households (agents) who are risk-neutral and have strictly increasing preferences over expected income. The agents differ in their initial endowments of a single investment good, $z$ where $z \geq 0$. They also differ in two productive characteristics: $x \in [0, \bar{x}]$ which can be thought of as ‘schooling’ or, more generally, ‘labor market skills’, and $\theta \in [\theta_{\text{min}}, \bar{\theta}]$ which will be interpreted as entrepreneurial ‘talent’ or ability.

There are two occupations/technologies. The first is a “business”or “entrepreneurship” technology, $E$ which requires capital investment $k > 0$ and one agent to operate and yields expected output

$$q^E(\theta, z) = \theta k^\alpha$$
where \( \alpha \in (0, 1) \). There is no minimum scale or fixed costs to start up a business, so anyone with \( z > 0 \), no matter how small can be an entrepreneur.

The second occupation or technology does not require capital and yields expected output

\[
q^A(x, z) = \mu(1 + x)^\gamma.
\]

Above, the parameter \( \mu > 0 \) corresponds to what a person with labor market skills \( x = 0 \) would earn while \( \gamma \geq 0 \) governs the sensitivity of \( q^A \) to increases in \( x \). We will interpret occupation \( A \) as the “alternative” or “non-business” occupation. It can include wage work or other similar activities, the income from which increases in \( x \).

### 2.2.2 Financial market

As in Evans and Jovanovic (1989), hereafter EJ, assume that the agents have access to a financial intermediary via which they can save or borrow at gross interest rate \( r \geq 1 \). The credit market is imperfect. Due to a limited enforcement problem, the maximum amount \( k \) that an agent can invest is \( \lambda z \) where \( \lambda > 0 \) is a parameter capturing the severity of credit constraint.\(^5\) Setting \( \lambda \) very large corresponds to perfect credit markets. Setting \( \lambda = 0 \) corresponds to a missing credit market (only saving is possible). If an agent has a sufficiently large wealth, she would be able to invest the optimal amount of capital in her firm and the credit constraint will not bind. In contrast, if an agent has relatively low wealth they will be credit-constrained and invest \( \lambda z \) even though the marginal product of capital exceeds the cost of funds \( r \) at that investment level.

Agents who are employed in the \( A \) occupation do not need capital \( k \), so they save their initial endowment which results in expected income of:

\[
y^A(x, z) = \mu(1 + x)^\gamma + rz
\]

Agents who are employed in the \( E \) occupation (entrepreneurs) would either save or borrow depending on their desired investment \( k \). That is, their expected income can be written as

\[
y^E(\theta, z) = \theta k^\alpha + r(z - k)
\]

The optimal investment level \( k \) will be determined below.

### 2.2.3 Involuntary entrepreneurship

In EJ (1989), given the agents’ preferences, they would prefer the occupation which yields higher expected income. That is, absent any constraints on their choice set, an agent would

\(^5\)The \( \lambda z \) upper bound can be easily micro-founded by a limited enforcement friction, see PTK (2006) for a discussion.
choose the occupation which achieves the maximum of $y^E(\theta, z)$ and $y^A(x, z)$. Here, we depart from EJ by assuming that, depending on the agent’s characteristic $x$ (education, labor market skills), the agent’s access to occupation $A$ is restricted with some probability. For instance, agents with lower level of $x$ find it harder to find wage work; government or private sector jobs require diplomas, qualifications, certificates, etc.

Specifically, let $P_x$ be the probability with which an agent with labor market skills $x$ does not have access to occupation $A$ in the current period. That is, with probability $P_x$ such agent only has access to occupation $E$, while with probability $1 - P_x$ she can choose either $E$ or $A$. If occupation $E$ is what this agent would have chosen to maximize her expected income, then this occupational choice (labor market) constraint is not binding. However, if $y^A > y^E$ for this agent, then she will be an “involuntary” entrepreneur – someone who engages in the $E$ occupation purely out of necessity.

In our baseline setting we assume that

$$1 - P_x = \left(\frac{1 + x}{1 + \bar{x}}\right)^\eta \tag{2.1}$$

where $\bar{x}$ is the largest possible value of $x$ and $\eta \geq 0$ is a parameter that governs the tightness of the choice constraint for different values of $x$. For example, the case $\eta = 0$ corresponds to $P_x = 0$, that is, all agents are able to choose freely between both occupations, as in the EJ model. In contrast, the case $\eta < 1$ corresponds to the occupational choice constraint becoming less tight quickly for $x$ relatively low, while $\eta > 1$ corresponds to the case when the constraint is relaxed only for relatively large values of $x$.

The economic interpretation of (2.1) is that agents with higher education or other labor market skills, $x$ are more likely to have access to both occupations in a given moment of time. This is also consistent with the “queuing for jobs” interpretation of Falco and Haywood (2013) mentioned in the introduction.

### 2.2.4 Optimal investment and occupational choice

Remember that income from entrepreneurship equals

$$y^E(\theta, z) = \theta k^\alpha - r(z - k).$$

Thus, if the credit constraint is not binding, an agent with initial wealth $z$ and ability $\theta$ would optimally want to invest the first-best (unconstrained) capital amount, $k_u(\theta)$ solving

$$k_u(\theta) \equiv \arg \max_k \{\theta k^\alpha - r k\} = \left(\frac{\theta \alpha}{r}\right) \frac{1}{1 - \alpha} \tag{2.2}$$

Note that $k_u(\theta)$ is increasing in $\theta$, implying that more talented entrepreneurs would want to invest more. The first-best capital amount does not depend on the entrepreneur’s initial
wealth \( z \). Intuitively, in a world without credit constraints, all productive projects will be financed at the levels of \( k \) that equalize marginal product to marginal cost.

With credit constraints, however, the first-best level of capital is only feasible if \( k_u(\theta) = (\theta^\alpha \lambda)^{1-\alpha} \leq \lambda z \). Define \( B(z) \) as the threshold level of entrepreneurial talent \( \theta \) at which \( k_u(\theta) = \lambda z \), that is,

\[
B(z) \equiv \frac{r}{\alpha} (\lambda z)^{1-\alpha} \quad (2.3)
\]

In other words, for a given initial wealth \( z \), the value \( B(z) \) is the maximum level of talent \( \theta \) at which an agent is financially unconstrained and able to invest \( k_u(\theta) \). For given wealth \( z \), note that the credit constraint is therefore more likely to bind for more talented entrepreneurs.

If \( \theta > B(z) \) the agent would optimally invest the maximum possible amount, \( \lambda z \) which is less than \( k_u(\theta) \) since the marginal product of capital exceeds the marginal cost.

We therefore obtain,

\[
y^E(\theta, z) = \begin{cases} 
\theta [k_u(\theta)]^\alpha + r(z - k_u(\theta)) & \text{if } \theta \leq B(z) \\
(1-\alpha)\theta^\frac{1-\alpha}{\alpha} (\frac{\theta}{r})^\frac{\alpha}{\alpha-1} & \text{if } \theta > B(z) 
\end{cases}
\]

or, \( y^E(\theta, z) - rz = \begin{cases} 
\theta(\lambda z)^\alpha + r(z - \lambda z) & \text{if } \theta \leq B(z) \\
(1-\alpha)\theta^\frac{1-\alpha}{\alpha} (\frac{\theta}{r})^\frac{\alpha}{\alpha-1} - \lambda rz & \text{if } \theta > B(z) 
\end{cases} \)

Alternatively, an agent who has access to the \( A \) (non-business) occupation earns,

\[
y^A(z, x) = \mu(1 + x)^\gamma + rz
\]

Define

\[
\Delta(z, \theta, x) \equiv y^E(\theta, z) - y^A(z, x)
\]

as the expected income differential between entrepreneurship and the alternative occupation. We obtain the following result, which essentially re-states the main trade-off from Evans and Jovanovic (1989)'s paper on credit-constrained occupational choice.

**Proposition 1**

An agent with initial wealth \( z \) and characteristics \( (\theta, x) \) who has access to both the entrepreneurial and alternative occupations, \( E \) and \( A \) would optimally choose entrepreneurship, \( E \) if

\[
\Delta(z, \theta, x) \geq 0 \iff \begin{cases} 
\theta \geq A(x) & \text{if } \theta \leq B(z) \\
\theta \geq C(x, z) & \text{if } \theta > B(z) 
\end{cases} \quad (2.4)
\]

where \( A(x) \equiv (\frac{\mu}{1+\alpha})^{1-\alpha}(1+x)^\gamma(1-\alpha)(\frac{x}{\alpha})^\alpha \), \( B(z) \equiv \frac{r}{\alpha} (\lambda z)^{1-\alpha} \), and \( C(z, x) \equiv (\lambda z)^{-\alpha} [\mu(1 + x)^\gamma + r\lambda z] \).

**Proof:** see Appendix B.
2.2.5 The probability of entrepreneurship

We follow the literature and assume that the entrepreneurial ability \( \theta \), while observable to the agents in the model, is unobservable to the econometrician. In contrast initial wealth \( z \) and the labor market characteristics \( x \) are known to all. Thus, for a given distribution of \( \theta \) and given \( z \) and \( x \), the model implies a probability that an agent chooses to be an entrepreneur (occupation \( E \)) or not (occupation \( A \)). In section 2.4, we compute and use these predicted probabilities to estimate the structural parameters of the model based on the observed occupational status of households in the Thai urban data. In addition, the model has implications for the probability/fraction of involuntary entrepreneurs which is a status unobserved in the data.

Proposition 1 implies that, suppressing the arguments in expressions \( A(x) \), \( B(z) \), \( C(z, x) \), and \( \Delta(\theta, z, x) \) and using \( P \) to denote probabilities,

\[
P(\Delta \geq 0) = P(\Delta \geq 0|\theta > B)P(\theta > B) + P(\Delta \geq 0|\theta \leq B)P(\theta \leq B)
\]

\[
= P(\theta > C|\theta > B)P(\theta > B) + P(\theta \geq A|\theta \leq B)P(\theta \leq B)
\]

\[
= P(\theta \geq C \land \theta > B) + P(\theta \geq A \land \theta \leq B) \tag{2.5}
\]

For a given initial wealth \( z \) and labor market skills \( x \), the exact ordering of \( A(x) \), \( B(z) \) and \( C(z, x) \) is completely determined.

To compute the probabilities in (2.5), we need an assumption on the distribution of the unobserved heterogeneity variable \( \theta \) (entrepreneurial talent). We follow PTK (2006) and assume,

\[
\ln \theta = \delta_0 + \delta_1 \ln z + \delta_2 \ln(1 + x) + \varepsilon \tag{2.6}
\]

where \( \varepsilon|z, x \sim N(0, \sigma) \)

The interpretation is that entrepreneurial ability can be correlated with initial wealth \( z \) and the observable characteristics \( x \) (in the estimation \( x \) is proxied by years of schooling of the principal earner) but we also allow a random talent component, \( \varepsilon \). The distributional parameters \( \delta_0, \delta_1, \delta_2 \) and \( \sigma \) will be estimated together with the rest of the structural parameters such as \( \alpha, \gamma, \lambda, \mu \).

Let \( 1_{B>A} \) denote the indicator function equaling one if \( B > A \) for given \( (x, z) \) and zero otherwise. It is easy to show that, for any given \( (x, z) \), the inequality \( B > A \) is mathematically equivalent to the inequality \( B > C \). Denote the conditional expectation of \( \ln \theta \) by

\[
\hat{\theta}(z, x) = \delta_0 + \delta_1 \ln z + \delta_2 \ln(1 + x),
\]

we obtain the following result.
Table 2.1: Voluntary and involuntary entrepreneurship

<table>
<thead>
<tr>
<th>$\Delta \geq 0$</th>
<th>$1_E = 0$</th>
<th>voluntary entrep.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta \leq 0$</td>
<td>$1_E = 1$</td>
<td>involuntary entrep.</td>
</tr>
</tbody>
</table>

**Lemma 1**

For an agent with characteristics $(z, x)$ who has access to both the entrepreneurial and alternative occupations $E$ and $A$ the probability (likelihood) of choosing entrepreneurship equals,

$$\tilde{P}_E \equiv P(\Delta(\theta, z, x) \geq 0) = 1_{B>A}(1 - \Phi(a)) + (1 - 1_{B>A})(1 - \Phi(c))$$  \hspace{1cm} (2.7)

where $a \equiv \frac{\ln A(x) - \tilde{\theta}(z, x)}{\sigma}$ and $c \equiv \frac{\ln C(z, x) - \tilde{\theta}(z, x)}{\sigma}$.

### 2.2.6 Involuntary entrepreneurship

Denote by $1_E$ the indicator function for choosing entrepreneurship in the model, conditional on $x$ and $z$. By the Law of total probability:

$$P(1_E = 1) = P(1_E = 1|\Delta \geq 0)P(\Delta \geq 0) + P(1_E = 1|\Delta < 0)P(\Delta < 0)$$

where $P(\Delta \geq 0)$ is given by (2.7) in Lemma 1.

According to the model, $P(1_E = 1|\Delta \geq 0) = 1$, since any agent who can choose between the two occupations and earns higher expected income by being entrepreneur ($\Delta \geq 0$) would choose occupation $E$. The model also implies that $P(1_E = 1|\Delta < 0) = P_x$ where $P_x$ was defined in (2.1) in Section 2.3. Therefore,

$$P_E \equiv P(1_E = 1) = P(\Delta \geq 0) + P_xP(\Delta < 0)$$  \hspace{1cm} (2.8)

That is, the overall probability of entrepreneurship $P_E$ can be decomposed into two terms. The first term, $P(\Delta \geq 0)$ corresponds to the probability of entrepreneurship conditional on $z$ and $x$ that would arise if all agents chose occupation $E$ based solely on expected income maximization, as assumed in the previous literature, for example EJ (1989). The second term $P_{IE} \equiv P_xP(\Delta < 0)$ is the additional probability of entrepreneurship, relative to the basic income maximization model, which we interpret as the probability/fraction of involuntary entrepreneurship. Note that $P_{IE}$ is a function of an agent’s initial wealth $z$, her labor market skills $x$, and the structural and distributional parameters of our model. The following table summarizes the above discussion.
2.3 Data

We use data from the Townsend Thai Project’s 2005 Urban Annual Resurvey. The main outcome variable of interest is household business ownership. We measure business ownership in the data in terms of whether a household reports that they own at least one business at the time of the survey. Thus, it is a binary variable equal to one if the household owns a business and zero otherwise. The corresponding variable in the model is $1_E$. In the estimation we also use the annual gross income of households, defined as their income excluding transfers from remittance, government programs, and interest. The model counterparts are $q^E = \theta k^\alpha$ and $q^A = \mu(1 + x)^\gamma$ for business and not business households, respectively.

Initial household wealth, $z$ is measured as the total value (in 2005 Thai baht) of landholdings, household durables and agricultural assets owned by a household five years prior to the survey. The reason for this back-dating is to avoid potential simultaneity problems between occupational status and current wealth. Recall that, in the model, initial wealth $z$ affects the investment potential of a household. We are therefore assuming that the level of pre-existing, year 2000 wealth measure we construct is free of reverse causality. Also, note that in the model, we allow initial wealth to be correlated with the level of entrepreneurial talent $\theta$, and therefore, we capture the possibility that more talented households may save more in anticipation of becoming business owners.

We proxy the model variable $x$, interpreted as education or other labor market characteristics or qualifications, by the years of schooling of the principal earner in the household. We use data on individual occupations and worker type within the household to identify the principal earner. For business households (regardless of whether the household also earns wage income or not), the principal earner is the member whose occupation and worker-type matches the reported business ownership type. For households running more than one business, the principal earner is defined as the business owner associated with the largest business in terms of assets. For non-business households, the principal earner is defined as the wage-earning member (for households with multiple wage-earners, the principal earner is the member earning the highest monthly wage income).

From the full survey sample, we construct a sub-sample that we use to estimate the model as follows. We exclude all households in the top one percentile of the initial wealth distribution, all households with zero initial wealth or zero gross income, and all households for which the principal earner could not be identified.

---

6Full details are available at cier.uchicago.edu.
7In Section XIII of the survey, a description of household business is reported in variables BA3A to BA3D. Information on the occupation of all household members above 15 years of age is reported in Section V. For business households, we identify the member whose occupation description (e.g. barber or rice farmer) and type (e.g. owner of business or employee) is closest to the description of the household business (e.g. salon or restaurant).
8Because of data limitations, we were not able to identify a principal earner for about 15% of all surveyed households.
Table 2.2: Occupation and source of income

<table>
<thead>
<tr>
<th>Self-reported business ownership</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>786</td>
<td>66.1</td>
</tr>
<tr>
<td>no</td>
<td>403</td>
<td>33.9</td>
</tr>
<tr>
<td>total</td>
<td>1,189</td>
<td>100</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Major source of annual gross income</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>business</td>
<td>588</td>
<td>49.4</td>
</tr>
<tr>
<td>wage</td>
<td>436</td>
<td>36.7</td>
</tr>
<tr>
<td>farming</td>
<td>27</td>
<td>2.3</td>
</tr>
<tr>
<td>transfers</td>
<td>108</td>
<td>9.1</td>
</tr>
<tr>
<td>other</td>
<td>30</td>
<td>2.5</td>
</tr>
<tr>
<td>total</td>
<td>1,189</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 2.2 shows that 66% of the households in our final sample report running a business. Using data on income sources, Table 2 also shows that business and wage work are the most important sources of income for households. 49.4% of all households in our sample derive the majority of their annual (gross) income from running a business while 36.7% of all households do so from wages. A small number, 2.3% of households derive the largest part of their income from farming (rice, other crops, and livestock-raising). Table 2.3 summarizes the key variables used in reduced form and structural estimation, with Table 2.4 reporting the coefficient estimates of a probit regression of business ownership on initial wealth, schooling, and additional household characteristics. The results indicate that both initial wealth (asset five years prior to survey year) and principal earners’ schooling respectively increase the probability of business ownership (at a decreasing rate). Households with female and older principal earners, and larger household size are more likely to be business owners.

2.4 Structural estimation

We have a sample of $N$ households, $i = 1, \ldots, N$ for whom we have data on their initial income $z_i$, the years of schooling of the principal earner, $x_i$ and occupational status, $E_i$ (with $E_i = 1$ if the household runs a business and zero otherwise). We estimate the structural parameters (technology and credit access) as well as the distributional parameters of $\theta$ via the generalized method of moments by matching a set of probability of entrepreneurship and income moments in the model to their data counterparts for given observed $x_i$ and $z_i$ (see below for details).

Specifically, the nine parameters we estimate are: $\alpha$ (the elasticity of business revenue with respect to capital), $\lambda$ (the credit constraint parameter), $\gamma$ (the elasticity of non-business income with respect to $x$), $\eta$ (the parameter governing the occupational choice constraint in
Table 2.3: Summary statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Business</th>
<th>Non-business</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>wealth 5 years ago (000s of baht)*</td>
<td>620.5</td>
<td>469.4</td>
<td>569.3</td>
</tr>
<tr>
<td></td>
<td>(814.8)</td>
<td>(682.3)</td>
<td>(775.5)</td>
</tr>
<tr>
<td></td>
<td>[335.1]</td>
<td>[235.1]</td>
<td>[305.0]</td>
</tr>
<tr>
<td>annual gross income (000s of baht)*</td>
<td>513.6</td>
<td>164.7</td>
<td>395.3</td>
</tr>
<tr>
<td></td>
<td>(1313.2)</td>
<td>(132.5)</td>
<td>(1074.9)</td>
</tr>
<tr>
<td></td>
<td>[276.8]</td>
<td>[126.0]</td>
<td>[200.8]</td>
</tr>
<tr>
<td>schooling of principal earner (years)*</td>
<td>7.3</td>
<td>9.8</td>
<td>8.1</td>
</tr>
<tr>
<td></td>
<td>(4.0)</td>
<td>(4.7)</td>
<td>(4.5)</td>
</tr>
<tr>
<td>age of principle earner*</td>
<td>49.4</td>
<td>41.2</td>
<td>46.6</td>
</tr>
<tr>
<td></td>
<td>(11.0)</td>
<td>(13.1)</td>
<td>(12.3)</td>
</tr>
<tr>
<td>male (gender of principal earner)*</td>
<td>0.45</td>
<td>0.59</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>(0.50)</td>
<td>(0.49)</td>
<td>(0.50)</td>
</tr>
<tr>
<td>household size</td>
<td>4.28</td>
<td>4.35</td>
<td>4.30</td>
</tr>
<tr>
<td></td>
<td>(1.90)</td>
<td>(1.83)</td>
<td>(1.87)</td>
</tr>
<tr>
<td>sample size</td>
<td>786</td>
<td>403</td>
<td>1189</td>
</tr>
<tr>
<td>sample proportion</td>
<td>66.1%</td>
<td>33.9%</td>
<td>100%</td>
</tr>
</tbody>
</table>

The sample excludes the top percentile of wealth, households with zero income, and where a principle earner could not be identified.

Mean, standard deviation (in paranthesis) for all variables, median (in brackets) for monetary values. Wealth and income are reported in thousands of 2005 Thai baht.

(*) difference-in-means test between business and non-business is significant at the 1% level.
Table 2.4: Probit regression of business ownership

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>wealth 5 years ago (mln Baht)</td>
<td>0.431***</td>
</tr>
<tr>
<td></td>
<td>(0.137)</td>
</tr>
<tr>
<td>square of wealth 5 years ago</td>
<td>-0.070***</td>
</tr>
<tr>
<td></td>
<td>(0.033)</td>
</tr>
<tr>
<td>schooling of principle earner</td>
<td>0.197***</td>
</tr>
<tr>
<td></td>
<td>(0.044)</td>
</tr>
<tr>
<td>schooling squared</td>
<td>-0.015***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
</tr>
<tr>
<td>age of principle earner</td>
<td>0.028***</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
</tr>
<tr>
<td>male (gender of principle earner)</td>
<td>-0.450***</td>
</tr>
<tr>
<td></td>
<td>(0.084)</td>
</tr>
<tr>
<td>household size</td>
<td>0.040**</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
</tr>
<tr>
<td>province = Chachoengsao</td>
<td>-0.127</td>
</tr>
<tr>
<td></td>
<td>(0.144)</td>
</tr>
<tr>
<td>province = Buriram</td>
<td>-0.437***</td>
</tr>
<tr>
<td></td>
<td>(0.142)</td>
</tr>
<tr>
<td>province = Sisaket</td>
<td>-0.006</td>
</tr>
<tr>
<td></td>
<td>(0.154)</td>
</tr>
<tr>
<td>province = Phrae</td>
<td>-0.478***</td>
</tr>
<tr>
<td></td>
<td>(0.138)</td>
</tr>
<tr>
<td>province = Satun</td>
<td>-0.762***</td>
</tr>
<tr>
<td></td>
<td>(0.145)</td>
</tr>
<tr>
<td>intercept</td>
<td>-0.992*</td>
</tr>
<tr>
<td></td>
<td>(0.319)</td>
</tr>
<tr>
<td>sample size</td>
<td>1189</td>
</tr>
</tbody>
</table>

Notes: Dependent variable is an indicator for whether household reports owning a business in 2005. Base category for province is Lopburi. * p<0.10, ** p<0.05, *** p<0.01.
$P_x$, $\mu$ (the scale parameter of non-business income), $\delta_0$ (the conditional mean of log talent), $\delta_1$ and $\delta_2$ (the elasticities of log ability with respect to initial wealth and schooling), and $\sigma$ (the standard deviation of the log-talent distribution). Call the vector of all estimated parameters $\phi \equiv (\alpha, \lambda, \gamma, \eta, \mu, \delta_0, \delta_1, \delta_2, \sigma)$.

We fix the interest rate parameter $r$ to 1.06, which corresponds to the median level of interest on household loans in the data.

2.4.1 GMM - matched moments and computation

The model parameters are estimated by minimizing the percentage deviation between various moments predicted by the model and their respective sample analogs. Given parameters $\phi$, denote the model-predicted moments by $h_j(z,x,\phi)$ for $j = 1, \ldots, J$ and their respective sample analogs by $h^d_j$. Define the percentage deviation of the model predicted moment from its sample analog as

$$q_j(z,x,\phi) \equiv \frac{h_j(z,x,\phi) - h^d_j}{h^d_j}, \quad j = 1, \ldots, J$$

Construct $q(z,x,\phi)$ as the $J \times 1$ vector of percentage deviations between the model-predicted moments and their sample analogs. The generalized method of moments (GMM) estimates are computed by minimizing the criterion function $q(z,x,\phi)'q(z,x,\phi)$ over $\phi$. We use an optimization routine robust to local extremes, initialized at the results from an extensive grid search over the parameter space.

For our baseline estimates, we match the 11 moments listed in Table 5 below. The moments are matched by choice of 9 parameters. The first seven moments correspond to the probabilities of business ownership in different sub-samples defined based on the terciles of years of schooling ($x$) and initial wealth ($z$). The model-predicted proportion of business owners for a subset of initial wealth levels $z_i \in Z$ and years of schooling $x_i \in X$ is

$$\frac{\sum_{i=1}^{N} 1\{z_i \in Z, x_i \in X\} P(1_E = 1|z_i, x_i, \phi)}{\sum_{i=1}^{N} 1\{z_i \in Z, x_i \in X\}}$$

where $P(1_E = 1|z_i, x_i, \phi)$ is computed using equation (2.8). The sample analog is the observed fraction of all business owners (those with $E_i = 1$) with $z_i \in Z$ and $x_i \in X$. The remaining four moments which we match correspond to the average gross incomes of

---

9Similar to the case in the first chapter, we found GMM as a suitable compromise; it is analytically and computationally easier and allows us to use information on both occupational choice and income from the data. In particular, while the optimization routine for MLE based only on occupational choice does converge, the average incomes of business and non-business households implied by the model evaluated at the MLE were far from those in the data.

10We first use a grid search over approximately 20,000 parameter configurations. We then initialize the Matlab minimizer `particleswarm` with an initial population of the 20 best-fitting parameter vectors from the grid search.
business and non-business households in the full sample or stratified by initial wealth and schooling. For example, the average expected gross income of a business household in the model is

$$\frac{\sum_{i=1}^{N} E(q|1_E = 1, z_i, x_i, \phi)}{\sum_{i=1}^{N} P(1_E = 1|z_i, x_i, \phi)}$$

where the expectation in the numerator is taken over $\varepsilon$. The sample analogs are obtained by replacing $P(1_E = 1|z_i, x_i, \phi)$ with the observed business status, $E_i$ and by replacing the conditional gross income expectation $E(q|1_E = o, z_i, x_i, \phi)$ for $o = \{0, 1\}$ by the average observed income of non-business or business households. The expected gross income $E(q|1_E = 1, z_i, x_i, \phi)$ is computed in the appendices.

### 2.4.2 Estimates and model fit

Table 2.6 reports the GMM parameter estimates. The return to capital in entrepreneurial income, $\alpha$, is estimated at 0.23, implying that a 10 percent increase in capital $k$ would lead to an approximately 2.2% percent increase in entrepreneurial income of unconstrained entrepreneurs, all else equal. The estimate of the credit constraint parameter $\lambda$ is 0.23, which implies that, for a household with initial wealth $z$ equal to the median, the maximum business investment it can make is about 70,000 Thai baht. As a comparison, the median business assets in the data (for business owning households) is about 19.7 thousand baht, or close to 6.5% of median initial wealth. The choice constraint parameter $\eta$ is estimated to be 0.41. At the modal schooling level $x = 4$, this implies a 41% probability that an agent is constrained in her income-maximizing occupational choice. Entrepreneurial talent $\theta$ is found to be weakly positively related with both initial wealth and years of schooling (the estimates of $\delta_1$ and $\delta_2$ are positive).

Table 2.7 reports the model predictions evaluated at the GMM estimates. We compute these statistics by simulating data from the model at the GMM parameter estimates by drawing 100 random values from the distribution of the shock $\varepsilon$ for each $i = 1, ..., N$. We then average, first over $\varepsilon$ for each $i$, and then over the chosen stratification of agents to compute the various statistics reported in Table 7. The proportion of involuntary entrepreneurs from all households in the sample, as defined in Table 2.1, is 10.8%. In other words, 16.6% of the business owners in the sample are classified as involuntary entrepreneurs. The remainder, 54.4% of all agents or 83.4% of all business owners are classified as voluntary entrepreneurs at our GMM estimates. Approximately 51% of all entrepreneurs are estimated to be credit constrained, that is their capital investment $k$ equals $\lambda$ times their initial wealth, $z$ and they invest less than their unconstrained optimum. The fraction of credit constrained is large among the voluntary entrepreneurs (57%), while much fewer (23%) of involuntary entrepreneurs are credit constrained. The reason is that voluntary entrepreneurs have higher entrepreneurial ability $\theta$ on average, and hence larger unconstrained capital level. Indeed, in
Table 2.5: List of matched moments

<table>
<thead>
<tr>
<th>Moment</th>
<th>Model</th>
<th>Sample analog</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. average probability of entrepreneurship</td>
<td>( \frac{1}{N} \sum_{i=1}^{N} P(1_{E_i} = 1\mid z_i, x_i, \phi) )</td>
<td>( \frac{1}{N} \sum_{i=1}^{N} E_i )</td>
</tr>
<tr>
<td>2. prob. of entrepreneurship, ( x \leq x_{t1} )</td>
<td>( \sum_{1 \leq i \leq r_1, z_i \leq e_1} P(1_{E_i} = 1\mid z_i, x_i, \phi) )</td>
<td>( \sum_{1 \leq i \leq r_1} E_i )</td>
</tr>
<tr>
<td>3. prob. of entrepreneurship, ( z \leq z_{t1} )</td>
<td>( \sum_{1 \leq i \leq r_1} P(1_{E_i} = 1\mid z_i, x_i, \phi) )</td>
<td>( \sum_{1 \leq i \leq r_1} E_i )</td>
</tr>
<tr>
<td>4. prob. of entrepreneurship, ( x \geq x_{t3} )</td>
<td>( \sum_{1 \leq i \leq r_3} P(1_{E_i} = 1\mid z_i, x_i, \phi) )</td>
<td>( \sum_{1 \leq i \leq r_3} E_i )</td>
</tr>
<tr>
<td>5. prob. of entrepreneurship, ( z \geq z_{t3} )</td>
<td>( \sum_{1 \leq i \leq r_3} P(1_{E_i} = 1\mid z_i, x_i, \phi) )</td>
<td>( \sum_{1 \leq i \leq r_3} E_i )</td>
</tr>
<tr>
<td>6. prob. of entrep., ( z \leq z_{t1}, x \leq x_{t1} )</td>
<td>( \sum_{1 \leq i \leq r_1, z_i \leq e_1} P(1_{E_i} = 1\mid z_i, x_i, \phi) )</td>
<td>( \sum_{1 \leq i \leq r_1, z_i \leq e_1} E_i )</td>
</tr>
<tr>
<td>7. prob. of entrep., ( z \geq z_{t3}, x \geq x_{t3} )</td>
<td>( \sum_{1 \leq i \leq r_3} P(1_{E_i} = 1\mid z_i, x_i, \phi) )</td>
<td>( \sum_{1 \leq i \leq r_3} E_i )</td>
</tr>
<tr>
<td>8. average gross income, entrep.</td>
<td>( \frac{1}{N} \sum_{i=1}^{N} E_i q_i E_i )</td>
<td>( \sum_{1 \leq i \leq r_1} q_i E_i )</td>
</tr>
<tr>
<td>9. average gross income, non-entrep.</td>
<td>( \frac{1}{N} \sum_{i=1}^{N} E_i q_i E_i )</td>
<td>( \sum_{1 \leq i \leq r_1} q_i E_i )</td>
</tr>
<tr>
<td>10. average gross income, entrep., ( z \leq z_m )</td>
<td>( \sum_{1 \leq i \leq r_1} z \leq z_m q_i E_i )</td>
<td>( \sum_{1 \leq i \leq r_1} z \leq z_m E_i )</td>
</tr>
<tr>
<td>11. average gross income, entrep., ( x \leq x_m )</td>
<td>( \sum_{1 \leq i \leq r_1} x \leq x_m q_i E_i )</td>
<td>( \sum_{1 \leq i \leq r_1} x \leq x_m E_i )</td>
</tr>
</tbody>
</table>

\( x \) = years of schooling; \( z \) = initial wealth.

Subscript \( m \) denotes median; \( t_1 = 33rd \) percentile; \( t_3 = 67th \) percentile.
Table 2.6: Structural estimates and model fit

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>s.e</th>
</tr>
</thead>
<tbody>
<tr>
<td>return to capital in business income</td>
<td>$\alpha$</td>
<td>0.227</td>
</tr>
<tr>
<td>credit constraint parameter</td>
<td>$\lambda$</td>
<td>0.233</td>
</tr>
<tr>
<td>return to schooling in non-business income</td>
<td>$\gamma$</td>
<td>0.747</td>
</tr>
<tr>
<td>tightness of the choice constraint</td>
<td>$\eta$</td>
<td>0.407</td>
</tr>
<tr>
<td>non-business income parameter</td>
<td>$\mu$</td>
<td>28.5</td>
</tr>
<tr>
<td>talent – constant</td>
<td>$\delta_0$</td>
<td>3.42</td>
</tr>
<tr>
<td>talent – elasticity w.r.t. initial wealth</td>
<td>$\delta_1$</td>
<td>0.129</td>
</tr>
<tr>
<td>talent – elasticity w.r.t. schooling</td>
<td>$\delta_2$</td>
<td>0.168</td>
</tr>
<tr>
<td>talent – standard deviation</td>
<td>$\sigma$</td>
<td>0.956</td>
</tr>
</tbody>
</table>

Standard errors are calculated from 99 bootstrap samples.

Table 2.7: Model predictions at the GMM estimates

<table>
<thead>
<tr>
<th>Model statistic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>entrepreneurs, % of all agents</td>
<td>65.2</td>
</tr>
<tr>
<td>involuntary entrepreneurs, % of all agents</td>
<td>10.8</td>
</tr>
<tr>
<td>involuntary entrepreneurs, % of all entrepreneurs</td>
<td>16.6</td>
</tr>
<tr>
<td>voluntary entrepreneurs, % of all agents</td>
<td>54.4</td>
</tr>
<tr>
<td>voluntary entrepreneurs, % of all entrepreneurs</td>
<td>83.4</td>
</tr>
<tr>
<td>credit constrained, % of all entrepreneurs</td>
<td>51.3</td>
</tr>
<tr>
<td>credit constrained, % of voluntary entrepreneurs</td>
<td>56.8</td>
</tr>
<tr>
<td>credit constrained, % of involuntary entrepreneurs</td>
<td>23.4</td>
</tr>
</tbody>
</table>

the simulated data the average log talent ($\log \theta$) at the GMM estimates is 5.2 for voluntary entrepreneurs, versus 3.5 for involuntary entrepreneurs and 3.7 for non-entrepreneurs.

The next table (Table 2.8) breaks down the distribution of voluntary and involuntary entrepreneurs in the model by initial wealth, $z$ and years of schooling, $x$ (both taken from the data). The reported percentages in the Table use the same simulated data from the model at the GMM estimates which was used in Table 7. We see that the majority (57.4%) of voluntary entrepreneurs have wealth above the median. This is intuitive, as larger wealth makes it less probable that an entrepreneur will be credit constrained and hence prefer the alternative occupation. This effect is emphasized for schooling above the median, since in that case the alternative income is larger and thus the agents need higher $z$ to be able to invest a sufficient amount to earn higher income as entrepreneurs. The distribution of voluntary entrepreneurs over years of schooling is closer to half and half. The smallest fraction of voluntary entrepreneurs is observed among agents with wealth below the median and schooling above the median. Intuitively, these agents are the most likely to be credit constrained and have a larger potential non-business income.

Looking at involuntary entrepreneurs (panel B of Table 2.8), we see that a large majority (over 70%) have years of schooling below the median (6 years) and also more than 60% have
wealth below the median. There are two reasons for this. First, from our assumptions, the occupational choice constraint which forces households into involuntary entrepreneurship is more restrictive for lower schooling \( x \). Second, having lower wealth \( z \) makes it more likely that one would be credit constrained if one chose to start a business, and hence prefer the alternative occupation. Indeed, in the simulated data 70% of all credit-constrained involuntary entrepreneurs have both wealth and schooling below the median while none of the credit-constrained involuntary entrepreneurs have wealth above the median.

Figure 2.1 shows the estimated relationship between initial wealth (in logs) and entrepreneurship and illustrates how our model differs from the purely income maximization model of EJ (1989). The left panel shows the relationship between initial wealth and entrepreneurship overall – it is positive but there is a lot of ‘noise’. In contrast, the relationship between initial wealth and voluntary entrepreneurship is strongly positive (the middle panel). This is the picture familiar from EJ (1989) and others, interpreted as indicative of the presence of financial constraints. We see that the relationship between initial wealth and entrepreneurship is made weaker by the presence of a negative relationship between initial wealth and involuntary entrepreneurship (the right-most panel).

Model fit

We next assess the model fit to the data, at the GMM parameter estimates. In Table 2.9, we report the model fit for the 11 chosen moments (defined in Table 2.5 above), that we target in the GMM estimation routine by minimizing the criterion function over the nine parameters in \( \phi \). We see that the seven moments based on the percentage of entrepreneurs are are all matched within 5% of their counterpart values in the data. The four income moments are matched even better, within 0.5% of their data counterparts.

We next look at additional moments, corresponding to other important dimensions of the model that we did not target in the GMM estimation (see Table 2.10). A good fit in these moments can be interpreted as a ‘validation’ of the model with data that have not

<table>
<thead>
<tr>
<th>A. Percent of voluntary entrepreneurs with wealth, ( z \leq \text{median} )</th>
<th>wealth, ( z &gt; \text{median} )</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>schooling, ( x \leq \text{median} )</td>
<td>29.2</td>
<td>27.1</td>
</tr>
<tr>
<td>schooling, ( x &gt; \text{median} )</td>
<td>15.4</td>
<td>28.3</td>
</tr>
<tr>
<td>total</td>
<td>44.6</td>
<td>55.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B. Percent of involuntary entrepreneurs with wealth, ( z \leq \text{median} )</th>
<th>wealth, ( z &gt; \text{median} )</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>schooling, ( x \leq \text{median} )</td>
<td>45.6</td>
<td>25.0</td>
</tr>
<tr>
<td>schooling, ( x &gt; \text{median} )</td>
<td>15.6</td>
<td>13.8</td>
</tr>
<tr>
<td>total</td>
<td>61.2</td>
<td>38.8</td>
</tr>
</tbody>
</table>
Figure 2.1: Probability of entrepreneurship as function of wealth

![Graph depicting the probability of entrepreneurship as a function of wealth.]

Table 2.9: Model fit: matched moments at GMM estimates

<table>
<thead>
<tr>
<th>moment</th>
<th>model</th>
<th>data</th>
<th>% deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. % entrepreneurs</td>
<td>65.2</td>
<td>66.1</td>
<td>-1.38</td>
</tr>
<tr>
<td>2. % entrepreneurs, x in bottom tercile</td>
<td>78.7</td>
<td>79.5</td>
<td>-1.03</td>
</tr>
<tr>
<td>3. % entrepreneurs, z in bottom tercile</td>
<td>59.5</td>
<td>58.9</td>
<td>1.03</td>
</tr>
<tr>
<td>4. % entrepreneurs, x in top tercile</td>
<td>50.5</td>
<td>52.0</td>
<td>-2.90</td>
</tr>
<tr>
<td>5. % entrepreneurs, z in top tercile</td>
<td>69.0</td>
<td>71.9</td>
<td>-4.13</td>
</tr>
<tr>
<td>6. % entr., z and x in bottom terciles</td>
<td>74.1</td>
<td>72.5</td>
<td>2.31</td>
</tr>
<tr>
<td>7. % entr., z and x in top terciles</td>
<td>57.0</td>
<td>54.3</td>
<td>4.90</td>
</tr>
<tr>
<td>8. average gross income – entrepreneurs</td>
<td>512.5</td>
<td>513.6</td>
<td>-0.22</td>
</tr>
<tr>
<td>9. average gross income – non-entrepreneurs</td>
<td>164.8</td>
<td>164.7</td>
<td>0.08</td>
</tr>
<tr>
<td>10. avg. gross income – entr., z below median</td>
<td>349.7</td>
<td>350.3</td>
<td>-0.17</td>
</tr>
<tr>
<td>11. avg. gross income – entr., x below median</td>
<td>387.1</td>
<td>385.7</td>
<td>0.38</td>
</tr>
</tbody>
</table>

GMM criterion value (sum of squared deviations) 5.9(10^{-3})

J-statistic 11.65

Notes: \( x_m \) = median \( x \), \( z_m \) = median \( z \); income levels in thousands

Table 2.10 indicates that the model fits well (within 5% deviation) in most of these additional dimensions. It fits slightly worse the percent of entrepreneurs with wealth below the median and schooling either above or below the median (lines 7 and 9 in Table 2.10). The model does not match so well the entrepreneurial income for households with both wealth and schooling below or above the median (lines 14 and 15 in Table 2.10).

Figure 2.2 further clarifies the patterns we see in Table 2.10 with regards to where the model succeeds or fails to match the predicted probability/fraction of entrepreneurship relative to what is observed in the data. The figure plots lowess regression lines and confidence

---

\[11\]Moments 5,6,12, and 13 in Table 2.10 can be constructed from the matched moments in Table 2.9 and so should fit well by construction if their complements from Table 2.9 are well fitted.
Table 2.10: Model fit: non-matched moments at GMM estimates

<table>
<thead>
<tr>
<th>moment</th>
<th>model</th>
<th>data</th>
<th>% deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. % entrepreneurs, $x$ below median</td>
<td>77.3</td>
<td>74.8</td>
<td>3.36</td>
</tr>
<tr>
<td>2. % entrepreneurs, $z$ below median</td>
<td>61.5</td>
<td>60.3</td>
<td>1.86</td>
</tr>
<tr>
<td>3. % entrepreneurs, $x$ above median</td>
<td>55.3</td>
<td>57.4</td>
<td>-3.62</td>
</tr>
<tr>
<td>4. % entrepreneurs, $z$ above median</td>
<td>71.2</td>
<td>71.9</td>
<td>-0.95</td>
</tr>
<tr>
<td>5. % entrepreneurs, $x$ in 2nd tercile</td>
<td>64.7</td>
<td>64.1</td>
<td>0.96</td>
</tr>
<tr>
<td>6. % entrepreneurs, $z$ in 2nd tercile</td>
<td>67.9</td>
<td>67.4</td>
<td>0.65</td>
</tr>
<tr>
<td>7. % entr., $z$ below median, $x$ below median</td>
<td>73.0</td>
<td>68.9</td>
<td>5.84</td>
</tr>
<tr>
<td>8. % entr., $z$ above median, $x$ above median</td>
<td>62.2</td>
<td>63.8</td>
<td>-2.52</td>
</tr>
<tr>
<td>9. % entr., $z$ below median, $x$ above median</td>
<td>46.3</td>
<td>49.0</td>
<td>-5.51</td>
</tr>
<tr>
<td>10. % entr., $z$ above median, $x$ below median</td>
<td>83.0</td>
<td>82.5</td>
<td>0.64</td>
</tr>
<tr>
<td>11. average gross income – all</td>
<td>390.8</td>
<td>395.3</td>
<td>-1.15</td>
</tr>
<tr>
<td>12. avg. gross income – entr., $z$ above median</td>
<td>638.3</td>
<td>651.0</td>
<td>-1.94</td>
</tr>
<tr>
<td>13. avg. gross income – entr., $x$ above median</td>
<td>669.7</td>
<td>680.6</td>
<td>-1.60</td>
</tr>
<tr>
<td>14. avg. gross income – entr., $z$ and $x$ below med.</td>
<td>294.2</td>
<td>335.5</td>
<td>-12.3</td>
</tr>
<tr>
<td>15. avg. gross income – entr., $z$ and $x$ above med.</td>
<td>778.8</td>
<td>858.0</td>
<td>-9.22</td>
</tr>
</tbody>
</table>

Note: income levels are in thousands baht

intervals around the data. We see that, at the GMM parameters the model matches well the overall level and slope of the lowess lines from the data (both with respect to initial wealth and schooling). However, the model struggles to match the data at very low levels of wealth (it under-predicts entrepreneurship) and at very low or high levels of schooling (it over-predicts entrepreneurship).\textsuperscript{12}

Figure 2.2: Probability of entrepreneurship - model vs. data

\textsuperscript{12}The test of overidentifying restrictions is rejected with J-statistic being equal to 11.6. The magnitude of the J-statistic is driven in particular by the seventh moment (percentage of entrepreneurs in the subsample with $z$ and $x$ in top terciles). At the weighted GMM estimates (obtained in two stages), the percentage deviation between model and data in that moment is 11% compared to deviations of less than 1.5% for the remainder of the moments.


Table 2.11: Relaxing the labor or credit constraints

<table>
<thead>
<tr>
<th></th>
<th>baseline free occ. choice ((\eta = 0))</th>
<th>relaxed credit (2(\lambda))</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Occupational choice</td>
<td></td>
<td></td>
</tr>
<tr>
<td>entrepreneurs</td>
<td>65.2%</td>
<td>54.4%</td>
</tr>
<tr>
<td>of which voluntary</td>
<td>83.4%</td>
<td>100%</td>
</tr>
<tr>
<td>of which involuntary</td>
<td>16.6%</td>
<td>0%</td>
</tr>
<tr>
<td>B. Net income</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean, all</td>
<td>378.0</td>
<td>+1.8%</td>
</tr>
<tr>
<td>10th percentile</td>
<td>188.0</td>
<td>+6.1%</td>
</tr>
<tr>
<td>30th percentile</td>
<td>281.6</td>
<td>+2.6%</td>
</tr>
<tr>
<td>median</td>
<td>353.6</td>
<td>+1.8%</td>
</tr>
<tr>
<td>70th percentile</td>
<td>433.5</td>
<td>+1.3%</td>
</tr>
<tr>
<td>90th percentile</td>
<td>595.1</td>
<td>+1.1%</td>
</tr>
<tr>
<td>mean, entrepreneurs</td>
<td>475.5</td>
<td>+16%</td>
</tr>
<tr>
<td>mean, voluntary entr.</td>
<td>553.9</td>
<td>no change</td>
</tr>
<tr>
<td>mean, involuntary entr.</td>
<td>82.6</td>
<td>n.a.</td>
</tr>
<tr>
<td>mean, non-business</td>
<td>195.3</td>
<td>-6.1%</td>
</tr>
</tbody>
</table>

2.4.3 Policy counterfactuals

Relaxing the labor or credit constraints

In the model, involuntary entrepreneurship arises if both of the following conditions are true: (i) the household does not have access to the alternative occupation (for example, a wage job), which we can interpret as a labor market constraint or friction and (ii) household income is maximized in the alternative occupation. The labor constraint is important for condition (i), while the credit constraint affects (ii). We compute the model at the GMM estimates to evaluate and isolate the effects of the two constraints on entrepreneurship – overall and its decomposition into voluntary and involuntary. We also report the effects of the two constraints on household income.

In the first counterfactual, we set the parameter \(\eta\) to zero while keeping all other parameters at their GMM estimates. This means that involuntary entrepreneurship is completely eliminated – all households have free occupational choice, for example, as in EJ (1989). This also affects average income in the economy since previously involuntary entrepreneurs are now able to choose the non-business occupation as income maximizing. Voluntary entrepreneurs are not affected by the elimination of the labor constraint as no new such entrepreneurs appear. The results are reported in Table 2.11 and are computed from the model-simulated data at the GMM estimates.\(^{13}\) In panel A we see that the elimination of the labor constraint reduces the rate of entrepreneurship to 55.3%. In panel B we compute the expected net income (mean, median, and percentiles) for each household defined as \(E(q^E - rk + (r - 1)z)\) for entrepreneurs, which equals output revenue minus the cost of cap-

\(^{13}\)In contrast, the statistics reported in Table 2.11 use the analytical expressions for mean income.
ital plus interest income, and the expectation is taken over the talent shock $\varepsilon$. Similarly we define net income as $q^A + (r - 1)z$ for non-entrepreneurs. Net income, as opposed to gross, is what households compare in making their occupational choice. We then average the net incomes across households. Table 2.11 shows that average net income is the highest for voluntary entrepreneurs and the lowest for involuntary entrepreneurs. This is intuitive, since by assumption, involuntary entrepreneurs would be more productive in the non-business occupation.

We see that eliminating the labor constraint increases income throughout the income distribution (see the second column in Panel B). The change from the baseline income indicates the change in the income distribution at the mean, median and different percentiles. As such it includes the effects of any mobility within the income distribution as a result of the counterfactual. For example, an ex-ante involuntary entrepreneur who is now free to work in the alternative occupation could move from the 10th to the 30th percentile, etc. We see that relaxing the labor constraint has strongest effect at the 10th income percentile (+6.1%), which are most likely to be involuntary entrepreneurs ex-ante. We also see a large positive effect on the mean entrepreneurial income (16% increase) accompanied with a fall in average income of non-business agents (-6.1%). The latter effect should not be confused with a negative impact on non-business income. Clearly, no one can lose from the relaxation of the labor constraint since everyone’s income weakly increases as one can either stay in their current occupation or switch to their preferred one. Instead, the reason for the fall in income of non-business agents is a pure composition effect – some unproductive entrepreneurs (with low talent $\theta$ and low schooling $x$) exit the business occupation and enter the alternative occupation. Finally, relaxing the labor constraint also affects the number of constrained entrepreneurs (those with $k = \lambda z$). The simulated data show that the percent of constrained entrepreneurs increases from 51.3% in the baseline to 56.8% (not reported in the table). The reason is that, with free occupational choice, all entrepreneurs are voluntary and have higher talent $\theta$ on average.

The second counterfactual we study is relaxing the credit constraint which we achieve by doubling the estimate of $\lambda$ from the baseline (from 0.23 to 0.46), keeping all other parameters at their GMM estimates. We see in the model-simulated data that relaxing the credit constraint has a minor effects on involuntary entrepreneurship (its share falls from 16.6% to 16.2%) and on entrepreneurship overall (it increases from 65.2% to 65.7%). This can be interpreted as suggesting that the labor market constraint rather than the credit constraint is more important in causing involuntary entrepreneurship. We do see, however, that relaxing the credit constraint has a significant impact on households’ income. The average increase in net income is more than double the increase from relaxing the labor constraint, with the impact being stronger across the whole income distribution. As in the labor constraint counterfactual, the 10th percentile experiences the largest income change.

\[14\] Of course, this is only true abstracting from general equilibrium effects that we do not model.
(+9.2%) as those agents become able to invest closer to their desired capital level. The voluntary entrepreneurs gain about the same as the average agent, while the involuntary entrepreneurs and non-entrepreneurs obtain only minor income gains, the former since they are mostly constrained by talent, the latter due to the small composition shift in the economy towards entrepreneurship.

Looking at the number of credit constrained households, those investing \( k = \lambda z \), in the simulated data (not reported in the table), unsurprisingly we see a large drop from 51.3% to 33.7% in the fraction of constrained entrepreneurs. Among voluntary entrepreneurs, the fraction of credit constrained falls from 56.8% to 37.5% while the corresponding impact among involuntary entrepreneurs is a decrease in the fraction of constrained from 23.4% to 14%.

We also illustrate the distribution of income gains from each of the two counterfactuals on Figure 2.3, stratified by log initial wealth, \( z \) and years of schooling, \( x \). Note that in our setting income gains can be interpreted as welfare gains. We use the simulated data from the model to compute the change in expected income (integrated over the talent shock \( \varepsilon \)) of each households with characteristics \((z_i, x_i)\) from the data, before vs. after the policy. The figure shows that relaxing the labor constraint (setting \( \eta = 0 \)) leads to much larger income gains for some low wealth individuals (up to 40%). These gains are monotonically decreasing on average in initial wealth and (except for very low \( x \) values) in the years of schooling, as it is less likely that a household would have been an involuntary entrepreneur for high \( z \) and \( x \). In contrast, the income gains from relaxing the credit constraint (doubling the estimated value of \( \lambda \)) are non-monotonic over initial wealth, with the households with intermediate wealth levels gaining the most. The reason is that they are most likely to be credit constrained entrepreneurs. The income gains from relaxing the credit constraint decline in schooling \( x \) on average, since households with larger values of \( x \) are more likely to have higher ability \( \theta \) and hence less likely to have been constrained.

**Microfinance**

We next consider the counterfactual of offering households the option to borrow and invest up to an additional \( M \) dollars. For example, this could be thought of in the context of a microfinance program. The requirement is that the loan can be only used to buy/rent business capital at the current interest rate \( r \). We analyze the effect of this policy counterfactual on the rate of involuntary entrepreneurship and household income. All model parameters are held at the baseline GMM estimates. We set the maximum microfinance loan size to 10% of the median gross income in our sample, \( M = 20,000 \) baht.

Households optimally choose \( k \) to solve

\[
\max_k \theta k^\alpha - rk \quad \text{subject to} \quad k \leq \lambda z + M
\]
and optimally choose to run a business if their income from entrepreneurship, using the re-optimized $k$, is higher than the alternative (unaffected by the policy). Clearly, all previously unconstrained households are unaffected by this policy while all constrained ones would have an incentive to participate.

Table 2.12 shows the effects of the policy on occupational choice and income, by various groups of households. We see that the fraction of entrepreneurs goes up by about 1 percentage point. In addition, within that increased number of businesses, the policy induces more voluntary entrepreneurship (+0.8%) while the rate of involuntary entrepreneurs falls from 16.6% to 15.8%. In terms of the policy effect on the distribution of income, the microfinance loan raises average income by 3% but the gains are unevenly spread among the different households. The poorest 10-th percentile benefits the most (a raise by 14% post vs. pre-policy) from the availability of additional credit while the top 10-th percentile benefits only marginally – these are households that are more likely to be unconstrained ex-ante.
Table 2.12: Policy counterfactual - microfinance

<table>
<thead>
<tr>
<th>Occupational choice</th>
<th>baseline</th>
<th>microfinance loan</th>
</tr>
</thead>
<tbody>
<tr>
<td>entrepreneurs</td>
<td>65.2%</td>
<td>66.3%</td>
</tr>
<tr>
<td>of which voluntary</td>
<td>83.4%</td>
<td>84.2%</td>
</tr>
<tr>
<td>of which involuntary</td>
<td>16.6%</td>
<td>15.8%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Net income</th>
<th>baseline</th>
<th>change from baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean, all</td>
<td>378.0</td>
<td>+3.3%</td>
</tr>
<tr>
<td>10th percentile</td>
<td>188.0</td>
<td>+14.2%</td>
</tr>
<tr>
<td>30th percentile</td>
<td>281.6</td>
<td>+5.3%</td>
</tr>
<tr>
<td>median</td>
<td>353.6</td>
<td>+2.5%</td>
</tr>
<tr>
<td>70th percentile</td>
<td>433.5</td>
<td>+1.7%</td>
</tr>
<tr>
<td>90th percentile</td>
<td>595.1</td>
<td>+1.0%</td>
</tr>
<tr>
<td>mean, entrepreneurs</td>
<td>475.5</td>
<td>+3.0%</td>
</tr>
<tr>
<td>mean, voluntary entr.</td>
<td>553.9</td>
<td>+2.2%</td>
</tr>
<tr>
<td>mean, involuntary entr.</td>
<td>82.6</td>
<td>+1.3%</td>
</tr>
<tr>
<td>mean, non-business</td>
<td>195.3</td>
<td>+0.6%</td>
</tr>
</tbody>
</table>

The mean income of entrepreneurs goes up (by 3%) for two reasons – first, the additional credit relaxes the financial constraint and allows entrepreneurs to earn more and second, there is a compositional shift from involuntary to voluntary entrepreneurs. The mean non-business income also goes up slightly as some agents with low schooling exit the occupation.

The effects of the microfinance loan policy on income are illustrated on Figure 2.4, stratified by log initial wealth, \( z \) and years of schooling, \( x \). We use the simulated data from the model and compute the change in expected income (integrated over the shock \( \varepsilon \)) of each households with characteristics \((z_i, x_i)\) from the data before and after the policy. We see that the policy benefits poorer households significantly (income gains of up to 60 percent relative to the baseline). The gains quickly dissipate for wealthier households since they are less likely to have been constrained and hence benefit from the access to microfinance. In contrast, the income gains are more spread out by years of schooling. This is due to the interaction of wealth and schooling in the data. The bottom panel of the figure shows that the households who gain the most from the policy are those with the lowest wealth and schooling. Low-wealth agents with high schooling do not gain much, as they are likely to be engaged in the alternative occupation for most values of \( \varepsilon \). Only the involuntary entrepreneurs among them stand to gain from the microfinance policy.

2.4.4 Robustness checks

We look at how sensitive the results are to the definition of business ownership, and how the labor market attribute \( x \) is measured. Column 2 in Table 2.13 reports estimates when we use whether a household derives the majority of their gross income from business to define business ownership. With this definition, the rate of business ownership in the sam-
## Table 2.13: Robustness checks

<table>
<thead>
<tr>
<th>Parameter/ Statistic</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>return to capital, $\alpha$</td>
<td>0.23</td>
<td>0.41</td>
<td>0.40</td>
<td>0.30</td>
<td>0.11</td>
<td>0.23</td>
<td>0.33</td>
</tr>
<tr>
<td>credit constraint, $\lambda$</td>
<td>0.23</td>
<td>0.43</td>
<td>0.28</td>
<td>0.17</td>
<td>0.06</td>
<td>1.32</td>
<td>0.20</td>
</tr>
<tr>
<td>return to schooling, $\gamma$</td>
<td>0.75</td>
<td>0.72</td>
<td>0.13</td>
<td>0.95</td>
<td>0.48</td>
<td>0.99</td>
<td>0.60</td>
</tr>
<tr>
<td>choice constraint, $\eta$</td>
<td>0.41</td>
<td>0.15</td>
<td>0.18</td>
<td>0.25</td>
<td>0.56</td>
<td>0.27</td>
<td>0.66</td>
</tr>
<tr>
<td>non-business income, $\mu$</td>
<td>28.5</td>
<td>38.7</td>
<td>120</td>
<td>18.5</td>
<td>49.9</td>
<td>14.3</td>
<td>46.8</td>
</tr>
<tr>
<td>talent – constant, $\delta_0$</td>
<td>3.42</td>
<td>3.36</td>
<td>3.66</td>
<td>3.27</td>
<td>4.14</td>
<td>2.09</td>
<td>4.11</td>
</tr>
<tr>
<td>talent – wealth, $\delta_1$</td>
<td>0.13</td>
<td>-0.03</td>
<td>-0.06</td>
<td>0.11</td>
<td>0.13</td>
<td>0.21</td>
<td>-0.02</td>
</tr>
<tr>
<td>talent – schooling, $\delta_2$</td>
<td>0.17</td>
<td>0.25</td>
<td>0.31</td>
<td>0.18</td>
<td>0.12</td>
<td>0.39</td>
<td>0.12</td>
</tr>
<tr>
<td>talent – std. deviation, $\sigma$</td>
<td>0.96</td>
<td>0.73</td>
<td>0.60</td>
<td>1.03</td>
<td>1.02</td>
<td>1.02</td>
<td>0.81</td>
</tr>
<tr>
<td>entrepreneurs, % of all</td>
<td>65.2</td>
<td>50.3</td>
<td>65.2</td>
<td>60.6</td>
<td>69.7</td>
<td>50.0</td>
<td>75.3</td>
</tr>
<tr>
<td>involuntary, % of all</td>
<td>10.8</td>
<td>5.7</td>
<td>6.9</td>
<td>7.1</td>
<td>13.8</td>
<td>8.0</td>
<td>18.2</td>
</tr>
<tr>
<td>involuntary, % of all entrep.</td>
<td>16.6</td>
<td>11.4</td>
<td>10.5</td>
<td>11.7</td>
<td>19.8</td>
<td>16.0</td>
<td>24.2</td>
</tr>
<tr>
<td>GMM criterion value</td>
<td>0.006</td>
<td>0.003</td>
<td>0.002</td>
<td>0.010</td>
<td>0.010</td>
<td>0.065</td>
<td>0.011</td>
</tr>
</tbody>
</table>

Notes:

1. baseline estimates
2. alternative definition of business ownership based on major source of income
3. alternative measure of schooling - head of household’s years of schooling
4. subsample = male principle earners
5. subsample = female principle earners
6. subsample = principle earners with below median age.
7. subsample = principle earners with above median (equal or higher) age
ple is 50%, compared to 65.2% of households that report owning a business. Compared to the 10.8% predicted for the baseline, the predicted probability of involuntary entrepreneurship with this new definition for business ownership is 5.7% for the sample (11.4% among entrepreneurs). A possible interpretation is that the proportion of income drawn from business is a proxy for entrepreneurial talent, such that households that rely more on business ownership have more talent, and therefore, are less likely to be involuntary entrepreneurs.

Column 3 uses years of schooling of the head of the household to measure $x$, instead of using the principle earner’s years of schooling. The probability of involuntary entrepreneurship is estimated to be 6.9% on the whole, and 10.5% for entrepreneurs. A possible interpretation could be that because heads of household have lower schooling on average compared to the principal earners, the implied alternative income is generally lower. As a result, we
estimate a lower probability of involuntary entrepreneurs when we use the head’s education as a proxy for $x$.

There could also be various other determinants of wage income. In the model, any variable that systematically affects occupational choice would be absorbed in the unobserved heterogeneity parameter $\theta$. We attempt to remedy this in part by estimating the model on various subsamples - by gender and age - to see whether the probability of involuntary entrepreneurship differs by these characteristics. Table 2.13 indicates that the estimates are roughly similar for male and female earners, however, principle earners above the median age are predicted to be are more likely to be involuntary entrepreneurs (18%), compared to their younger counterparts (8%).

2.5 Conclusions

The classical theory of occupational choice is predicated on the observed choice being better than the next best alternative. In this paper, we empirically explore the idea that some of the observed occupational choices are involuntary, especially in the context of developing countries. We model and estimate structurally the possibility that agents do not have access to wage employment, allowing the standard model of free occupational choice as a nested case. We define involuntary entrepreneurs as those entrepreneurs who maximize their income in a non-entrepreneurial job, but who are not able to access that job due to frictions in the labor market. Our structural estimation results classify about 17% of all business owners in the Thai urban data as involuntary entrepreneurs.

We use the structural estimates to evaluate the effect of relaxing the credit and labor constraints, as well as the impact of a microfinance policy on the rate of entrepreneurship (voluntary and involuntary) and household income, on average and stratified by wealth and schooling. Our results suggest that there are large potential income gains, especially for poorer households, from relaxing either the labor market or credit constraints or from providing access to microcredit, however the fraction of involuntary entrepreneurs can only be significantly reduced by addressing the labor market constraint.
Chapter 3

A Structural Econometric Analysis of Multiple Occupations - Evidence from Urban Thai Data

3.1 Introduction

There are at least three economic arguments for why the holding of multiple occupations is an indication of incomplete markets. First, imperfect insurance in consumption could lead to individuals diversifying income sources. Second, imperfect credit markets could lead to individuals not being able to “raise the capital they would need to run a business that would occupy them fully” (Banerjee and Duflo, 2007, p162). Third, and somewhat related is the idea that apart from credit constraints, underdeveloped labor markets could lead to individuals taking up a low productivity side activity to occupy what would otherwise be idle time.

I explore the latter two arguments. The primary goal of this paper is to understand why we observe multiple occupations, and more specifically, to quantify the importance of the financial constraint channel as opposed to lack of entrepreneurial skill channel in explaining the phenomenon. This is necessary to characterize whether multiple occupations is a sign of economic inefficiency (and therefore whether we should aim to discourage it), and whether and what kind of policies are likely to increase household welfare in this context. For example, if multiple occupations is a result mainly of lack of access to credit, policies that improve this market will enhance welfare by allowing households to maximize income. However, if it is mainly a result of underdeveloped labor markets in conjunction with being particularly ill-suited for entrepreneurship (leading to a lack of specialization in either occupation), encouraging further specialization in entrepreneurship will be ineffective in raising incomes, or in encouraging specialization in the first place. Therefore, the magnitude of
each type has implications for what kind of policies might be the most effective in increasing incomes.

I develop and structurally estimate a model in which risk neutral agents optimally allocate capital and labor based on their entrepreneurial ability and employability elsewhere to maximize their total income, subject to a collateral constraint and a time constraint. In much of the previous literature that applies classical occupational choice models to study entrepreneurship in developing countries, agents are able to choose only one occupation, and consequently do not distinguish between those who specialize solely in business ownership and those that take on another occupation in addition.

I relax the assumption that occupational choice is a binary and a mutually exclusive choice. Each agent in the model has access to two technologies. The entrepreneurial technology allows the agent to combine capital and labor, and income is assumed to be zero if either of the inputs are zero.\(^1\) The second technology only uses labor as an input. I call this technology wage-work for simplicity, although it could also represent subsistence occupations where capital plays no role. Agents maximize expected total income from the two technologies by choosing capital, entrepreneurial labor, and wage-work labor. Following the existing occupational literature in development, I also allow a collateral constraint for capital investment. In a nutshell, the model in this paper is an extension of Evans and Jovanovic (1989) – similar to their model, agents are heterogeneous in their entrepreneurial talent, initial wealth and wage, are expected income maximizers, and can only afford a positive multiple of their collateralizable wealth as capital investment. However, in addition to this basic framework, I allow households to allocate a fixed time endowment between entrepreneurship and wage-work.

The model assumes a Cobb-Douglas production for entrepreneurship with labor and capital as variable inputs that are gross complements. In addition, it assumes that the marginal products of capital and labor with respect to entrepreneurial output are higher (everywhere) for agents with higher entrepreneurial talent. The model predicts multiple occupations for individuals whose equilibrium entrepreneurial labor allocation does not exhaust the time constraint. This can arise both for individuals who are credit-constrained, and for those who are not credit-constrained. In the former case, capital investment in business is lower than it would be in the first-best scenario. Due to the complementarity in the two inputs, entrepreneurial labor is also suboptimal and does not exhaust the time-

---

\(^1\)I use a Cobb-Douglas production technology, that has decreasing returns to scale in capital and labor. An important implication is that optimal entrepreneurial labor and capital is never zero, unless a constraint on capital makes it zero. As a result, the model predicts that everyone is an entrepreneur. I focus only on business owners when I take the model to the data. Assuming a fixed cost of operating a business would produce exclusive wage-workers. However, estimation of such a model on the entire sample of business owners and wage-workers produces extremely poor model fit (for example, at the estimates, the model predicts similar probabilities of multiple occupation holding across all wealth levels, which is not true in the data). I leave the incorporation of exclusive wage-workers through more suitable extensions of the model for future work.
constraint. I call these multiple occupation holders “credit-constrained”. The second type of multiple job holders are those who are predicted to be unable to occupy their time endowment fully with entrepreneurship, even though they are not credit-constrained at their level of entrepreneurial talent. In that sense, I call them “skill-constrained”.

I estimate the model on data from a 2005 survey of urban households in Thailand (Townsend Thai Project). About 61% of households report owning at least one business. Among business owning households, I identify the individual business owner, and find that 20.4% of them also report a secondary occupation. On average, business owners with multiple occupations in the data are richer (in terms of household asset five years prior to the survey) and have higher years of schooling, compared to business owners that only have the one occupation. However, multiple occupation holders have lower gross business income on average compared to those with single occupation.

I structurally estimate the model via GMM to match the proportion of multiple occupation holders (overall and among different subsamples), and the average business incomes of multiple and single occupation holders respectively. I find that 85% of the business owners with multiple occupations (or 18% of the sample) are skill-constrained. That is, they are predicted to hold two jobs because given their level of entrepreneurial talent, the optimal entrepreneurial labor does not fully occupy them even though they are predicted to invest the first-best level of capital. The remaining 15% of multiple occupation holders (or 3.2% of the sample) are credit-constrained. The model predicts that on average, individuals with multiple jobs are less entrepreneurially skilled, and as observed in the data, have higher starting wealth and years of schooling, compared to those with a single occupation. Within the multiple occupations group, the skill-constrained are predicted to be less entrepreneurially talented, considerably richer, and have higher years of schooling compared to those who are predicted to be credit-constrained.

Consistent with the relative importance of lack of entrepreneurial skill as a cause for multiple occupations, a key implication of the model (at the GMM estimates) is that improving access to credit alone might not result in a significant decrease in the holding of multiple occupations. In particular, I estimate that eliminating the credit-constraint enables 2.1% of the sample to specialize in business ownership, and decreases the proportion of multiple occupation holders from 21% to 18.9%. It does however lead to an increase in the incomes of the poorest households (that are most likely to be credit-constrained). While the model implies that credit market inefficiency is not an important cause of multiple occupations among business owners in rural Thailand, the existence of skill-constrained multiple occupations could still be a symptom of underdevelopment, for example, of the wage-labor market.

More detail on the Townsend Thai Project can be found at http://cier.uchicago.edu/.
Related literature

One of the first papers to theoretically explore individuals holding more than one job is Shishko and Rostker (1976). In their model, multiple job holding occurs due to a time constraint on the main job, as long as the second job pays a wage above the marginal rate of substitution of income for leisure at the intersection of the primary wage line and the allowable hours on the first job. The paper estimates elasticities of moonlighting on US data (Income Dynamics Panel, 1968-70), and concludes that it decreases with wages and hours worked in the primary job, and increases with the wage on the second job. Similar to Shishko and Rostker (1976), the cause of moonlighting in Krishnan (1990) is the assumption that hours of work are fixed (institutionally) in the primary job, and examines married men’s decision to take a second job using the Survey of Income and Program Participation. The paper finds that those who held a second job (5% of married men) worked fewer hours on the first job, longer hours in total, were younger and had larger families on average. In my model, hours-constraint due to fixed employment contracts is not an explicit cause of multiple occupations. However, constraints in the market for wage employment, including the inability to find full-time jobs, could lead to multiple occupations for individuals who are not talented enough as entrepreneurs to exhaust their time endowment in business ownership. In that sense, skill-constrained multiple occupation holders (as defined in my model) could answer that they would like to work more in their non-business jobs. Similarly, the credit-constrained business owners (as defined in my model), who would like to employ both more labor and capital in their business could report that they would like to work more hours in their business. However given the credit constraint, they find it optimal to allocate their hours in a second job.

More recently, Choe et al. (2015) have modeled multiple occupations as a result of both the “hours-constraint” and a preference for job differentiation by using a Stone-Geary utility function. They motivate this by noting that about 60% of dual-job holding episodes among male workers in the British Household Panel Survey are self-reportedly unconstrained by hours on the main job. The preference for job differentiation could be a result of preference for heterogeneous jobs (Conway and Kimmel, 2001), hedging against unemployment (Bell et al., 1997) or to transition to a new primary job (Panos et al., 2011). These reasons could encapsulate both skill-constrained multiple occupation holders (as I define it in this paper) and risk-diversification motives.

The practice of multiple occupations is especially prevalent in developing countries. In their review of poor urban households, Banerjee and Duflo (2007) report that 47% in Cote d’Ivoire and Indonesia, 36% in Pakistan, 24% in Mexico and 20% in Peru derive their income from more than one source. They argue that individuals might take up a second job to occupy what would otherwise be wasted time, and that credit constraint and lack of

---

3“Moonlighting” is a term used frequently in the literature to refer to dual or multiple job holding.
skill could be the two main reasons for having left-over time. I model occupational choice and multiple occupations exactly along these lines. In particular, unlike much of the prior work, the model in this paper explicitly generates the holding of multiple occupations due to lack of entrepreneurial skill (or a wage-job that employs them full time). Banerjee and Duflo (2007) also notes that while risk diversification is a possible reason, most households in their surveys tend to employ themselves in relatively safe jobs. Similar to the urban survey used in this paper, the data in their paper come from urban settings.

Nonetheless, risk diversification is possibly an important reason for occupational diversification, and has been the focus of much of the empirical work behind occupational diversification in developing countries – especially for regions and households that deal with considerable income risk. Bandyopadhyay and Skoufias (2015) study how households in flood-prone areas in rural Bangladesh cope with risk, and find that households in areas of high rainfall variability have more occupational diversity. Shenoy (2015) uses the rural surveys of the Townsend Thai data (I use the urban surveys from the same project), and estimates that Thai rice farmers expecting a harvest in the next three months take on one extra activity when the volatility in international rice prices rises by 21%. In addition, the paper finds no evidence that occupational diversity decreased following the introduction of the Million Baht Program in 2001, a credit injection that gave one million baht for public lending in every village. The latter result is consistent with the finding in my paper that credit-constraint is not a major cause of multiple occupations in the Townsend data. In a related finding, Bianchi and Bobba (2012) report that current occupational choices are more responsive to future cash transfers compared to current cash transfers in Mexico, and speculate that the program created an increase in business ownership by enhancing the willingness to bear risk as opposed to simply relaxing current liquidity constraints. The model in this paper assumes risk neutral agents, and therefore does not account for risk diversification as a motive behind holding multiple occupations. It however does provide a framework that can be extended to include this motive, for example, by assuming risk aversion or by allowing variance in total income to enter the objective function negatively.

The paper is organized as follows. Section 3.2 presents the model, defining each type of multiple occupation holding, followed by descriptive summaries of multiple occupations in the Townsend Thai Data’s urban survey in Section 3.3. I then proceed to structurally

---

4 Higher rainfall variability is also associated with lower consumption; the use of credit (household took a loan in the past 12 months) or safety nets (a household member availed of safety-net programs in the last 12 months) are found not to completely mitigate the negative effects of rainfall variability on consumption, whereas access to markets does seem to offset them.

5 A behavioral model is used in understanding the motives behind multiple occupations in Hlouskova et al. (2015). The paper considers how loss aversion, the value of the reference level of income, and the expected return to risk affect the decision to hold multiple jobs. Their model predicts that a worker will not seek a risky job if she has an income reference level equal to what she can earn from a safe job. At any other reference level, the worker seeks new ventures provided she is compensated with a higher expected wage and is sufficiently loss averse.
estimate the model, present the main results and analysis in Section 3.4, and finally conclude in Section 3.5.

3.2 Model

I start with the assumption that agents are endowed with initial wealth $z$, years of schooling $x$ and entrepreneurial talent $\theta$. They can derive total income from an entrepreneurial production function and a non-entrepreneurial wage function.

$$y(z, x, \theta) = y^E(z, x, \theta) + y^W(x).$$

Additionally, total income from entrepreneurship $y^E(z, \theta)$ depends on talent and initial wealth in the following manner:

$$y^E(z, \theta) = \theta k^\alpha h^\beta + r(z - k), \quad \alpha > 0, \quad \beta > 0, \quad \alpha + \beta < 1,$$

where $\alpha$ and $\beta$ are elasticities of entrepreneurial income with respect to capital and labor respectively, and $r$ is the gross rate of interest at which agents can borrow and lend. The restriction that $\alpha + \beta < 1$ implies that the production function has decreasing returns to scale. A credit constraint is modeled in the form of a ceiling for capital investment:

$$k \leq \lambda z, \quad \lambda > 0.$$

This form of credit constraint can be micro-founded from a market failure of limited enforcement leading to collateralized loans. Additionally, labor allocation $h$ is subject to a time constraint:

$$h \leq T,$$

where $T$ represents the maximum time available for work. I assume that $T$ is not endogenous, and that it is fixed for all agents. This ignores the possibility of a market for hiring paid workers (more talented entrepreneurs might want to endogenously hire additional workers). It also ignores the fact that some business households employ unpaid family workers.\footnote{The majority of businesses observed in the data do not hire paid employees, which could reflect low productivity or a labor market failure (for example, moral hazard concerns).}

I also abstract away from heterogeneity in preference for leisure.\footnote{The model can be extended, at the cost of additional parameters, to allow the time endowment $T$ to depend on variables that could potentially affect preference for leisure, such as number of children and household size.}

The second source of income $y^W(x)$ is assumed to only depend on years of schooling, and is assumed to be linear in the time allocated to the activity, $(T - h)$:

$$y^W(x) = w(T - h),$$
with a Mincerian wage function, \( w = \mu(1 + x)^\gamma \). The parameter \( \mu \) represents the wage for individuals with zero years of schooling, and \( \gamma \) represents the elasticity of wage income with respect to schooling.

The agent faces the following optimization problem.

\[
\max_{k,h} w(T - h) + \theta k^\alpha h^\beta + r(z - k) \quad \text{s.t.} \quad h \in [0, T], \ k \in [0, \lambda z]
\] (3.1)

There are four threshold values of entrepreneurial talent. Suppose a household is able to optimally set \( k \) and \( h \) at their first-best levels, \( k_u \) and \( h_u \). Ignoring the constraints for now, the first-order conditions imply that the first-best levels of entrepreneurial capital and labor are

\[
k_u = \left(\frac{\theta \alpha}{r}\right)^{1-\beta}(\frac{\theta \beta}{w})^{1-\alpha-\beta}, \quad h_u = \left(\frac{\theta \alpha}{r}\right)^{1-\alpha}(\frac{\theta \beta}{w})^{1-\alpha-\beta}
\] (3.2)

The model predicts that entrepreneurial talent \( \theta \) is positively associated with the level of inputs in the first-best case, while wage (or equivalently \( x \)) is negatively related to the two. Therefore, individuals with higher \( \theta \), for a given level of wealth and schooling, are more likely to be credit-constrained and time-constrained. Individuals with higher \( w \), for a given level of wealth and talent, are less likely to be credit-constrained and time-constrained.

Define \( A(x) \) as the minimum level of \( \theta \) at which an individual with \( x \) years of schooling has \( h_u > 1 \).

\[
i.e. \quad h_u > T \iff \theta > T^{1-\alpha-\beta}(\frac{r}{\alpha}(\frac{w}{\beta})^{1-\alpha}) \equiv A(x)
\] (3.3)

Similarly, define \( B(z, x) \) as the minimum level of \( \theta \) at which an individual with characteristics \((z, x)\) finds \( k_u > \lambda z \).

\[
i.e. \quad k_u > \lambda z \iff \theta > (\lambda z)^{1-\alpha-\beta}(\frac{r}{\alpha}(\frac{w}{\beta})^{1-\alpha}) \equiv B(z, x)
\] (3.4)

Suppose \( k \) is set at \( \lambda z \). The optimal entrepreneurial labor choice can be found to be equal to

\[
\hat{h} = \left(\frac{\theta (\lambda z)^{\alpha} \beta}{w}\right)^{1-\beta}
\] (3.5)

and similarly, if \( h \) is set at \( T \), the optimal capital input can be found to be equal to

\[
\hat{k} = \left(\frac{\theta \alpha T^\beta}{r}\right)^{1-\alpha}
\] (3.6)

We again need to ensure that \( \hat{k} \) and \( \hat{h} \) do not violate the constraints. Define \( C(z) \) as the level of talent above which \( \hat{k} > \lambda z \).

\[
\hat{k} > \lambda z \iff \theta > (\lambda z)^{1-\alpha} \frac{r}{\alpha T^\beta} \equiv C(z)
\] (3.7)

Finally, define \( D(z, x) \) as the level of talent above which \( \hat{h} > 1 \).
Figure 3.1: Occupational map in \((z, \theta)\) plane for a given level of \(x\)

\[
\hat{h} > T \iff \theta > T^{1-\beta} \frac{w}{(\lambda z)^\alpha} \equiv D(z, x) \tag{3.8}
\]

Figure 3.1 plots the occupational map for different values of initial wealth \(z\) and talent \(\theta\), for a given value of schooling, \(x\), at plausible parameters.\(^8\) The solid line in the figure divides the plane into multiple occupation holders \((M = 1)\) and single occupation holders \((M = 0)\), equal to \(D(z, x)\) if the agent is credit-constrained, and equal to \(A(x)\) if the agent is not credit-constrained. The dashed line in the figure determines whether an agent is credit-constrained or not, and is equal to \(B(z, x)\) if the agent holds multiple occupations and \(C(z)\) if the agent specializes in business ownership. Consequently, the model predicts four types of business owners.

As a solution to the optimization problem defined in 3.1, the first-best levels of capital and entrepreneurial labor are allocated if both are simultaneously affordable. That is,

\[
(k^*, h^*) = (k_u(\theta), h_u(\theta)) \quad \text{if} \quad k_u < \lambda z, \; h_u < T \\
= (k_u(\theta), h_u(\theta)) \quad \text{if} \quad \theta < A(x), \; \theta < B(z, x) \nonumber.
\]

\(^8\)The parameters are set at the GMM estimates presented later in the paper.
The first type of multiple occupations therefore arises when an agent is of type \((\theta, z, x)\) such that \(\theta < \min\{A(x), B(z, x)\}\) - see the area labeled as \(M=1, \text{skill-constrained}\) in Figure 3.1. Given the level of initial wealth \(z\) and schooling \(x\), the level of entrepreneurial talent is sufficiently low to require capital and entrepreneurial labor allocations that do not exhaust the time constraint. As a result, they take a second job to maximize total income.\(^9\) In Figure 3.1, it can be seen that individuals with relatively high wealth but low talent are likely to be of this type.

A second type of multiple occupations is illustrated as \(M=1, \text{credit-constrained}\) in Figure 3.1. It arises when an agent is of type \((\theta, z, x)\) and the following conditions apply:

\[
(k^*, h^*) = \begin{cases} 
(\lambda z, \hat{h}(\theta)) & \text{if } k_u(\theta) > \lambda z, \hat{h}(\theta) < T \\
(\lambda z, \hat{h}(\theta)) & \text{if } B(z, x) < \theta < D(z, x)
\end{cases}
\]

In this case, given the level of initial wealth \(z\) and wage \(w\), the level of entrepreneurial talent is high enough for the credit constraint to bind (talent \(\theta\) is above the threshold \(B(z, x)\)). At the same time, the constrained-optimal level of equilibrium entrepreneurial labor does not exhaust the time constraint (talent \(\theta\) is below the threshold \(D(z, x)\)). Consequently, individuals take a second job to maximize total income; those with low wealth and relatively low talent are likely to be in this category. For a given level of talent, whether an increase in wealth leads to specialization in business depends on whether talent \(\theta\) is above or below the threshold \(A(x)\). That is, when \(\theta > A(x)\) (take \(\theta = 150\) in Figure 3.1 as an example), an increase in wealth could lead a multiple occupation holder \((M = 1)\) to specialize in business \((M = 0)\). In contrast, if \(\theta < A(x)\), any increase in wealth \(z\) does not lead to a switch to specialization. It only makes a credit-constrained multiple occupation holder into a skill-constrained multiple occupation holder.

Similarly, the model predicts two types of single-occupation business owners, depending on whether the business owner can invest the first-best level of capital or not. First, an agent of type \((\theta, z, x)\) such that \(A(w) < \theta < C(z)\), finds her first-best labor choice in business is higher than the time endowment \(T\) while unconstrained by credit (see the area labeled \(M=0, \text{time constrained}\) in Figure 3.1). Therefore,

\[
(k^*, h^*) = \begin{cases} 
(\hat{k}(\theta), T) & \text{if } \hat{k}(\theta) < \lambda z, h_u(\theta) > T \\
(\hat{k}(\theta), T) & \text{if } A(x) < \theta < C(z)
\end{cases}
\]

Finally, an agent with \((\theta, z, w)\) such that \(\theta > \max\{D(z, w), C(z)\}\), finds her first-best labor choice in business is higher than time endowment, in addition to being credit-constrained

---

\(^9\)Note that the “skill” here refers to entrepreneurial skill. Everything else equal, an individual could be skill-constrained because of a high wage \(w\) - higher \(w\) decreases the first-best entrepreneurial labor allocation \(h_u\) as specified in equation 3.2.
(labeled as $M=0$, credit-constrained in Figure 3.1). That is,

$$(k^*, h^*) = (\lambda z, T) \text{ if } \hat{k}(\theta) > \lambda z, \hat{h}(\theta) > T$$

$$(\lambda z, T) \text{ if } \theta > D(z, x), \theta > C(z).$$

A formal solution to the optimization problem defined in equation 3.1 is presented in Appendix C.1. For simplicity, assume for now that there is no correlation between talent $\theta$ and the observables $(z, x)$; I will relax this assumption when I estimate the model later. For a given level of schooling $x$, the model has the following implications:

(i). less entrepreneurially talented entrepreneurs are more likely to divide their time between their business and a second occupation

(ii). low-talent-high-asset entrepreneurs are likely to hold multiple occupations because they are skill-constrained

(iii). medium-talent-low-asset entrepreneurs are likely to hold multiple occupations because they are credit-constrained

(iv). high-talent-low-asset entrepreneurs are likely to specialize in business and be constrained by credit

(v). high-talent-high-asset entrepreneurs are likely to specialize in business and be constrained by the time endowment

**Predicted probability of multiple occupations**

In order to derive probabilities of multiple occupations predicted by the model, I assume a distribution for the unobserved heterogeneity in talent $\theta$. Specifically, I follow the literature on structural estimation of occupational choice models (Paulson et al., 2006; Karaivanov, 2012 for example), and assume that talent is possibly correlated with initial wealth $z$ and years of schooling $x$ and that the error term is log-normally distributed.

$$\ln \theta = \delta_0 + \delta_1 \ln z + \delta_2 \ln(1 + x) + \epsilon \quad (3.9)$$

$$\epsilon \sim N(0, \sigma)$$

For an agent with type $(z, x)$, the probability of holding multiple occupations is equal to the probability that the equilibrium entrepreneurial labor allocation $h^*$ is less than the time endowment $T$, which is, as discussed previously, the sum of the probability of being in skill-constrained multiple occupations and the probability of being in credit-constrained multiple occupations. The following proposition derives the probability of multiple occupations formally.

**Proposition 2**
Define $M$ as the indicator for holding two occupations. The model predicts that the probability of holding two occupations, conditional on characteristics $(z, x)$ and model parameters $\psi$ is

$$P(M = 1) = 1_{A > B} \Phi(b) + (1_{A < B}) \Phi(a) + 1_{B < D} \{ \Phi(d) - \Phi(b) \}$$

where $a \equiv (\ln A(x) - \bar{\theta})/\sigma$, $b \equiv (\ln B(z, x) - \bar{\theta})/\sigma$, $d \equiv (\ln D(z, x) - \bar{\theta})/\sigma$, and $\bar{\theta} = \delta_0 + \delta_1 \ln z + \delta_2 \ln(1 + x)$. $\Phi(\cdot)$ denotes the standard normal cumulative distribution function.

**Proof**

$$P(M = 1 | z, x, \psi) = P(h^* < T | z, x, \psi) = P(\theta < B(z, x), \theta < A(x)) + P(\theta > B(z, x), \theta < D(z, x))$$

$$= P(\theta < \min(B(z, x), A(x))) + P(B(z, x) < \theta < D(z, x))$$

$$= 1_{A > B} \Phi(b) + (1_{A < B}) \Phi(a) + 1_{B < D} \{ \Phi(d) - \Phi(b) \}$$

Note that the probability that talent $\theta$ is jointly below the thresholds $B(z, x)$ and $A(x)$ gives the probability of skill-constrained multiple occupations, and the probability that $\theta$ is between the thresholds $B(z, x)$ and $D(z, x)$ gives the probability of credit-constrained multiple occupations. Appendix C.1 derives model predictions for expected gross business income, conditional on multiple occupations and single occupation respectively; these will constitute additional moments that I use to structurally estimate the parameters of the model in Section 3.4.

### 3.3 Data

I use data from the Townsend Thai Project Initial Household Survey 2005 (Urban Area), a household survey of six provinces of urban Thailand.\footnote{The survey picked 16 communities in municipal areas of each amphoe or county under the ongoing Townsend's project (Rural Survey), totaling 96 communities overall. From each community, the survey randomly selected households who are present in the community-fund list (those that applied to Government Housing Bank or Bank of Agriculture and Agricultural Cooperatives). Each community fund must have no less than 95% of all households in their lists. More detail on the Townsend Thai Project can be found at http://cier.uchicago.edu/.

For simplicity, I focus on households that own a single business, and evaluate the occupational choice of the business owner. The indicator variable for multiple occupations, $M$, is equal to one if the individual has a secondary occupation. The proxy for $z$ is total asset of the household five years prior to the survey year, the proxy for $x$ is years of schooling of the business owner, and $q^E = \theta k^\alpha$ is measured by business income. I exclude households in the top percentile of wealth and business income. Key variables are summarized in Table 3.1. About 20% of business owners in the sample hold a secondary occupation, and are therefore defined as multiple occupation holders. They have significantly higher initial wealth $z$ and
Table 3.1: Summary statistics

<table>
<thead>
<tr>
<th></th>
<th>Multiple occ.</th>
<th>Single occ.</th>
</tr>
</thead>
<tbody>
<tr>
<td>wealth 5 years ago (000's)</td>
<td>723.0*</td>
<td>554.5*</td>
</tr>
<tr>
<td></td>
<td>(722.5)</td>
<td>(735.8)</td>
</tr>
<tr>
<td></td>
<td>[499.9]</td>
<td>[306.7]</td>
</tr>
<tr>
<td>business income (gross)</td>
<td>208.8*</td>
<td>317.1*</td>
</tr>
<tr>
<td></td>
<td>(298.9)</td>
<td>(437.7)</td>
</tr>
<tr>
<td></td>
<td>[108.0]</td>
<td>[156.6]</td>
</tr>
<tr>
<td>total earned income (gross)</td>
<td>320.9*</td>
<td>399.0*</td>
</tr>
<tr>
<td></td>
<td>(334.2)</td>
<td>(439.9)</td>
</tr>
<tr>
<td></td>
<td>[216.0]</td>
<td>[261.8]</td>
</tr>
<tr>
<td>years of schooling</td>
<td>8.0*</td>
<td>7.2*</td>
</tr>
<tr>
<td></td>
<td>(4.4)</td>
<td>(4.0)</td>
</tr>
<tr>
<td>household size</td>
<td>4.3</td>
<td>4.3</td>
</tr>
<tr>
<td></td>
<td>(1.6)</td>
<td>(2.0)</td>
</tr>
<tr>
<td>age</td>
<td>48.3</td>
<td>49.4</td>
</tr>
<tr>
<td></td>
<td>(11.0)</td>
<td>(11.0)</td>
</tr>
<tr>
<td>gender (male)</td>
<td>0.6*</td>
<td>0.4*</td>
</tr>
<tr>
<td></td>
<td>(0.5)</td>
<td>(0.5)</td>
</tr>
<tr>
<td>sample size</td>
<td>145</td>
<td>566</td>
</tr>
<tr>
<td>proportion</td>
<td>79.6%</td>
<td>20.4%</td>
</tr>
</tbody>
</table>

Mean, standard deviation (in parentheses) median (in brackets).

* indicates difference-in-means test is significant at 5% level.

higher years of schooling on average $x$, but have lower average business income (and average total gross income) than those with a single occupation.\(^{11}\)

Table 3.2 reports the coefficient estimates respectively from probit and linear probability regressions of $M$ (indicator for holding multiple occupations) on initial wealth $z$ and its square, years of schooling $x$ of the business owner and its square, the business owner’s gender and age, and household size. The estimates indicate that initial wealth increases the probability of holding multiple occupations (although at a decreasing rate). The effect of initial wealth $z$ on the probability of multiple occupations is theoretically ambiguous according to the model in Section 3.2. For a given level of talent $\theta$ and schooling $x$, the probability of credit-constrained multiple occupations is higher for relatively poorer households, whereas for the same level of talent $\theta$ and schooling $x$, the probability of skill-constrained multiple occupations is higher for relatively richer households (see Figure 3.1). With the observed positive correlation between multiple occupations and initial wealth, the model will likely map the multiple occupation holders into the skill-constrained type.

The regression results also indicate that years of schooling $x$ is associated with lower probability of multiple occupations for lower years of schooling, and higher probability for

\(^{11}\)Multiple occupation business owners are more likely to be male, but there is no statistically significant difference in average age of the principle earner and average household size between the two groups.
Table 3.2: Reduced-form regressions of multiple occupations

<table>
<thead>
<tr>
<th>Dependent variable: indicator variable for multiple occ. (M)</th>
<th>Probit</th>
<th>Linear prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>wealth 5 years ago (mln)</td>
<td>0.741***</td>
<td>0.207***</td>
</tr>
<tr>
<td></td>
<td>(0.199)</td>
<td>(0.053)</td>
</tr>
<tr>
<td>wealth squared (mln)</td>
<td>-0.190***</td>
<td>-0.052***</td>
</tr>
<tr>
<td></td>
<td>(0.066)</td>
<td>(0.016)</td>
</tr>
<tr>
<td>years of schooling</td>
<td>-0.118***</td>
<td>-0.033***</td>
</tr>
<tr>
<td></td>
<td>(0.058)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>schooling squared</td>
<td>0.007**</td>
<td>0.002***</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>gender (male)</td>
<td>0.392***</td>
<td>0.108***</td>
</tr>
<tr>
<td></td>
<td>(0.109)</td>
<td>(0.030)</td>
</tr>
<tr>
<td>age</td>
<td>-0.010*</td>
<td>-0.003*</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>household size</td>
<td>-0.007</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>intercept</td>
<td>-0.432</td>
<td>0.313***</td>
</tr>
<tr>
<td></td>
<td>(0.422)</td>
<td>(0.116)</td>
</tr>
<tr>
<td>sample size</td>
<td>711</td>
<td>711</td>
</tr>
</tbody>
</table>

*p < 0.10, **p < 0.05, ***p < 0.01

According to the model in Section 3.2, the effect of \( x \) on the probability of multiple occupations is positive for a given level of initial wealth \( z \) and talent \( \theta \); the opportunity cost of allocating an additional hour in business is higher for individuals with more years of schooling. However, this relationship could become ambiguous when we allow for a correlation between entrepreneurial talent and schooling.

Other variables that significantly affect the probability of holding multiple jobs include gender (being a male business owner is associated with a higher probability of multiple jobs) and age (older business owners are less likely to have multiple jobs). Household size does not significantly affect the probability of having multiple jobs.

Next, I look at the businesses and second occupations of the business owners in the sample, and summarize the key variables for each group. Table 3.3 shows that about 60% of the business owners are traders\(^{13}\), 32.5% are involved in small services (tailor, laundry, restaurant or noodle shop, repair shop, rental taxis etc), 3% are small producers (furniture makers etc., including fish farmers) and 4% raise livestock. Given that the majority of business owners are traders, the average business owner is most representative of this group.

\(^{12}\)For example, from the linear probability regression, the estimates indicate that the marginal effect of an additional year of schooling on the dependent variable is negative for years of schooling lower than 9.5, and positive otherwise. Values for years of schooling range from zero to seventeen.

\(^{13}\)Traders refer to those who declare themselves as one or are involved in buy-sale of goods under the category “other” as their occupation.
Table 3.3: Types of businesses among multiple occupation holders

<table>
<thead>
<tr>
<th>% of sample</th>
<th>trader</th>
<th>service</th>
<th>producer</th>
<th>livestock</th>
<th>overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>trader service producer livestock overall</td>
<td>59.9%</td>
<td>32.5%</td>
<td>3.4%</td>
<td>4.2%</td>
<td>100%</td>
</tr>
<tr>
<td>% with multiple occupation</td>
<td>18.8%</td>
<td>14.3%</td>
<td>41.7%</td>
<td>73.3%</td>
<td>20.4%</td>
</tr>
<tr>
<td>wealth 5 years ago (000s)</td>
<td>552.7</td>
<td>627.5</td>
<td>510.5</td>
<td>866.8</td>
<td>588.8</td>
</tr>
<tr>
<td>gross business income (000s)</td>
<td>366.1</td>
<td>197.2</td>
<td>339.9</td>
<td>122.1</td>
<td>300.0</td>
</tr>
<tr>
<td>years of schooling</td>
<td>7.1</td>
<td>7.6</td>
<td>9.6</td>
<td>7.3</td>
<td>7.4</td>
</tr>
<tr>
<td>gender (male indicator)</td>
<td>0.4</td>
<td>0.5</td>
<td>0.8</td>
<td>0.8</td>
<td>0.4</td>
</tr>
<tr>
<td>age</td>
<td>49.6</td>
<td>48.2</td>
<td>47.1</td>
<td>53.5</td>
<td>49.2</td>
</tr>
<tr>
<td>household size</td>
<td>4.4</td>
<td>4.0</td>
<td>4.2</td>
<td>4.4</td>
<td>4.3</td>
</tr>
</tbody>
</table>

Table 3.4: Types of non-business occupations among multiple occupation holders

<table>
<thead>
<tr>
<th>Second occupation</th>
<th>farmer</th>
<th>wage - profes.</th>
<th>wage - other</th>
<th>other</th>
<th>overall</th>
<th>M = 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of M = 1</td>
<td>55.9%</td>
<td>11.7%</td>
<td>18.6%</td>
<td>13.8%</td>
<td>100%</td>
<td>-</td>
</tr>
<tr>
<td>wealth 5 years ago (000s)</td>
<td>719.3</td>
<td>1263.8</td>
<td>397.2</td>
<td>717.9</td>
<td>723.0</td>
<td>544.5</td>
</tr>
<tr>
<td>business income (000s)</td>
<td>198.9</td>
<td>272.8</td>
<td>143.6</td>
<td>282.3</td>
<td>208.8</td>
<td>317.1</td>
</tr>
<tr>
<td>years of schooling</td>
<td>6.4</td>
<td>13.5</td>
<td>7.7</td>
<td>10.2</td>
<td>8.0</td>
<td>7.2</td>
</tr>
<tr>
<td>gender (male indicator)</td>
<td>0.6</td>
<td>0.6</td>
<td>0.5</td>
<td>0.4</td>
<td>48.3</td>
<td>49.4</td>
</tr>
<tr>
<td>age</td>
<td>50.6</td>
<td>50.4</td>
<td>43.9</td>
<td>43.3</td>
<td>0.6</td>
<td>0.4</td>
</tr>
<tr>
<td>household size</td>
<td>4.5</td>
<td>4.2</td>
<td>4.0</td>
<td>4.0</td>
<td>4.3</td>
<td>4.3</td>
</tr>
</tbody>
</table>
Those in services are the least likely to hold a second job, while small producers and livestock raisers are the most likely to have two occupations.

Table 3.4 reports the types of second occupations held by business owners in the sample. The majority of business owners with a second occupation are involved in farming (56%). Among the 30% who are wage-earners, those working in professional occupations (government employee, teacher etc.) are considerably richer and have more years of schooling on average compared to any other group, and particularly compared to non-professional wage-workers.

3.4 Structural Estimation

3.4.1 GMM - matched moments and computation

I have a sample of $N$ households, $i = 1, ..., N$ for whom initial wealth $z_i$, years of schooling $x_i$, and multiple occupation status $M_i$ are observed in the data. The parameters of the model include the entrepreneurial technology parameters $(\alpha, \beta)$ and the non-business income parameters $(\mu, \gamma)$, the credit constraint parameter $(\lambda)$, as well as the distributional parameters of talent $\theta$. I fix the interest rate parameter $r$ at 1.06, which corresponds to the median level of interest charged for loans in the data and the total time endowment $T$ is normalized to 1. Denote by $\phi$ the set of remaining nine parameters: $\phi = \{\alpha, \beta, \gamma, \mu, \lambda, \delta_0, \delta_1, \delta_2, \sigma\}$.

I estimate $\phi$ by minimizing the sum of squared percentage deviations between nine model predicted moments and their sample analogs, as listed in Table 3.5. The first moment is the proportion of multiple occupation overall, with the next six moments matching the proportion of multiple occupation holders for different subsamples (whether $z$ belongs to one of the three quartiles and whether $x$ belongs to one of the three quartiles, defined in Table 3.5). The remaining two moments are the expected gross income from entrepreneurship, conditional on $M = 1$, and conditional on $M = 0$. The model predictions for these moments are derived in propositions 2, 3 and 4.

Given parameters $\phi$, denote the model-predicted moments in the second column of Table 3.5 by $h_j(z, x, \phi)$, their sample analogs in the third column by $h^d_j$, and the percentage deviation between the two as $q_j(z, x, \phi) \equiv \frac{h_j(z, x, \phi) - h^d_j}{h^d_j}$, $j = 1, ..., J$. Finally, denote by $q(z, x, \phi)$ the $J \times 1$ vector of $q_j$'s. The GMM estimates are the solution to minimizing $q(z, x, \phi)'q(z, x, \phi)$ over $\phi$.

Even if the second job is seasonal, the model in Section ?? allows for that as long as engaging in it diverts time away from entrepreneurship. Holding seasonal jobs could therefore indicate that the individual is insufficiently skilled to run a business that occupies her full time.

The final group of multiple occupation holder, “other”, is made up of individuals engaged in piece-rate work. This includes 4 individuals involved in land rentals as their second occupation.

The proportion of multiple occupation holders for $z$ in the last quartile is a linear combination of the first four moments, and is therefore omitted. The proportion of multiple occupation holders for the fourth quartile of schooling is omitted for the same reason.
Table 3.5: List of matched moments

<table>
<thead>
<tr>
<th>Moment</th>
<th>Model</th>
<th>Sample analog</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. average probability of multiple occ</td>
<td>( \frac{1}{N} \sum_{i=1}^{N} P(M_i = 1</td>
<td>z_i, x_i, \phi) )</td>
</tr>
<tr>
<td>2. prob. of multiple occ, ( z \leq z_{25} )</td>
<td>( \frac{N}{\sum_{i=1}^{N} 1_{z_i \leq 25}} \sum_{1_{z_i \leq 25}} P(M_i = 1</td>
<td>z_i, x_i, \phi) )</td>
</tr>
<tr>
<td>3. prob. of multiple occ, ( z \in (z_{25}, z_{50}) )</td>
<td>( \frac{N}{\sum_{i=1}^{N} 1_{z_i \in (z_{25}, z_{50})}} \sum_{1_{z_i \in (z_{25}, z_{50})}} P(M_i = 1</td>
<td>z_i, x_i, \phi) )</td>
</tr>
<tr>
<td>4. prob. of multiple occ, ( z \in (z_{50}, z_{75}) )</td>
<td>( \frac{N}{\sum_{i=1}^{N} 1_{z_i \in (z_{50}, z_{75})}} \sum_{1_{z_i \in (z_{50}, z_{75})}} P(M_i = 1</td>
<td>z_i, x_i, \phi) )</td>
</tr>
<tr>
<td>6. prob. of multiple occ, ( x \leq x_{25} )</td>
<td>( \frac{N}{\sum_{i=1}^{N} 1_{x_i \leq 25}} \sum_{1_{x_i \leq 25}} P(M_i = 1</td>
<td>z_i, x_i, \phi) )</td>
</tr>
<tr>
<td>7. prob. of multiple occ, ( x \in (x_{25}, x_{50}) )</td>
<td>( \frac{N}{\sum_{i=1}^{N} 1_{x_i \in (x_{25}, x_{50})}} \sum_{1_{x_i \in (x_{25}, x_{50})}} P(M_i = 1</td>
<td>z_i, x_i, \phi) )</td>
</tr>
<tr>
<td>8. prob. of multiple occ, ( x \in (x_{50}, x_{75}) )</td>
<td>( \frac{N}{\sum_{i=1}^{N} 1_{x_i \in (x_{50}, x_{75})}} \sum_{1_{x_i \in (x_{50}, x_{75})}} P(M_i = 1</td>
<td>z_i, x_i, \phi) )</td>
</tr>
<tr>
<td>9. average entrep. income ( q^E, M = 1 )</td>
<td>( \frac{\sum_{i=1}^{N} E(q^E_i</td>
<td>M_i = 1, z_i, x_i, \phi) P(M_i = 1</td>
</tr>
<tr>
<td>10. average entrep. income ( q^E, M = 0 )</td>
<td>( \frac{\sum_{i=1}^{N} E(q^E_i</td>
<td>M_i = 0, z_i, x_i, \phi) P(M_i = 0</td>
</tr>
</tbody>
</table>

\( M \) is the indicator variable for multiple occupations.
\( z \) is initial wealth, \( z_j \) denotes the \( j^{th} \) percentile of \( z \).
\( x \) is years of schooling of business owners, \( x_j \) denotes the \( j^{th} \) percentile of \( x \).
Entrepreneurial income refers to the gross output, \( q^E = 9k^\alpha h^\beta \).
Recall that $M$ is measured as the indicator for whether the business owners report a secondary occupation ($M = 1$) or not ($M = 0$). Since I only look at households with a single business, business income $q^E$ is measured as business income reported by the household. Initial wealth $z$ is measured in the data as the total household asset held by the household five years prior to the survey in the form of land, household durables and agricultural assets. Finally, years of schooling $x$ is measured as the years of schooling of the business owner. The estimation sample excludes the top percentile of households in terms of initial wealth and business income to reduce the impact of outliers.

### 3.4.2 Estimates and model fit

The structural estimates are reported in Table 3.6. The estimates of the entrepreneurial income parameters, $\alpha$ and $\beta$, imply that a 10% increase in capital $k$ or a 10% increase in entrepreneurial labor $h$, would respectively lead to approximately 1% and 4% increase in the entrepreneurial income of unconstrained entrepreneurs, all else equal. The credit constraint parameter, $\lambda$ is estimated to be 0.17, implying that households can invest up to 17% of their initial wealth as capital. At the median wealth, this implies that capital $k$ should be less than 57 thousand baht.

The conditional mean and standard deviation of log entrepreneurial talent is estimated to be 5.6 and 0.2 respectively. Entrepreneurial talent $\theta$ is estimated to be negatively correlated with initial wealth $z$, with an elasticity of -0.08, and positively correlated with schooling $x$, with an elasticity of 0.1. The estimate of $\mu$ implies that for a business owner with zero years of schooling, income from an alternative non-business source would be about 77 thousand baht if all of the labor was allocated to the alternative occupation. The elasticity of this income with respect to schooling, $\gamma$, is estimated to be about 0.1.

Table 3.6 reports the value of the GMM criterion function, following the percentage deviations of each of the nine moments used to estimate the structural parameters. The model is able to reproduce the proportion of multiple occupation business owners closely, within 3% of the observed proportion. It is also able to match the proportion with multiple occupations for the first two quartiles of initial wealth within 5% of the observed analogs. The multiple occupation proportion for the subsample with wealth in the third quintiles are underestimated by 21%, with a 5.6 percentage points difference; this is the largest deviation among all targeted moments. The proportions of multiple occupation for the various subsamples based on quartiles of schooling are estimated within 10% (or within two percentage points difference) of their observed counterparts. The model is able to match

---

superscript

$^{17}$Evan and Jovanovic (1989) interprets the parameter $\delta_1$ by stating that “[it] may reflect greater past savings by those high-$\theta$ people, who, knowing their $\theta$, expected to become entrepreneurs one day. Or, if we stretch the interpretation a bit, it may reflect lower absolute risk aversion of wealthy people, making them more inclined to become entrepreneurs ...” (p816). They estimate $\delta_1$ to be negative and statistically significant for their sample of male workers in the United States, suggesting that high-asset people tend to be relatively poor entrepreneurs.
Table 3.6: Structural estimates and model fit

Structural estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>s.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>return to capital in $y^E$</td>
<td>$\alpha$</td>
<td>0.1083 (0.0139)</td>
</tr>
<tr>
<td>return to labor in $y^E$</td>
<td>$\beta$</td>
<td>0.3817 (0.0410)</td>
</tr>
<tr>
<td>credit constraint</td>
<td>$\lambda$</td>
<td>0.1701 (0.2527)</td>
</tr>
<tr>
<td>talent - intercept</td>
<td>$\delta_0$</td>
<td>5.6110 (0.5639)</td>
</tr>
<tr>
<td>talent - correlation with wealth</td>
<td>$\delta_1$</td>
<td>-0.0848 (0.0087)</td>
</tr>
<tr>
<td>talent - correlation with schooling</td>
<td>$\delta_2$</td>
<td>0.1261 (0.0128)</td>
</tr>
<tr>
<td>talent - standard deviation</td>
<td>$\sigma$</td>
<td>0.1861 (0.0444)</td>
</tr>
<tr>
<td>alternative income parameter</td>
<td>$\mu$</td>
<td>76.94 (9.81)</td>
</tr>
<tr>
<td>return to education in $y^W$</td>
<td>$\gamma$</td>
<td>0.1184 (0.0439)</td>
</tr>
</tbody>
</table>

Standard errors are calculated from 99 bootstrap samples.

Model fit

<table>
<thead>
<tr>
<th>Moment</th>
<th>Predicted</th>
<th>Observed</th>
<th>% Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>probability of multiple occ, $M = 1$</td>
<td>0.2106</td>
<td>0.2039</td>
<td>2.37</td>
</tr>
<tr>
<td>prob. of $M = 1$, $z \leq z_{25}$</td>
<td>0.1291</td>
<td>0.1348</td>
<td>-4.28</td>
</tr>
<tr>
<td>prob. of $M = 1$, $z \in (z_{25}, z_{50}]$</td>
<td>0.1285</td>
<td>0.1236</td>
<td>3.95</td>
</tr>
<tr>
<td>prob. of $M = 1$, $z \in (z_{50}, z_{75}]$</td>
<td>0.2147</td>
<td>0.2712</td>
<td>-20.8</td>
</tr>
<tr>
<td>prob. of $M = 1$, $x \leq x_{25}$</td>
<td>0.2062</td>
<td>0.1900</td>
<td>8.51</td>
</tr>
<tr>
<td>prob. of $M = 1$, $x \in (x_{25}, x_{50}]$</td>
<td>0.1847</td>
<td>0.1972</td>
<td>-6.32</td>
</tr>
<tr>
<td>prob. of $M = 1$, $x \in (x_{50}, x_{75}]$</td>
<td>0.2132</td>
<td>0.1946</td>
<td>9.56</td>
</tr>
<tr>
<td>average entrep. income $q^E$, $M = 1$</td>
<td>209.3</td>
<td>213.0</td>
<td>-1.72</td>
</tr>
<tr>
<td>average entrep. income $q^E$, $M = 0$</td>
<td>332.0</td>
<td>322.3</td>
<td>3.00</td>
</tr>
</tbody>
</table>

sum of squared % deviations (GMM criterion function) 0.0694

$M$ is the indicator variable for multiple occupations.

$z$ is initial wealth, $z_j$ denotes the $j^{th}$ percentile of $z$.

$x$ is years of schooling of business owners, $x_j$ denotes the $j^{th}$ percentile of $x$.

Entrepreneurial income refers to the gross output, $q^E = \theta h^\alpha k^\beta$. 

85
Figure 3.2: Model fit - unmatched moments

Local polynomial fit by initial wealth

Local polynomial fit by schooling
Figure 3.3: Predicted probability of multiple occupation as function of wealth

Probability of skill-constrained multiple occupations

Probability of credit-constrained multiple occupations
very closely the average business incomes of multiple occupation holders, and the average non-business incomes of single occupation holders, within 3% of the observed averages.

Figure 3.2 plots local polynomial fits of observed and predicted probabilities of multiple occupation at the GMM estimates by percentiles of initial wealth, and the entire range of years of schooling, both of which are not matched directly in estimating the parameters. In both graphs, the predicted fits are within the 95% confidence intervals of the observed fits, except in the case at the very bottom percentiles and the top decile of initial wealth. The figure also shows that evaluated at the GMM estimates, the model predicts that the probability of multiple occupations is decreasing at first and then increasing in initial wealth $z$. This is because, as Figure 3.3 shows, the probability of skill-constrained multiple occupations is increasing in initial wealth, however the probability of credit-constrained multiple occupations is decreasing in initial wealth.

For a given level of entrepreneurial talent $\theta$ (and schooling $x$), an agent with higher initial wealth is more likely to be skill-constrained if talent is below a certain threshold, as shown in Figure 3.1. In addition, part of the positive correlation between initial wealth and the predicted probability of skill-constrained multiple occupations can be attributed to the negative estimated correlation between $\theta$ and initial wealth $z$ (as the estimated value of $\delta_1$). The negative correlation between initial wealth and the predicted probability of credit-constrained multiple occupations is straightforward - poorer households are more likely to have binding credit constraints.

### 3.4.3 Analysis

By estimating the magnitude of each type of multiple occupation holding, the goal of the paper is to understand the causes of multiple occupations and analyze the implications for policy. The predicted proportions for various types of entrepreneurs are reported in Table 3.7. These statistics are calculated by simulating the model at the GMM estimates from Table 3.6, drawing at random 100 values for the shock $\varepsilon$ for each observation $i = 1, \ldots, N$.

The predicted proportion of business owners with multiple occupations is 21%, of which 85% (or 18% of the sample) are able to allocate the first-best levels of capital ($k_u$) and labor ($h_u$), as these are respectively lower than the investment limit ($\lambda z$) and the total time endowment. The remaining 15% (or 3% of the sample) holds a second job while being credit-constrained, and their constrained-optimum labor allocation does not exhaust the time endowment. In total, about 33% of business owners are predicted to be credit-constrained, however, only a few of them engage in multiple occupations.

Most of the sample, 79%, is predicted to specialize in business ownership, of which 38% (or 30% of the sample) specialize in business ownership and are credit-constrained. Relaxing the credit constraint therefore will increase their incomes. The remaining 62% (or 49% of the sample) specialize in business because the first-best level of their entrepreneurial labor $h_u$ is greater than the time endowment, and the corresponding constrained-optimal level
Table 3.7: Model predictions at GMM estimates

**Estimated types of entrepreneurs**

<table>
<thead>
<tr>
<th>Model statistic</th>
<th>Predicted value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Multiple occupation (M = 1)</strong></td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>21.0%</td>
</tr>
<tr>
<td>skill-constrained, $P(k_u &lt; \lambda z, h_u &lt; T)$</td>
<td>17.8%</td>
</tr>
<tr>
<td>credit-constrained, $P(k_u &gt; \lambda z, \hat{h} &lt; T)$</td>
<td>3.2%</td>
</tr>
<tr>
<td><strong>Single occupation (M = 0)</strong></td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>79.0%</td>
</tr>
<tr>
<td>time-constrained, $P(\hat{k} &lt; \lambda z, h_u &gt; T)$</td>
<td>48.8%</td>
</tr>
<tr>
<td>credit-constrained, $P(\hat{k} &gt; \lambda z, \hat{h} &gt; T)$</td>
<td>30.2%</td>
</tr>
</tbody>
</table>

**Characteristics of entrepreneurs by predicted type**

<table>
<thead>
<tr>
<th>average $\theta$</th>
<th>average $z$</th>
<th>average $x$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Multiple occupation (M = 1)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>166.2</td>
<td>941.6</td>
</tr>
<tr>
<td>skill-constrained</td>
<td>161.4</td>
<td>1099.3</td>
</tr>
<tr>
<td>credit-constrained</td>
<td>193.6</td>
<td>51.3</td>
</tr>
<tr>
<td><strong>Single occupation (M = 0)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>239.3</td>
<td>494.9</td>
</tr>
<tr>
<td>time-constrained</td>
<td>223.1</td>
<td>744.9</td>
</tr>
<tr>
<td>credit-constrained</td>
<td>265.4</td>
<td>92.5</td>
</tr>
</tbody>
</table>

of capital is lower than the investment limit $\lambda z$. If they could relax the time constraint to afford $h_u$ (for example by hiring workers), they could potentially find themselves credit-constrained as their desired level of capital would also increase. Otherwise, they would not benefit from a relaxation of the credit limit.

Figure 3.7 summarizes the characteristics of each type of entrepreneur. Those with multiple occupations are predicted to be less entrepreneurially talented on average, have higher average initial wealth, and higher average years of schooling than those with a single occupation (the latter two points are true in the observed data as well). However, within each group, there is a small fraction that is considerably different from the average. The credit-constrained multiple occupation holders are more talented on average, have much lower median income and lower years of schooling compared to the average multiple occupation entrepreneur. Among single occupation entrepreneurs, those that are credit-constrained are also more talented on average, poorer and have lower years of schooling on average. Note that the differences within multiple occupation holders (or single occupation holders) in
terms of average initial wealth and years of schooling, in addition to entrepreneurial talent, are unobserved in the data.

### Counterfactual analysis

I next analyze the effect of two counterfactual scenarios - relaxing the credit constraint and increasing entrepreneurial productivity - on the proportion of multiple occupation holders and income to further illustrate the implications of the estimated model. The second column in Table 3.8 reports the baseline statistics calculated at the GMM estimates, and the remaining columns report the corresponding statistics under the counterfactuals.

In the first counterfactual policy, I assume that the credit-constraint parameter $\lambda$ is high enough for everyone to be able to invest the first-best level of capital $k_u(\theta, r)$, such that the credit constraint is virtually eliminated for the given level of entrepreneurial talent and opportunity cost of funds $r$. An increase in $\lambda$ can be interpreted as lower collateral requirements due to improvements in contract enforcement. The third column Table 3.8 shows that when the first-best level of capital is invested, the probability of multiple occupations falls by only about 2 percentage points from 21% to 18.9%. This reflects the fact that relaxing the credit constraint does not affect the 17.8% of skill-constrained multiple occupation holders in the baseline. Some of the previously credit-constrained

---

18 I achieve this by multiplying the baseline $\lambda$ by a large number - in this case, by 1000.

19 Credit injection through microcredit loans for example would have the same effect of increasing equilibrium capital, with the difference being that access to credit is not tied to initial wealth.
multiple occupation holders now become skill-constrained, as the first-best labor allocation that complements with the first-best capital does not exhaust the time endowment. By definition, the probability of credit-constrained single occupation goes to zero from 30%.

The gain in income from eliminating the credit constraint is estimated to be about 3% on average. However, Figure 3.4 shows that there is considerable heterogeneity in how this policy affects individuals along the wealth distribution. Those in the first quintile of wealth increase their income by 16% on average. However, the effect of the remainder of the sample is predicted to be almost zero. The implication is that only individuals in the first quintile of initial wealth are predicted to be credit-constrained.

Figure 3.4 shows that the effect of eliminating the credit constraint is slightly lower for those with higher years of schooling. Intuitively, the opportunity cost of entrepreneurial labor is higher for individuals with higher years of schooling (through higher wages $w$). As a result, the complementary first-best capital $k_u$ is lower, and for a given level of wealth $z$, the probability that $k_u \leq \lambda z$ is also lower – implying that everything else equal, individuals with higher $x$ are less likely to be credit-constrained. Schooling also affects entrepreneurial talent $\theta$ positively through $\delta_2$, making individuals with higher years of schooling more likely to be credit-constrained. However, the direct effect of higher opportunity cost of entrepreneurial labor dominates on average.

The next counterfactual I consider is a sample-wide 10% increase in entrepreneurial talent $\theta$. This policy could be interpreted as a business training program that increases the marginal productivity of capital and labor inputs. It could also be interpreted as an exercise where the baseline is compared with a more productive economy. The results of a 10% increase in $\theta$ are reported in the last column of Table 3.8. The probability of multiple occupations falls from 21% to 10% - a higher $\theta$ increases equilibrium entrepreneurial

\[ \text{Total income is calculated as } \arg \max_{k,h} \theta k^a h^b - rk + w(T - h) + rz - z. \]

\[ \text{This is reflected in the fact that } B(z, x) \text{ is increasing in } x. \]
labor such that specializing in business is now profitable. In the baseline, 33.4% of the sample are credit-constrained among both single and multiple occupation holders. With the 10% increase in entrepreneurial productivity, the proportion that is predicted to be credit-constrained increases slightly to 36.2%.

The average gain in income is estimated to be about 10% for the whole sample. Figure 3.5 shows that the income gains from increasing $\theta$ is relatively homogeneous across wealth and schooling ranging from 5% to 11% (unlike in the case of relaxing the credit constraint where income gains ranged from 0% to 80%). Since everyone in the sample is an entrepreneur, a higher $\theta$ affects everyone’s income regardless of a switch in occupational choice. Within that, the gain in income for households with higher wealth are smaller as a percentage of their baseline income.

In a nutshell, the GMM estimates suggest that although a significant proportion of business owners are credit-constrained, relaxing the credit constraint would neither decrease the practice of multiple occupations by much, nor increase income in a significant manner for the average household. Completely eliminating the credit constraint is however predicted to increase the incomes of the poorest households by a significant amount. Given that most of the multiple occupation holders are predicted to be skill-constrained, increasing entrepreneurial productivity would be an effective policy to increase income.

### 3.5 Conclusions

The lack of specialization is a salient feature of occupational choice in developing countries. This paper examines the role of financial constraints and entrepreneurial skill in explaining the presence of multiple occupations, using survey data on household heads in semi-urban areas in Thailand. Structural estimation of this model is necessary to clarify why we
observe this phenomenon, which in turn is crucial when discussing the role for policy. For example, if it turns out that individuals do not specialize in their occupations due to credit constraints, policies that improve access to credit might enhance welfare by allowing households to maximize income. However, if it turns out that individuals do so in response to underdeveloped labor market (leading to irregular, seasonal or contractually limited paid employment), in conjunction with being particularly ill-suited for entrepreneurship, encouraging further specialization in their businesses might be ineffective in raising incomes, or equivalently, in reducing the holding of multiple jobs.

Using data from the Townsend Thai Project’s urban survey of 2005, I structurally estimate a model that predicts more than one occupation for individuals whose equilibrium entrepreneurial labor allocation does not exhaust the time constraint. I find that 85% of the business owners with multiple occupations (or 18% of the sample) are skill-constrained. They hold two jobs because given their level of entrepreneurial talent, the optimal entrepreneurial labor does not fully occupy them even though they invest the first-best level of capital. While this type of multiple occupation holding is not a sign of credit market inefficiency, it could still be indicative of underdeveloped labor markets that are unable to employ low-talent entrepreneurs in full-time occupations. The remaining 15% of multiple occupation holders (or 3.2% of the sample) are credit-constrained. A key implication of the estimated magnitudes of the two types of multiple occupation holders is that relaxing the credit constraint would not significantly decrease the practice of multiple occupations. For instance, I estimate that eliminating the credit-constraint only enables 2.1% of the sample to specialize in business ownership. These findings are also consistent with the idea that many entrepreneurs in developing countries are skill-constrained, and take up business ownership as a supplemental income source.

This paper can be extended in at least two important ways. First, the entrepreneurial production function can be altered to allow for the possibility of exclusive wage-workers. Second, the framework can also be extended to include risk aversion, utility from leisure, or disutility from variance in income to study diversification motive behind the holding of multiple jobs.

22 Preliminary estimation of a version with fixed costs in capital leads to a large decrease in model fit.
Conclusion

In each of the three chapters, I estimate extensions of a well-known occupational choice model (Evans and Jovanovic, 1989) to study entrepreneurship among Thai households.

While I use the same underlying structural model and the same country, the models in each chapter differ from each other in important ways. The model in the first chapter is closest to the baseline model, extending it to accommodate two periods with varying outside options of entrepreneurship. In the second chapter, the probability of entrepreneurship consists of an additional probability relative to the baseline due to a labor market constraint. In order to model multiple occupations among business owners in the third chapter, I include labor as an input in the entrepreneurial production function, and alter the interpretation of the outside option of entrepreneurship (i.e., the parameter $\mu$ no longer captures the minimum aggregate opportunity cost of entrepreneurship).

An additional source of variation in the estimates of the model parameters comes from the fact that I use different samples. The first chapter uses rural households of the Townsend data from 1997 and 1998, the second chapter uses urban survey from 2005, and the third chapter is estimated only on business owners from the 2005 urban survey.

Nevertheless, the primary benefit of estimating structural models is that structural parameters are expected to be stable across policy interventions. In particular, I discuss the comparability of capital’s share of income and return to schooling implied by the structural estimates in each chapter, and compare the estimates of the credit constraint parameter $\lambda$.

The elasticity of business income with respect to capital investment, $\alpha$, is estimated to be 0.44, 0.23 and 0.11 respectively in the first, second and the third chapter. For the first two chapters, $\alpha$ is capital’s share of income. For the third chapter, due to the presence of capital and labor in the production function, capital’s share is equal to the ratio of $\alpha$ and $\alpha + \beta$. Consequently, estimates for capital’s share of income is equal to 0.44 for the rural data set for the period 1997-1998, and equal to 0.23 and 0.22 for the urban data set for 2005. These are in the range of estimates found in the macroeconomic literature for capital’s share of income.

The parameter $\gamma$ is the elasticity of the outside option of entrepreneurship with respect to schooling in the first two chapters, and is estimated to be 0.17 and 0.75 respectively. Given the importance of farming as an alternative source of income in the rural data set,
the effect of schooling on the outside option of entrepreneurship is predictably lower than in the urban data. Consistent with the importance of wage income as the dominant non-business alternative (see Table 2.2), the estimate of 0.75 for $\gamma$ for the urban data set is also comparable with the estimates for return to schooling in the education literature. For example, an additional year of schooling from the sample average of eight years leads to an increase of around 8% in wage income. In the third chapter, $\gamma$ measures the elasticity of non-business income (the second occupation) of entrepreneurs only, and is estimated to be 0.12. It is plausible that return to schooling with respect to wage income is lower among business owners, and higher among exclusive wage-earners (usually the focus of papers in the literature pertaining to the return to schooling).

The parameter $\lambda$ measures the tightness of the credit constraint in all three models, representing the proportion of wealth that an entrepreneur can invest as capital. In the first chapter, I estimate that $\lambda$ is equal to 0.67, implying that households in the rural sample can invest up to 67% of their household wealth as capital. For the urban data, $\lambda$ (in the second and third chapters) are estimated relatively close to each other at 0.23 and 0.17. The implication however is that the credit market constraint is tighter for urban households in 2005 compared to rural households in 1997-1998. While we might expect access to credit market to be better in urban areas, the presence and the importance of informal sources in rural areas might help explain the present results.

Despite these differences, I find consistent evidence for the idea that entrepreneurship in developing countries arises out of imperfect labor market conditions. For example, I estimate that while endogenously starting a business during the Asian financial crisis mitigated about 40% of the income loss among rural households in my data, low entrepreneurial productivity limits the extent to which improving access to credit can further offset the income loss. In effect, the insurance value provided by entrepreneurship is limited for the households, and policies that encourage entrepreneurship are imperfect substitutes for unemployment insurance. Similarly, in the second chapter, we find that although there are large potential gains from relaxing the credit constraint, involuntary entrepreneurship, which in the model is indicative of occupational choice misallocation and ensuing inefficiency, can only significantly be reduced by addressing the labor market constraint. In the third chapter, I find that a significant proportion of business owners in the sample have two occupations because they are skill-constrained, and are unable to occupy themselves in full-time occupations. In particular, the results could be indicative of underdeveloped labor markets. Overall, results from all three chapters suggest that we need to look closely at the outside options of entrepreneurs, and that addressing labor market constraints is crucial for understanding the forces behind entrepreneurship and for affecting household well-being.

There are two important caveats. First, the models are static, and therefore are unable to fully capture dynamic decision-making by households and individuals. For example, a dynamic model might be able to capture non-linear relationships between wealth and en-
entrepreneurship. Buera (2009) examines the relationship between wealth and entrepreneurship in a dynamic model and predicts that the probability of entrepreneurship is increasing for low wealth levels and decreasing for high wealth levels. In addition, if the labor market constraints that restrict access to wage-employment originate from search frictions, such a friction alone might not plausibly be expected to be able to generate and explain long-term involuntary entrepreneurship (since search frictions are generally assumed to be temporary).

Second, the analyses are conducted in partial equilibrium. Some policy interventions might have different effects when general equilibrium effects are considered. For example, relaxing credit constraints or increasing entrepreneurial productivity might not just raise an entrepreneur's income, but also raise labor demand (as they might become employers), and therefore increase wages. The incidence of higher wages and its effect on entrepreneurial profit might have important distributional effects. Buera et al. (2014) studies the aggregate and distributional impact of microfinance in a dynamic model of occupational choice with financial frictions, and highlights the general equilibrium implications. Extending the analyses in dynamic models that take into account general equilibrium effects are avenues for future research.
Bibliography


Appendix A

Entrepreneurship During an Economic Crisis

A.1 Endogenizing the credit constraint

Imperfect enforcement of credit contracts due to voluntary default of borrower can lead to a constraint of the form $k \leq \lambda z$. Suppose when borrowers default, they are caught with a probability of $p$ and a fraction $\beta$ of their wealth $z$ is forfeited. If they do not default, they pay $r(z - k)$.

The incentive compatibility constraint is

\[ r(z - k) \leq p\beta z \]

\[ \implies z \geq \frac{r}{r + p\beta} k \]

\[ \implies k \leq (1 + \frac{\beta p}{r}) z. \]

The parameter $\lambda$, which represents the credit constraint parameter is equal to $1 + \frac{\beta p}{r}$ in this case.

A.2 Proofs

Model quantities are conditional on observed data $(z, x)$ and model parameters, denoted by $\phi$, which I suppress for notational ease. The definition of threshold values of talent $A, B$ and $C$ are given respectively in equations 3.6, 1.7 and 3.8. Let $\Phi(\cdot)$ denote the standard normal CDF. For notational simplicity, define

\[ a_t = \frac{\ln A_t - \bar{\theta}}{\sigma}, \quad b_t = \frac{\ln B_t - \bar{\theta}}{\sigma}, \quad c_t = \frac{\ln C_t - \bar{\theta}}{\sigma}. \quad (A.1) \]

Proposition 1
Define \( E \) as the indicator that a household is an entrepreneur. Conditional on initial wealth \( z \), labor market attribute \( x \) and model parameters \( \phi \), the probability of entrepreneurship is

\[
P(E = 1|z, x, \phi) = 1_{B > A}[1 - \Phi(a)] + 1_{B < A}[1 - \Phi(c_t)] \tag{A.2}
\]

**Proof**

\[
P(E = 1) = P(\pi > 0) = P(\pi > 0)\theta > B)P(\theta > B) + P(\pi > 0)\theta < B)P(\theta < B)
\]

\[
= P(\theta > C)\theta > B)P(\theta > B) + P(\theta > A)\theta < B)P(\theta < B)
\]

\[
= \frac{P(\theta > C)\theta > B)}{P(\theta > B)}P(\theta > B) + \frac{P(\theta > A)\theta < B)}{P(\theta < B)}P(\theta < B)
\]

\[
= 1_{B > C}P(\ln \theta > \ln B) + 1_{B < C}P(\ln \theta > C) + 1_{A < B}P(\ln A < \theta < \ln B)
\]

\[
= 1_{B > C}P(\frac{\xi}{\sigma} > \frac{\ln B - \bar{B}}{\sigma}) + 1_{B < C}P(\frac{\xi}{\sigma} > \frac{\ln C - \bar{B}}{\sigma}) + 1_{A < B}P(\ln A - \theta < \ln B - \bar{B})
\]

\[
= 1_{B > C}(1 - \Phi(b)) + 1_{B < C}(1 - \Phi(c)) + 1_{A < B}(\Phi(b) - \Phi(a))
\]

Note that \( B > A \Leftrightarrow B > C \). Therefore,

\[
P(E = 1) = 1_{B > A}(1 - \Phi(b) + \Phi(b) - \Phi(a)) + 1_{B < A}(1 - \Phi(c))
\]

\[
= 1_{B > A}[1 - \Phi(a)] + 1_{B < A}[1 - \Phi(c_t)]
\]

□

**Proposition 2** Define \( E_{01} \) as the indicator that a household is not in business in period \( t \) and is in business in period \( s \). For \( \mu_t \neq \mu_s \) and for a household with characteristic \((z_t, z_s, x)\), the probability of that \( E_{01} = 1 \) is given by

\[
P(E_{01} = 1|z_t, z_s, x, \phi) = \max\left[0, \Phi(c_t) - \Phi(\max(b_t, b_s, c_s))\right] + \max\left[0, \Phi(\min(a_t, b_t)) - \Phi(\max(b_s, c_t))\right]
\]

\[
+ \max\left[0, \Phi(\min(b_t, b_s)) - \Phi(\max(b_t, a_t))\right] + \max\left[0, \Phi(\min(b_t, b_s, a_t)) - \Phi(a_s)\right]
\]

**Proof**

\[
P(E_{01} = 1) = P(E_{01} = 1, \theta > B_t, \theta > B_s) + P(E_{01} = 1, \theta < B_t, \theta > B_s)
\]

\[
+ P(E_{01} = 1, \theta > B_t, \theta < B_s) + P(E_{01} = 1, \theta < B_t, \theta < B_s)
\]

The first term can be written as,

\[
P(E_{01} = 1, \theta > B_t, \theta > B_s) = P(\theta > B_t, \theta > B_s, \theta < C_t, \theta > C_s)
\]

\[
= 1_{[\max(B_t, B_s, C_s) < C_t]}P\left(\max(B_t, B_s, C_s) < \theta < C_t\right)
\]

Using similar steps for the remaining terms, we get

\[
P(E_{01} = 1) = 1_{[\max(B_t, B_s, C_s) < C_t]}P\left(\max(B_t, B_s, C_s) < \theta < C_t\right)
\]

\[
+ 1_{[\max(B_t, B_s, C_s) < \min(B_t, A_t)]}P\left(\max(B_s, C_s) < \theta < \min(B_t, A_t)\right)
\]

\[
+ 1_{[\max(B_t, A_s) < \min(B_t, C_t)]}P\left(\max(B_t, A_s) < \theta < \min(B_t, C_t)\right)
\]

\[
+ 1_{[A_s < \min(B_t, B_s, A_t)]}P\left(A_s < \theta < \min(B_t, B_s, A_t)\right)
\]

103
For any constant \(d\), note that 
\[
P(\theta < d) = P(\ln \theta < \ln d) = P\left(\frac{\xi}{\sigma} < \frac{\ln d - \bar{\theta}}{\sigma}\right) = \Phi\left(\frac{\ln d - \bar{\theta}}{\sigma}\right).
\]

Transforming the terms appropriately, we get
\[
P(E_{01} = 1) = \max\left[0, \Phi(c_t) - \Phi(\max(b_t, b_s, c_s))\right] + \max\left[0, \Phi(\min(a_t, b_t)) - \Phi(\max(b_s, c_s))\right] + \max\left[0, \Phi(\min(b_t, b_s)) - \Phi(\max(b_t, a_t))\right] - \Phi(a_s).
\]

\[\square\]

**Proposition 3** Define \(E_{11}\) as the indicator that a household is in business in period \(t\) and in business in period \(s\). For a household with characteristic \((z_t, z_s, x)\), the probability of this event is given by
\[
P(E_{11} = 1 | z_t, z_s, x, \phi) = \left[1 - \Phi(\max(b_t, b_s, c_t, c_s))\right] + \max\left[0, \Phi(b_t) - \Phi(\max(a_t, b_s, c_s))\right] + \max\left[0, \Phi(b_s) - \Phi(\max(b_t, c_t, a_s))\right] + \max\left[0, \Phi(\min(b_t, b_s)) - \Phi(\max(a_t, a_s))\right].
\]

\[(A.4)\]

**Proof**
\[
P(E_{11} = 1) = P(E_{11} = 1, \theta > B_t, \theta > B_s) + P(E_{11} = 1, \theta < B_t, \theta < B_s) + P(E_{11} = 1, \theta > B_t, \theta < B_s) + P(E_{11} = 1, \theta < B_t, \theta > B_s)
\]
\[
= P\left(\theta > \max(B_t, B_s, C_t, C_s)\right) + 1_{\max(B_s, A_t, C_t) < B_t} P\left(\max(B_s, A_t, C_t) < \theta < B_t\right)
\]
\[
+ 1_{\max(B_t, C_t, A_s) < B_s} P\left(\max(B_t, C_t, A_s) < \theta < B_s\right)
\]
\[
+ 1_{\max(A_t, A_s) < \min(B_t, B_s)} P\left(\max(A_t, A_s) < \theta < \min(B_t, B_s)\right)
\]
\[
= \left[1 - \Phi(\max(b_t, b_s, c_t, c_s))\right] + \max\left[0, \Phi(b_t) - \Phi(\max(a_t, b_s, c_s))\right] + \max\left[0, \Phi(b_s) - \Phi(\max(b_t, c_t, a_s))\right] + \max\left[0, \Phi(\min(b_t, b_s)) - \Phi(\max(a_t, a_s))\right].
\]

\[\square\]

**Proposition 4** Define \(E_{00}\) as the indicator that a household is not in business in period \(t\) and not in business in period \(s\). For \(\mu_t \neq \mu_s\) and for a household with characteristic \((z_t, z_s, x)\), the probability of this event is given by
\[
P(E_{00} = 1 | z_t, z_s, x) = \max\left[0, \Phi(\min(c_t, c_s)) - \Phi(\max(b_t, b_s))\right] + \max\left[0, \Phi(\min(a_t, b_t, c_s)) - \Phi(b_s)\right] + \max\left[0, \Phi(\min(c_t, a_s, b_s)) - \Phi(\min(c_t, c_s, b_s))\right] + \max\left[0, \Phi(\min(b_t, b_s, a_t, a_s))\right].
\]

\[(A.5)\]
Proof

\[ P(E_{00} = 1) = P(E_{00} = 1, \theta > B_t, \theta > B_s) + P(E_{00} = 1, \theta < B_t, \theta > B_s) + P(E_{00} = 1, \theta > B_t, \theta < B_s) + P(E_{00} = 1, \theta < B_t, \theta < B_s) \]
\[ = 1_{[\max(B_t, B_s) < \min(C_t, C_s)]} P\left( \max(B_t, B_s) < \theta < \min(C_t, C_s) \right) \]
\[ + 1_{[B_t < \min(A_t, B_t, B_s)]} P\left( B_t < \theta < \min(C_t, A_t, B_s) \right) + P\left( \theta < \min(B_t, B_s, A_t, A_s) \right) \]
\[ = \max\left[ 0, \Phi(\min(c_t, c_s)) - \Phi(\max(b_t, b_s)) \right] + \max\left[ 0, \Phi(\min(a_t, b_t, c_s)) - \Phi(b_t) \right] \]
\[ + \max\left[ 0, \Phi(\min(c_t, a_t, b_s)) - \Phi(\min(c_t, a_t, b_s)) \right] + \left[ \Phi(\min(b_t, b_t, a_t, a_s)) \right] \]

□

Proposition 5

Let \( q^E = \theta k^x \theta \) be entrepreneurial revenue, or gross income from entrepreneurship. The expectation of \( q^E \), conditional on \( E = 1 \) (and on \( z, x \) and \( \phi \) ) is given by

\[
E(q^E|E = 1, z, x, \phi) = 1_{A > B}(\lambda z)^{\alpha} \exp(\bar{\theta} + \frac{\sigma^2}{2} \Phi(\sigma - c) - \frac{\alpha}{2} \sigma^2) \\
+ 1_{A < B} \exp\left( \frac{\bar{\theta}}{1 - \alpha} + \frac{\sigma^2}{2(1 - \alpha)^2} \right)^{\Phi(\sigma - c) - \Phi(\sigma - b)} \\
+ 1_{A < B}(\lambda z)^{\alpha} \exp(\bar{\theta} + \frac{\sigma^2}{2} \Phi(\sigma - b) - \frac{\alpha}{2} \sigma^2) \tag{A.6}
\]

Proof

\[
E(q^E|E = 1) = E(q^E|\pi > 0) \\
= 1_{A > B} E(q^E|\pi > 0, A > B) + 1_{A < B} E(q^E|\pi > 0, A < B)
\]

The first term is,

\[
E(q^E|\pi > 0, A > B) = E(q^E|\pi > 0, \theta < B, A > B) P(\theta < B|\pi > 0, A > B) \\
+ E(q^E|\pi > 0, \theta > B, A > B) P(\theta > B|\pi > 0, A > B) \\
= E(q^E|\theta > C, \theta < B, A > B) P(\theta < B|\theta > C, A > B) \\
+ E(q^E|\theta > C, \theta > B, A > B) P(\theta > B|\theta > C, A > B) \\
= E(q^E|\theta > C) \\
= (\lambda z)^{\alpha} E(\theta|\theta > C) \\
= (\lambda z)^{\alpha} \Phi(\sigma - c) \\
= (\lambda z)^{\alpha} \exp(\bar{\theta} + \frac{\sigma^2}{2} \Phi(\sigma - c))
\]

where I used the fact that \( P(\theta < B|\theta > C, A > B) = 0 \), \( P(\theta > B|\theta > C, A > B) = 1 \), and since \( A > B \iff C > B \), the condition that \( (\theta > C \text{ and } \theta > B) \) can be simplified to \( \theta > C \).
Next,

$$\mathbb{E}(q^E|\pi > 0, A < B) = \mathbb{E}(q^E|\pi > 0, \theta < B, A < B)P(\theta < B|\pi > 0, A < B)$$

$$+ \mathbb{E}(q^E|\pi > 0, \theta > B, A < B)P(\theta > B|\pi > 0, A < B)$$

$$= \mathbb{E}(q^E|\theta > A, \theta < B, A < B)P(\theta > B|\theta > A, A < B)$$

$$+ \mathbb{E}(q^E|\theta > A, \theta > B, A < B)P(\theta > B|\theta > A, A < B)$$

$$= \mathbb{E}(q^E|\theta > A, \theta < B, A < B)\frac{P(A<\theta<B)}{P(\theta>B)}$$

$$+ \mathbb{E}(q^E|\theta > A, \theta > B, A < B)\frac{P(\theta>B)}{P(\theta>A)}$$

$$= \left(\frac{a}{\pi}\right)^{\alpha} \frac{\alpha}{1-\alpha} \exp\left(\frac{\theta}{1-\alpha} + \frac{\sigma^2}{2(1-\alpha)^2}\right) \frac{\Phi(\frac{\sigma}{\alpha}(\sigma-a)) - \Phi(\frac{\sigma}{\alpha}(\sigma-b))}{1-\Phi(\alpha)}$$

where I have used the fact that $P(\theta > A, \theta > B) = P(\theta > B)$ when $A < B$, and therefore,

$$\frac{P(A<\theta<B)}{P(\theta>B)} = \frac{\Phi(b)-\Phi(a)}{1-\Phi(\alpha)}$$

When $A < \theta < B$, $q^E = \theta k^a \vartheta$, and

$$\mathbb{E}(q^E|\theta > A, \theta < B, A < B) = \left(\frac{a}{\pi}\right)^{\alpha} \frac{\alpha}{1-\alpha} \mathbb{E}(\theta|\theta > A, A < B)$$

$$= \left(\frac{a}{\pi}\right)^{\alpha} \frac{\alpha}{1-\alpha} \mathbb{E}(\theta|\theta > A, A < B)$$

$$= \frac{\Phi(\sigma|\alpha(a)) - \Phi(\sigma|\alpha(b))}{\Phi(b)-\Phi(a)}$$

For $A < B$, $(\theta > A, \theta > B) \Leftrightarrow \theta > B$, and $q^E = \theta(\lambda z)^a \vartheta$, and

$$\mathbb{E}(q^E|\theta > A, \theta > B, A < B) = \left(\lambda z\right)^a \mathbb{E}(\theta|\theta > B)$$

$$= \left(\lambda z\right)^a \mathbb{E}(\theta|\theta > B)$$

$$= \frac{\Phi(\sigma|\alpha(a)) - \Phi(\sigma|\alpha(b))}{\Phi(b)-\Phi(a)}$$

$\square$

**Moments**

For $j \in \{1, 2, 3\}$

$$q_j(z, x, \phi) \equiv \frac{1}{N} \sum_{i=1}^{N} P(G_{\kappa,i} = 1|z_i, x_i, \phi) - \frac{1}{N} \sum_{i=1}^{N} G_{\kappa,i}^d,$$

where $G_{\kappa,i}^d$ denotes the observed counterpart of $G_\kappa$. For $j \in \{4, 5, 6\}$,

$$q_j(z, x, \phi) \equiv \frac{1}{N} \sum_{i=1}^{N} \left(1\{x_i < x_m\} P(G_{\kappa,i} = 1|z_i, x_i, \phi) - \frac{1}{N} \sum_{i=1}^{N} \left(1\{x_i < x_m\} \times G_{\kappa,i}^d\right)\right),$$

where $1\{x_i < x_m\}$ is an indicator function for having less than the median four years of schooling. The remaining four moments match the average gross business income in
each year, conditional on being a business owner, overall and for below-median schooling subsamples. For $j \in \{7, 8\}$,

$$h_j(z, x, \phi) = \frac{\sum_i \left\{ \mathbb{E}(q^E_{it} | E_{it} = 1, z_{it}, x_i, \phi) P(E_{it} = 1 | z_{it}, x_i, \phi) \right\}}{\sum_i P(E_{it} = 1 | z_{it}, x_i, \phi)}, \quad t \in \{97, 98\}$$

$$h^d_j = \frac{\sum_{i=1}^N (q^E_{it} \times E^d_{it})}{\sum_{i=1}^N E^d_{it}}, \quad t \in \{97, 98\}$$

and for $j \in \{9, 10\}$,

$$h_j(z, x, \phi) = \frac{\sum_i \left\{ \mathbb{1}\{x_i < x_m\} \mathbb{E}(q^E_{it} | E_{it} = 1, z_{it}, x_i, \phi) P(E_{it} = 1 | z_{it}, x_i, \phi) \right\}}{\sum_i \mathbb{1}\{x_i < x_m\} P(E_{it} = 1 | z_{it}, x_i, \phi)}, \quad t \in \{97, 98\}$$

$$h^d_j = \frac{\sum_{i=1}^N \mathbb{1}\{x_i < x_m\} \times q^E_{it} \times E^d_{it}}{\sum_{i=1}^N \mathbb{1}\{x_i < x_m\} \times E^d_{it}}, \quad t \in \{97, 98\}$$

where again, $q^E_{it}$ denotes the observed counterpart of $q^E_{it}$.

### A.3 Additional tables
## Table A.1: Business ownership over the years in Townsend Thai Data

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>% business owners</td>
<td>16.9%</td>
<td>36.6%</td>
<td>36.9%</td>
<td>42.4%</td>
<td>42.4%</td>
</tr>
<tr>
<td>% entering</td>
<td>23.0%</td>
<td>7.1%</td>
<td>10.3%</td>
<td>7.2%</td>
<td></td>
</tr>
<tr>
<td>% exiting</td>
<td>3.2%</td>
<td>6.8%</td>
<td>4.8%</td>
<td>7.2%</td>
<td></td>
</tr>
<tr>
<td>net change (p.p)</td>
<td>19.8%</td>
<td>0.2%</td>
<td>5.5%</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>% change w.r.t. previous year</td>
<td>117.6%</td>
<td>0.6%</td>
<td>20.5%</td>
<td>0%</td>
<td></td>
</tr>
</tbody>
</table>

Based on a balanced panel sample of 835 households, observed every survey year from 1997 to 2001 in the Townsend Thai Data.

## Table A.2: Sources of household income

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>rice farming</td>
<td>58.5 60.0</td>
<td>37.4 39.3</td>
<td>78.5 79.3</td>
</tr>
<tr>
<td>wage $^*,c,n$</td>
<td>66.7 60.1</td>
<td>74.4 70.9</td>
<td>60.4 50.8</td>
</tr>
<tr>
<td>business $^*,c,n$</td>
<td>16.8 36.7</td>
<td>21.2 45.3</td>
<td>12.8 28.5</td>
</tr>
<tr>
<td>other agriculture $^n$</td>
<td>41.4 43.9</td>
<td>50.0 50.0</td>
<td>33.2 38.1</td>
</tr>
</tbody>
</table>

$^*$ indicates that the difference-in-means test between years for the whole sample is significant at 10%.

$c$ indicates that the difference-in-means test between years for the central provinces is significant at 10%.

$n$ indicates that the difference-in-means test between years for the Northeast is significant at 10%.
Table A.3: Loan characteristics

<table>
<thead>
<tr>
<th>loan source</th>
<th>obs</th>
<th>interest rate (median)</th>
<th>loan size (median)</th>
<th>duration (median)</th>
<th>for business (median)</th>
<th>for consumption (median)</th>
</tr>
</thead>
<tbody>
<tr>
<td>relatives, neighbors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>250</td>
<td>10%</td>
<td>10.0</td>
<td>12</td>
<td>8.8%</td>
<td>50.0%</td>
</tr>
<tr>
<td>1998</td>
<td>464</td>
<td>0%</td>
<td>6.0</td>
<td>12</td>
<td>6.3%</td>
<td>70.0%</td>
</tr>
<tr>
<td>BAAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>414</td>
<td>10%</td>
<td>20.0</td>
<td>12</td>
<td>8.9%</td>
<td>28.5%</td>
</tr>
<tr>
<td>1998</td>
<td>384</td>
<td>10%</td>
<td>27.7</td>
<td>12</td>
<td>10.2%</td>
<td>63.0%</td>
</tr>
<tr>
<td>village funds</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>204</td>
<td>10%</td>
<td>15.0</td>
<td>12</td>
<td>5.9%</td>
<td>37.8%</td>
</tr>
<tr>
<td>1998</td>
<td>343</td>
<td>11%</td>
<td>15.0</td>
<td>12</td>
<td>4.4%</td>
<td>72.2%</td>
</tr>
<tr>
<td>moneylender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>204</td>
<td>22%</td>
<td>13.4</td>
<td>12</td>
<td>4.9%</td>
<td>45.6%</td>
</tr>
<tr>
<td>1998</td>
<td>343</td>
<td>22%</td>
<td>19.0</td>
<td>12</td>
<td>7.0%</td>
<td>58.9%</td>
</tr>
<tr>
<td>bank</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>34</td>
<td>7%</td>
<td>100.0</td>
<td>48</td>
<td>23.5%</td>
<td>38.2%</td>
</tr>
<tr>
<td>1998</td>
<td>29</td>
<td>8%</td>
<td>120.0</td>
<td>60</td>
<td>13.8%</td>
<td>62.1%</td>
</tr>
<tr>
<td>other</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>84</td>
<td>1%</td>
<td>20.0</td>
<td>24</td>
<td>9.5%</td>
<td>44.0%</td>
</tr>
<tr>
<td>1998</td>
<td>121</td>
<td>1%</td>
<td>10.0</td>
<td>24</td>
<td>7.4%</td>
<td>75.2%</td>
</tr>
<tr>
<td>total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>1121</td>
<td>10%</td>
<td>20.0</td>
<td>12</td>
<td>6.6%</td>
<td>31.3%</td>
</tr>
<tr>
<td>1998</td>
<td>1499</td>
<td>10%</td>
<td>15.0</td>
<td>12</td>
<td>6.5%</td>
<td>57.2%</td>
</tr>
</tbody>
</table>

Each loan observation, within or across households, is counted uniquely.

Units: annual interest rate, loan size in 000s of 1997 Thai baht, duration in months.

The last two columns report loan use for business investment or consumption.
Appendix B

Involuntary Entrepreneurship

B.1 Proofs

Proof of Proposition 1
Using the definitions of $y^E(\theta, z)$, $y^A(z, x)$, and $\Delta(z, \theta, x)$, we obtain,

$$\Delta(z, \theta, x) \geq 0 \iff \begin{cases} (1 - \alpha)\theta \frac{1}{\sqrt{\sigma}(z)} \frac{\alpha}{\sqrt{\sigma}} - \mu(1 + x)^\gamma \geq 0 & \text{if } \theta \leq B(z) \\ \theta(\lambda z)^\alpha - r\lambda z - \mu(1 + x)^\gamma \geq 0 & \text{if } \theta > B(z) \end{cases}$$

(B.1)

which, in terms of the agent’s entrepreneurial ability $\theta$, is equivalent to,

$$\Delta(z, \theta, x) \geq 0 \iff \begin{cases} \theta \geq \left(\frac{\mu - 1\alpha(1 + x)^\gamma}{1 - \alpha}\right) & \text{if } \theta \leq B(z) \\ \theta \geq (\lambda z)^{-\alpha}[\mu(1 + x)^\gamma + r\lambda z] & \text{if } \theta > B(z) \end{cases}$$

Proof of Lemma 1:
Using (2.5), we have, since $1_{B>A} = 1_{B>C}$,

$$\tilde{P}_E = P(\Delta \geq 0) = 1_{B>A}\left\{P\left(\frac{\varepsilon}{\sigma} \geq \frac{\ln B - \tilde{\theta}}{\sigma}\right) + P\left(\frac{\ln A - \tilde{\theta}}{\sigma} \leq \frac{\varepsilon}{\sigma} \leq \frac{\ln B - \tilde{\theta}}{\sigma}\right)\right\} +$$
$$+(1 - 1_{B>A})P\left(\frac{\varepsilon}{\sigma} \geq \frac{\ln C - \tilde{\theta}}{\sigma}\right)$$

Let $\Phi(\cdot)$ be the standard normal cumulative density function. We then obtain,

$$P(\Delta \geq 0) = 1_{B>A}\left\{1 - \Phi\left(\frac{\ln B - \tilde{\theta}}{\sigma}\right) + \Phi\left(\frac{\ln B - \tilde{\theta}}{\sigma}\right) - \Phi\left(\frac{\ln A - \tilde{\theta}}{\sigma}\right)\right\} +$$
$$+(1 - 1_{B>A})\left\{1 - \Phi\left(\frac{\ln C - \tilde{\theta}}{\sigma}\right)\right\} =$$

$$1_{B>A}\left\{1 - \Phi\left(\frac{\ln A - \tilde{\theta}}{\sigma}\right)\right\} + (1 - 1_{B>A})\left\{1 - \Phi\left(\frac{\ln C - \tilde{\theta}}{\sigma}\right)\right\}$$

which is equivalent to the Lemma statement. □
Derivation of the income moments

We derive the expected gross income conditional on business ownership. Recall that the expected entrepreneurial output is defined as $q^E(\theta, z) = \theta k^\alpha$. The expected output, conditional on being an entrepreneur (and conditional on observables $z$ and $x$, but we suppress these dependencies for notational ease) is,

$$E(q^E|1_E = 1) = \int q^E(\theta)f(\theta|1_E = 1)d\theta = \int q^E(\theta)\frac{f(\theta, 1_E = 1)}{P(1_E = 1)}d\theta =$$

$$= \int q^E(\theta)f(\theta, \Delta \geq 0) + P_x f(\theta, \Delta < 0)\frac{P(1_E = 1)}{P(1_E = 1)}d\theta =$$

$$= \frac{P(\Delta \geq 0)}{P(1_E = 1)} \int q^E(\theta)f(\theta, \Delta \geq 0)d\theta + \frac{P_x P(\Delta < 0)}{P(1_E = 1)} \int q^E(\theta)f(\theta, \Delta < 0)d\theta =$$

$$= \frac{P(\Delta \geq 0)}{P(1_E = 1)} E(q^E|\Delta \geq 0) + \frac{P_x P(\Delta < 0)}{P(1_E = 1)} E(q^E|\Delta < 0)$$

The probabilities $P(\Delta < 0)$ and $P(\Delta \geq 0)$ were computed in Lemma 1. We also have

$$E(q^E|\Delta < 0) = E(q^E|\Delta < 0, \theta > B)P(\theta > B|\Delta < 0) + E(\ln q^E|\Delta < 0, \theta \leq B)P(\theta \leq B|\Delta < 0)$$

$$= E(q^E|\Delta < 0, \theta > B)\frac{P(\theta > B, \Delta < 0)}{P(\Delta < 0)} + E(q^E|\Delta < 0, \theta \leq B)\frac{P(\theta \leq B, \Delta < 0)}{P(\Delta < 0)}$$

$$= E(q^E|\Delta < 0, \theta > B)\frac{P(\theta > B, \theta < C)}{P(\Delta < 0)} + E(q^E|\Delta < 0, \theta \leq B)\frac{P(\theta \leq B, \theta < A)}{P(\Delta < 0)}$$

$$= E(q^E|\Delta < 0, \theta > B)\frac{(1 - 1_{B>A})(\Phi(c) - \Phi(b))}{P(\Delta < 0)} + E(q^E|\Delta < 0, \theta \leq B)\frac{\Phi(\min(a, b))}{P(\Delta < 0)}$$

and

$$E(q^E|\Delta \geq 0) = E(q^E|\Delta \geq 0, \theta > B)P(\theta > B|\Delta \geq 0) + E(q^E|\Delta \geq 0, \theta \leq B)P(\theta \leq B|\Delta \geq 0)$$

$$= E(q^E|\Delta \geq 0, \theta > B)\frac{P(\theta > B, \Delta \geq 0)}{P(\Delta \geq 0)} + E(q^E|\Delta \geq 0, \theta \leq B)\frac{P(\theta \leq B, \Delta \geq 0)}{P(\Delta \geq 0)}$$

$$= E(q^E|\Delta \geq 0, \theta > B)\frac{P(\theta > B, \theta \geq C)}{P(\Delta \geq 0)} + E(q^E|\Delta \geq 0, \theta \leq B)\frac{P(\theta \leq B, \theta \geq A)}{P(\Delta \geq 0)}$$

$$= E(q^E|\Delta \geq 0, \theta > B)\frac{1 - \Phi(\max(b, c))}{P(\Delta \geq 0)} + E(q^E|\Delta \geq 0, \theta \leq B)\frac{1_{B>A}(\Phi(b) - \Phi(a))}{P(\Delta \geq 0)}$$

and where:

1. $E(q^E|\Delta < 0, \theta > B) = E\left\{ (\lambda z)^\alpha \theta | B < \theta < C \right\} = (\lambda z)^\alpha E(\theta|B < \theta < C)$

2. $E(q^E|\Delta < 0, \theta \leq B) = E\left\{ (\frac{\theta}{z})^{1-\alpha} \theta^{1-\alpha} | \theta \leq \min(A, B) \right\}$. 

111
\[
\alpha r \alpha_1 - \alpha E(\theta - \alpha_1) \Theta(\min(a, b) - \frac{\sigma}{(1 - \alpha)}) \], where \( E(\theta^{1 - \alpha}) = \exp\left(\hat{\theta} + \frac{\sigma^2}{2(1 - \alpha)^2}\right) \).

3. \( E(q^E|\triangle \geq 0, \theta > B) = E\left\{ (\lambda z)^\alpha \theta | \theta \geq \max(B, C) \right\} = (\lambda z)^\alpha E(\theta) \frac{\Phi(\max(b, c) - \sigma)}{\Phi(-\max(b, c))} \).

4. \( E(q^E|\triangle \geq 0, \theta \leq B) = E\left\{ \left(\frac{\alpha r}{\theta - \alpha(1 - \alpha)}\right) \theta^{1 - \alpha} \mid A \leq \theta \leq B \right\} = (\frac{\alpha r}{\theta - \alpha(1 - \alpha)}) E(\theta^{1 - \alpha}) \frac{\Phi(\frac{\sigma}{(1 - \alpha)} - a) - \Phi((\frac{\sigma}{(1 - \alpha)}) - b)}{\Phi(b) - \Phi(a)} \).
Appendix C

A Structural Analysis of Multiple Occupations

C.1 Proofs

Proposition 1

The solution to the optimization defined in 3.1 is

\[ h^* = \begin{cases} 
  h_u & \text{if } (k_u < \lambda z & h_u < T) \Leftrightarrow (\theta < B(z, w) & \theta < A(w)) \\
  \hat{h} & \text{if } (k_u > \lambda z & \hat{h} < T) \Leftrightarrow (\theta > B(z, w) & \theta < D(z, w)) \\
  T & \text{if } \begin{cases} 
    (\hat{k} < \lambda z & h_u > T) \Leftrightarrow (\theta < C(z) & \theta > A(w)) \\
    (\hat{k} > \lambda z & \hat{h} > T) \Leftrightarrow (\theta > D(z, w) & \theta > C(z)) 
  \end{cases}
\end{cases} \]

\[(C.1)\]

\[ k^* = \begin{cases} 
  k_u & \text{if } (k_u < \lambda z & h_u < T) \Leftrightarrow (\theta < B(z, w) & \theta < A(w)) \\
  \hat{k} & \text{if } (\hat{k} < \lambda z & h_u > T) \Leftrightarrow (\theta < C(z) & \theta > A(w)) \\
  \lambda z & \text{if } \begin{cases} 
    (k_u > \lambda z & \hat{h} < T) \Leftrightarrow (\theta > B(z, w) & \theta < D(z, w)) \\
    (\hat{k} > \lambda z & \hat{h} > T) \Leftrightarrow (\theta > D(z, w) & \theta > C(z)) 
  \end{cases}
\end{cases} \]

\[(C.2)\]

Proof

The Lagrangian for the problem is

\[ \max_{h_b, k, \lambda} L(h_b, k) = w(T - h) + \theta k^\alpha h^\beta + r(z - k) + \lambda_1 k + \lambda_2 h + \lambda_3 [\lambda z - k] + \lambda_4 [T - h] \]

The Kuhn-Tucker conditions:

\[ k \geq 0, \lambda_1 \geq 0, \lambda_1 k = 0 \]
\[ h \geq 0, \lambda_2 \geq 0, \lambda_2 h = 0 \]
\[ \lambda z - k \geq 0, \lambda_3 \geq 0, \lambda_3 [\lambda z - k] = 0 \]
\[T - h \geq 0, \lambda_4 \geq 0, \lambda_4[T - h] = 0\]
\[
d\frac{L}{dh} = -w + \theta k^\alpha h^{\beta - 1} + \lambda_2 - \lambda_4 = 0
\]
\[
d\frac{L}{dk} = \theta k^{\alpha - 1}h^\beta - r + \lambda_1 - \lambda_3 = 0
\]

Case 1:

Suppose \(k\) and \(h\) are interior solutions. This means that \(\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0\) as none of the constraints bind. At the optimum, marginal product is equal to marginal cost of each input.

\[
\theta \alpha k^{\alpha - 1}h^\beta = r
\]
\[
\theta k^\alpha h^{\beta - 1} = w
\]

Let \(y = k^\alpha h^\beta\).

\[
\implies \theta \alpha \frac{y}{k} = r, \quad \theta \beta \frac{y}{h} = w
\]
\[
\implies k^* = \frac{\theta \alpha}{r} y, \quad h^* = \frac{\theta \beta}{w} y
\]

At the optimum,

\[
y^* = \left(\frac{\theta \alpha}{r} y^*\right)^\alpha \left(\frac{\theta \beta}{w} y^*\right)^\beta
\]
\[
\implies y^{*1-\alpha-\beta} = \left(\frac{\theta \alpha}{r}\right)^\alpha \left(\frac{\theta \beta}{w}\right)^\beta
\]
\[
\implies y^* = \left(\frac{\theta \alpha}{r}\right)^{1-\alpha-\beta} \left(\frac{\theta \beta}{w}\right)^{\beta/(1-\alpha-\beta)}
\]

Capital investment is

\[
k^* = \frac{\theta \alpha}{r} \left(\frac{\theta \alpha}{r}\right)^{\frac{\alpha}{1-\alpha-\beta}} \left(\frac{\theta \beta}{w}\right)^{\frac{\beta}{1-\alpha-\beta}} = \left(\frac{\theta \alpha}{r}\right)^{\frac{1-\alpha-\beta}{1-\alpha-\beta}} \left(\frac{\theta \beta}{w}\right)^{\frac{\beta}{1-\alpha-\beta}}
\]
\[
\implies k^* = \frac{\theta \alpha}{r} \left(\frac{1-\alpha-\beta}{1-\alpha-\beta}\right) \left(\frac{\theta \beta}{w}\right)^{\frac{\beta}{1-\alpha-\beta}} = k_u
\]

Entrepreneurial labor is

\[
h^* = \frac{\theta \beta}{w} \left(\frac{\theta \alpha}{r}\right)^{\frac{\alpha}{1-\alpha-\beta}} \left(\frac{\theta \beta}{w}\right)^{\frac{\beta}{1-\alpha-\beta}} = \left(\frac{\theta \alpha}{r}\right)^{\frac{1-\alpha-\beta}{1-\alpha-\beta}} \left(\frac{\theta \beta}{w}\right)^{\frac{1-\alpha}{1-\alpha-\beta}}
\]
\[
\implies h^* = \left(\frac{\theta \alpha}{r}\right)^{\frac{1-\alpha-\beta}{1-\alpha-\beta}} \left(\frac{\theta \beta}{w}\right)^{\frac{1-\alpha}{1-\alpha-\beta}} = h_u
\]

\(k^*\) and \(h^*\) are guaranteed to be positive if \(\theta, \alpha, \beta\) and \(r\) are positive. To be feasible, \(k^*\) must be less than \(\lambda z\) and \(h^* < T\).

\[
h_u \leq T
\]
\[
\left(\frac{\theta \alpha}{r}\right)^{\frac{1-\alpha-\beta}{1-\alpha-\beta}} \left(\frac{\theta \beta}{w}\right)^{\frac{1-\alpha}{1-\alpha-\beta}} \leq T
\]
\[
\left(\frac{\theta}{r}\right)^{\frac{1-\alpha-\beta}{1-\alpha-\beta}} \left(\frac{\beta}{w}\right)^{\frac{1-\alpha}{1-\alpha-\beta}} \leq T
\]
\[
\theta^{1-\alpha-\beta} \leq T \left( \frac{r}{\alpha} \right)^{1-\alpha-\beta} \left( \frac{w}{\beta} \right)^{1-\alpha-\beta}
\]
\[
\theta \leq T^{1-\alpha-\beta} \left( \frac{r}{\alpha} \theta^{\alpha} \right) (w/\beta)^{1-\alpha} \equiv A(w)
\]

Therefore,
\[
h_u \leq T \iff \theta \leq A(w)
\]

and
\[
k_u \leq \lambda z
\]
\[
\theta \left( \frac{\theta}{r} \right)^{1-\alpha-\beta} \left( \frac{\theta}{w} \right)^{1-\alpha-\beta} \leq \lambda z
\]
\[
\theta^{1-\alpha-\beta} \left( \frac{\theta}{r} \right)^{1-\alpha-\beta} \left( \frac{\theta}{w} \right)^{1-\alpha-\beta} \leq \lambda z
\]
\[
\theta^{1-\alpha-\beta} \leq \lambda z \left( \frac{r}{\alpha} \right)^{1-\alpha-\beta} \left( \frac{w}{\beta} \right)^{1-\alpha-\beta}
\]
\[
\theta \leq \left( \lambda z \right)^{1-\alpha-\beta} \left( \frac{r}{\alpha} \right)^{1-\beta} \left( \frac{w}{\beta} \right)^{\beta} \equiv B(z, w)
\]

Therefore,
\[
k_u \leq \lambda z \iff \theta \leq B(z, w)
\]

Case 2:

Suppose at the optimum, a household chooses \( k = \lambda z \) and \( 0 < h < T \). For this to be an equilibrium,

\[
k > 0, \lambda_1 = 0
\]
\[
h > 0, \lambda_2 = 0
\]
\[
\lambda z - k = 0, \lambda_3 \geq 0
\]
\[
T - h > 0, \lambda_4 = 0
\]
\[
dL dh = -w + \theta(\lambda z)^{\alpha} h^{\beta-1} = 0
\]
\[
\implies h = \left( \frac{\theta(\lambda z)^{\alpha}}{w} \right)^{\frac{1}{\beta-1}} \equiv \hat{h}
\]

For \( \hat{h} < T \)

\[
\left( \frac{\theta(\lambda z)^{\alpha}}{w} \right)^{\frac{1}{\beta-1}} < T
\]
\[
\implies \theta < T^{1-\beta} \frac{w}{(\lambda z)^{\alpha} \beta} \equiv D(z, w)
\]

Therefore,
\[
\hat{h} < T \iff \theta < D(z, w)
\]

With respect to capital;
\[
dL dk = \theta \alpha (\lambda z)^{\alpha-1} h^\beta - r - \lambda_3 = 0
\]

115
\[
\theta \alpha (\lambda z)^{\alpha - 1} \left( \frac{\theta (\lambda z)^{\alpha \beta}}{w} \right) \frac{\beta}{1 - \beta} - r - \lambda_3 = 0 \\
\theta^{1+ \frac{\beta}{1 - \beta}} (\lambda z)^{\alpha - 1} \frac{w^\alpha}{w} \left( \frac{\beta}{w} \right) \frac{\beta}{1 - \beta} - r = \lambda_3
\]

where \( \lambda_3 \geq 0 \).

\[
\Rightarrow \theta^{1+ \frac{1}{1 - \beta}} (\lambda z)^{\alpha + \beta - 1} \frac{w^\alpha}{w} \left( \frac{1}{\lambda z} \right)^{\alpha + \beta - 1} - r = \lambda_3 \geq 0 \\
\theta^{1 - \beta} \geq \frac{r \left( \frac{w}{\beta} \right)^{\frac{\beta}{1 - \beta}} \left( \frac{1}{\lambda z} \right)^{\frac{1 - \beta}{1 - \beta}}}{\alpha} \\
\theta \geq \left( \frac{r}{\alpha} \right)^{1 - \beta} \left( \frac{w}{\beta} \right)^{\beta} \left( \frac{1}{\lambda z} \right)^{\alpha + \beta - 1} \\
\theta \geq \left( \frac{r}{\alpha} \right)^{1 - \beta} \left( \frac{w}{\beta} \right)^{\beta} (\lambda z)^{1 - \alpha - \beta} \equiv B(z, w)
\]

Households with \( \theta \geq B(z, w) \) and \( \theta < D(z, w) \) will invest \( k = \lambda z \) and \( h = \hat{h} \) in their business.

Case 3:

Suppose at the optimum, \( 0 < k < \lambda z \) and \( h = T \). If so, the following must hold;

\[
k > 0, \lambda_1 = 0 \\
h > 0, \lambda_2 = 0 \\
\lambda z - k > 0, \lambda_3 = 0 \\
T - h = 0, \lambda_4 \geq 0 \\
\frac{dL}{dk} = \theta \alpha k^{\alpha - 1} T^\beta - r = 0 \\
\Rightarrow \theta \alpha k^{\alpha - 1} T^\beta = r \\
k^* = \left( \frac{r}{\theta \alpha T^\beta} \right)^{\frac{1}{\alpha - 1}} \\
k^* = \hat{k} \equiv \left( \frac{\theta \alpha T^\beta}{r} \right)^{\frac{1}{1 - \alpha}}
\]

\( \hat{k} \) must be feasible. That is, \( \hat{k} \leq \lambda z \);

\[
\left( \frac{\theta \alpha T^\beta}{r} \right)^{\frac{1}{1 - \alpha}} \leq \lambda z \\
\Rightarrow \theta < (\lambda z)^{1 - \alpha} \frac{r}{\alpha T^\beta} \equiv C(z)
\]

Therefore,

\( \hat{k} \leq \lambda z \iff \theta < C(z) \)

Optimality condition with respect to \( h \) is
\[
\frac{dL}{dh} = -w + \theta k^{\alpha \beta} - \lambda_4 = 0
\]
where \(\lambda_4\) must be non-negative. If so,
\[
-w + \theta \left(\frac{\theta \alpha T^\beta}{r}\right)^{\frac{\alpha}{1-\alpha}} \beta \geq 0
\]
\[
\theta^{1+\frac{\alpha}{1-\alpha}} \left(\frac{\alpha}{r}\right)^{\frac{\alpha}{1-\alpha}} T^{\beta+\frac{1}{1-\alpha}} - w \geq 0
\]
\[
\theta^{\frac{1}{1-\alpha}} \geq \frac{w}{\beta} \left(\frac{\alpha}{r}\right)^{\frac{\alpha}{1-\alpha}} \frac{1}{T^{\alpha+\beta-1}}
\]
\[
\theta \geq T^{1-\alpha \beta} \left(\frac{r}{\alpha}\right)^{\alpha} \left(\frac{w}{\beta}\right)^{1-\alpha} \equiv A(w)
\]

Case 4:
Finally, suppose \(k = \lambda z\) and \(h_b = T\) at the optimum. If so,
\[
k > 0, \lambda_1 = 0
\]
\[
h > 0, \lambda_2 = 0
\]
\[
\lambda z - k = 0, \lambda_3 \geq 0
\]
\[
T - h = 0, \lambda_4 \geq 0
\]
\[
\frac{dL}{dk} = -w + \theta(\lambda z)^{\alpha \beta} T^{\beta-1} - \lambda_4 = 0
\]
where \(\lambda_4\) must be non-negative
\[
\implies \theta(\lambda z)^{\alpha \beta} T^{1-\beta} \geq w_a
\]
\[
\theta \geq \frac{w_a}{(\lambda z)^{\alpha \beta}} T^{1-\beta} \equiv D(z, w)
\]
Finally, with respect to \(k\):
\[
\frac{dL}{dk} = \theta(\lambda z)^{\alpha-1} T^\beta - r + \lambda_3 = 0
\]
where \(\lambda_3\) must be non-negative.
\[
\implies \theta(\lambda z)^{\alpha-1} T^\beta \geq r
\]
\[
\theta \geq \frac{r}{\alpha(\lambda z)^{\alpha-1} T^\beta}
\]
\[
\theta \geq (\lambda z)^{1-\alpha} \left(\frac{r}{\alpha T^\beta}\right) \equiv C(z)
\]
For households with \(\theta \geq D(z, w)\) and \(\theta \geq C(z)\), \(k^* = \lambda z\) and \(h^* = T\).
\[\Box\]
Other proofs

Proposition 3

Define gross income from entrepreneurship as \( q^E \equiv \theta k^\alpha \). The model predicts that the expected gross income conditional on \( M = 0 \), characteristics \((z, x)\) and the model parameters \( \psi \) is

\[
E(q^E|M = 0) = \left( \frac{\alpha}{r} \right) \frac{\alpha - \beta}{1 - \alpha} T^{\beta - \alpha} E(\theta^{\frac{1}{1 - \alpha}}) \frac{\Phi(\frac{\alpha - \beta - a}{\beta} - \Phi(\frac{\alpha - c}{\beta} - \Phi(a))}{\Phi(c) - \Phi(a)} P_{q1}^E + (\lambda z)^{\alpha} T^{\beta} E(\theta^{\frac{1}{1 - \alpha}}) \frac{\Phi(\sigma - \max(c,d))}{\Phi(\sigma - \max(c,d))} P_{q2}^E
\]

where

\[
E(\theta^{\frac{1}{1 - \alpha}}) = \exp(\frac{\theta}{1 - \alpha} + 2(1 - \alpha)^2)
\]

\[
E(\theta) = \exp(\theta + \frac{\alpha}{2})
\]

\[
P_{q1}^E = \frac{1_{A > B} \{ \Phi(c) - \Phi(a) \}}{1 - P(M = 1)}
\]

\[
P_{q2}^E = \frac{1_{A > B} \{ \Phi(c) + 1_{A < B} \{ 1 - \Phi(d) \}}{1 - P(M = 1)}
\]

Proof

\[
E(q^E|M = 0) = E(q^E|M = 0, h^* = T, k^* = \hat{k}) P(h^* = T, k^* = \hat{k}|M = 0) + E(q^E|M = 0, h^* = T, k^* = \lambda z) P(h^* = T, k^* = \lambda z|M = 0)
\]

From Proposition 1, \( h^* = T, k^* = \hat{k} \) is equivalent to \( A(w) < \theta < C(z) \), and \( h^* = T, k^* = \lambda z \) is equivalent to \( \theta > \max(C(z), D(z, w)) \). Therefore,

\[
E(q^E|M = 0, h^* = T, k^* = \hat{k}) = E(\theta^{\frac{2\alpha T^\beta}{r}} \frac{\alpha - \beta}{1 - \alpha} T^{\beta - \alpha} | h^* = T, k^* = \hat{k}, M = 0)
\]

\[
= \left( \frac{\alpha}{r} \right) \frac{\alpha - \beta}{1 - \alpha} T^{\beta - \alpha} E(\theta^{\frac{1}{1 - \alpha}} | h^* = T, k^* = \hat{k}, M = 0)
\]

\[
= \left( \frac{\alpha}{r} \right) \frac{\alpha - \beta}{1 - \alpha} T^{\beta - \alpha} E(\theta^{\frac{1}{1 - \alpha}}) \frac{\Phi(\frac{\alpha - \beta - a}{\beta} - \Phi(\frac{\alpha - c}{\beta} - \Phi(a))}{\Phi(c) - \Phi(a)}
\]

\[
E(q^E|M = 0, h^* = T, k^* = \lambda z) = E(\theta(\lambda z)^{\alpha} T^{\beta} | h^* = T, k^* = \lambda z, M = 0)
\]

\[
= (\lambda z)^{\alpha} T^{\beta} E(\theta | h^* = T, k^* = \lambda z, M = 0)
\]

\[
= (\lambda z)^{\alpha} T^{\beta} E(\theta^{\frac{1}{1 - \alpha}}) \frac{\Phi(\sigma - \max(c,d))}{\Phi(\sigma - \max(c,d))}
\]

Finally,

\[
P(h^* = T, k^* = \hat{k}|M = 0) = \frac{P(A(w) < \theta < C(z))}{P(M = 0)}
\]

\[
= \frac{1_{A < B} \{ \Phi(c) - \Phi(a) \}}{1 - P(M = 1)}
\]

\[
P(h^* = T, k^* = \lambda z|M = 0) = \frac{P(\theta > \max(C(z), D(z, w)))}{P(M = 0)}
\]

\[
= \frac{1_{A > B} \{ 1 - \Phi(c) \} + 1_{A < B} \{ 1 - \Phi(d) \}}{1 - P(M = 1)}
\]

\[
\square
\]

Proposition 4

Define gross income from entrepreneurship as \( q^E \equiv \theta k^\alpha \). The model predicts that the expected gross income conditional on \( M = 1 \), characteristics \((z, x)\) and the model parameters \( \psi \) is

\[
E(q^E|M = 1) = \left( \frac{\alpha}{r} \right) \frac{\alpha - \beta}{1 - \alpha} (\frac{\beta}{w})^{\alpha - \beta} T^{\beta - \alpha} E(\theta^{\frac{1}{1 - \alpha}}) \frac{\Phi(\min(a,b) - \frac{\alpha - \beta}{\beta} - \Phi(\frac{\alpha - c}{\beta} - \Phi(a))}{\Phi(a) - \Phi(\min(a,b))} P_{q1}^M + (\lambda z)^{\alpha} (\frac{\beta}{w})^{\alpha - \beta} T^{\beta - \alpha} E(\theta^{\frac{1}{1 - \alpha}}) \frac{\Phi(\sigma - a^\beta - \Phi(\frac{\alpha - c}{\beta} - \Phi(d))}{\Phi(\sigma - a^\beta - \Phi(\sigma - \max(d,c))} P_{q2}^M
\]

118
where
\[
E(\theta^{1-\alpha-\beta}) = \exp\left(\frac{\theta^\alpha}{1-\alpha-\beta} + \frac{\sigma^2}{2(1-\alpha-\beta)^2}\right)
\]
\[
E(\theta^{1-\beta}) = \exp\left(\frac{\theta^\beta}{1-\beta} + \frac{\sigma^2}{2(1-\beta)^2}\right)
\]
\[
P_{q_1^M} = \frac{1_{A>B}(\Phi(b) + 1_{A\leq B}(\Phi(a))}{P(M=1)}
\]
\[
P_{q_2^M} = \frac{1_{B=A}(\Phi(d) - \Phi(b))}{P(M=1)}
\]

**Proof**

\[
E(q^E|M=1) = E(q^E|M=1, h^* = h_u, k^* = k_u)P(h^* = h_u, k^* = k_u|M=1)
+ E(q^E|M=1, h^* = \hat{h}, k^* = \lambda z)P(h^* = \hat{h}, k^* = \lambda z|M=1)
\]

From Proposition 1, \(h^* = h_u, k^* = k_u\) is equivalent to \(\theta < \min(A(w), B(z, w))\), and \(h^* = \hat{h}, k^* = \lambda z\) is equivalent to \(B(z, w) < \theta < D(z, w)\). Therefore,

\[
E(q^E|M=1, h^* = h_u, k^* = k_u) = E\left(\theta^{1-\alpha-\beta}\left(\frac{\alpha}{w}\right)^{\frac{\alpha}{\theta}} - \left(\frac{\beta}{w}\right)^{\frac{\beta}{\theta}}|h^* = h_u, k^* = k_u, M = 1\right)
\]

\[
= \left(\frac{\alpha}{\theta}\right)^{\frac{\alpha}{\theta}} - \left(\frac{\beta}{\theta}\right)^{\frac{\beta}{\theta}} E\left(\theta^{1-\alpha-\beta}\right) = (\theta)^{\frac{\alpha}{\theta}} - \left(\frac{\beta}{\theta}\right)^{\frac{\beta}{\theta}} E\left(\theta^{1-\alpha-\beta}\right)
\]

\[
E(q^E|M=1, h^* = \hat{h}, k^* = \lambda z) = E\left(\theta(\lambda z)^{\alpha}\hat{h}^\beta|h^* = \hat{h}, k^* = \lambda z, M = 1\right)
\]

\[
= (\lambda z)^{\alpha} - \left(\frac{\beta}{\theta}\right)^{\frac{\beta}{\theta}} E\left(\theta^{1-\alpha-\beta}\right) = (\lambda z)^{\alpha} - \left(\frac{\beta}{\theta}\right)^{\frac{\beta}{\theta}} E\left(\theta^{1-\alpha-\beta}\right)
\]

Finally,

\[
P(h^* = h_u, k^* = k_u|M=1) = \frac{P(\theta<\min(A(w), B(z, w)))}{P(M=1)}
\]

\[
P(h^* = \hat{h}, k^* = \lambda z|M=1) = \frac{1_{A \leq B}(\Phi(b) + 1_{A < B}(\Phi(a))}{P(M=1)}
\]

\[
P(h^* = \hat{h}, k^* = \lambda z|M=1) = \frac{1_{B < D}(\Phi(d) - \Phi(b))}{P(M=1)}
\]