Why Form Business Partnerships?

by

Jungho Lee
Singapore Management University

Date: August 16, Tuesday
Time: 4 - 5:30 pm
Venue: Silcock Journal Room (AS2 06-23)
Chair: Park Jungjae
Why Form Business Partnerships?

Jungho Lee†

December 30, 2015

Abstract

I estimate a matching model of partnerships among startup owners to quantify the relative importance of productivity gain, financing gain, and moral hazard to the business partnership formation. The productivity gain accounts for 85% of the gain from the observed partnerships. For partners in the first quintile of wealth distribution, however, financing accounts for 80% of the gain. The moral hazard discourages 57% of partnerships. Despite potential gains from financing, most financially constrained agents do not form a partnership due to moral hazard and productivity loss.

JEL Classification: D23, D24, L22, L26, M13

Keywords: partnership, productivity, financial constraints, moral hazard, entrepreneurship, matching

*This paper is based on the first chapter of my PhD dissertation at Washington University in St. Louis. I am greatly indebted to Barton H. Hamilton for his guidance and advice. I am also very grateful to Tat Chan, George-Levi Gayle and Carl Sanders for their help and feedback on this project. I thank Mariagiovanna Baccara, Kyoung Jin Choi, Ignacio Esponda, Yumi Koh, Sunha Myong, JungJae Park, Bernardo Silveira, and Aloysius Siow for helpful comments. All errors are my own.

†School of Economics, Singapore Management University, 90 Stamford Road, Singapore 178903; Email: jungholee@smu.edu.sg.
1 Introduction

Finding a partner is one of the most important decisions for potential startup owners. Successful business owners often argue that finding the right partner is the key for a firm’s success.\footnote{See, for example, Cohen and Eisner (2010); Kawasaki (2004).} Indeed, many successful companies started as partnerships. Examples include Hewlett-Packard, Procter & Gamble, and Ben & Jerry’s. Yet, despite successful examples, only 15% of non-family owned businesses began with partnerships.\footnote{The Survey of Income and Program Participation (SIPP). See Section 2 in this paper.} Who forms business partnerships, and why? Answering this question can be valuable not only for potential startup owners but also for policy makers who try to boost entrepreneurship via partnerships.

Theoretical studies have identified gains and losses from business partnerships. Working with a partner may increase a firm’s productivity by a knowledge transfer between partners. Financially constrained entrepreneurs can increase financing capacity by finding a wealthy partner. The gains however can be offset by moral hazard; that is, the problem of inducing optimal effort by partners when the private return to effort is smaller than the marginal product of effort. The overall impact on entrepreneurs of the option to form a partnership will depend on the magnitude of productivity gain, financing gain, and moral hazard.

In this paper, I quantify the relative importance of productivity gain, financing gain, and moral hazard in business partnership formation to assess the role of the option to form a partnership on new firm formation. Using the Survey of Income and Program Participation (SIPP), a nationally representative household-based survey of the United States population, I first document the pattern of partnership choice of startup owners. First, business earnings of partners are significantly greater than those of measurably similar single owners. Second, the proportion of partners among startup owners is significantly greater below the 20th percentile than between the 20th and the 50th percentile of wealth distribution. Third, despite the high earnings among the observed partners, only 15% of startup owners form partnerships.

The high earnings of a partner may be driven either by productivity gains or by the increased investment from the other partner. The second observation in the preceding paragraph may further support the view that partnerships are driven by a financial motive since low-wealth individuals are more likely to be financially constrained (Evans and Jovanovic (1989)).
Moral hazard could be a reason behind the small proportion of partners among startup owners.

To disentangle the effect of financing and productivity motives and to quantify the cost of moral hazard, I develop and estimate a simple model of partnership formation. The model is an extension of Evans and Jovanovic (1989), an occupational choice model between a worker and an entrepreneur with collateral constraint. I incorporate an option to form a partnership into their model. By forming a partnership, financially constrained agents can increase borrowing capacity. I also allow for the possibility that working with a partner requires an additional ability such as collaborative skills. A partnership can be formed if there exist mutual gains in productivity or one partner’s gain in productivity and the other partner’s gain in financing despite the cost of moral hazard. Who is matched with whom is determined in a stable matching in which no partner would prefer to be a single owner or a worker, and one cannot find two partners who would both rather form a partnership with each other than remain with their current partners. The model is estimated by the method of simulated moments.

The estimated model implies three sets of results. First, among those who choose to be partners, the increase in productivity accounts for the major gains from partnerships: 85% of the aggregate gains for all partners is attributed to the productivity gains. However, the gains from financing are the major gains for low-wealth partners. For the partners in the first quintile of wealth distribution, financing accounts for 80% of the entire gain from partnerships. The cost of moral hazard is high among the observed partnerships. This cost corresponds to 39% of the entire gain from partnerships.

Second, when financial friction is introduced, most financially constrained agents do not form a partnership despite potential gains from financing. First of all, the cost of moral hazard is high at the extensive margin. The number of partners decreases 57% due to moral hazard. Moreover, working with a partner leads to a productivity loss for most individuals.

Third, financial friction generates “mismatch” among partners. The welfare loss associated with the mismatch is 3.3% of the aggregate welfare losses due to financial friction. However, there exists a group of people – mostly agents with enough wealth – who gain from the mismatch. The welfare gain for these agents is 2.3% of the aggregate welfare losses due to financial friction.

I compare the effectiveness of two government sponsored funding programs, one for all
startup firms and the other for only partnership firms. The former is more effective both in helping financially constrained startup owners to create businesses and in improving the match quality among partners.

This paper contributes to the literature on business partnerships. Extensive theoretical studies have examined business partnerships. For example, Lazear (1998) and Argote and Ingram (2000) argue that the knowledge transfer between partners can increase a firm’s productivity. Related to this view, other studies show that the partnership can be the optimal contract if the knowledge or a tacit human capital are not observable or not contractible. On the other hand, Jensen and Meckling (1992) and Becker and Murphy (1992) point out that a firm’s productivity can decrease due to difficulties in coordination and communication between partners. Regarding the cost of partnerships, moral hazard between partners has been the primary theoretical concern since Holmstrom (1982).

Contrary to extensive theoretical studies, empirical studies on business partnerships are rare. Gaynor and Pauly (1990) show that compensation incentives influence the productivity in medical group practices. Gaynor and Gertler (1995) and Lang and Gordon (1995) show that the extent of moral hazard among partners in medical group practices and in law partnerships depends on the degree to which the partners spread risk. Åstebro and Serrano (2015) show that for independent inventors having a partner can significantly increase the probability of commercialization success. These studies examine already established partnerships or partnerships in special groups (lawyers, physicians, or inventors). I examine the formation of business partnerships among potential startup owners in the general population. Moreover, no previous paper has studied how financial friction affects partnership formation among startup owners.

This paper also contributes to the literature on entrepreneurs and the creation of new

---

3 For example, see Teece (1980); Garicano and Santos (2004); Morrison and Wilhelm (2004); Bar-Isaac (2007).

4 The papers that study moral hazard in teams include Radner (1986); Rasmusen (1987); Legros and Matsushima (1991); Legros and Matthews (1991); Miller (1997); Strausz (1999); Battaglini (2006); Rahman and Obara (2010).

5 Another line of literature focuses on established professional service partnerships as an alternative to corporations (e.g., Levin and Tadelis (2005); Kaya and Vereshchagina (2014)).

6 Another key difference from Åstebro and Serrano (2015) is that I allow heterogeneous productivity gains (or losses) across agents induced by a matching equilibrium while Åstebro and Serrano (2015) assume a homogeneous productivity gain from adding a partner across all inventors.
businesses. The traditional view is that individuals are selected into entrepreneurship based on their entrepreneurial productivity (Lucas (1978)) and preference (Hamilton (2000)), while financial constraint may hinder the optimal allocation of entrepreneurial talent (Evans and Jovanovic (1989)).

Although relevant to potential startup owners, the option to form a partnership is rarely examined empirically. In particular, despite its potential role for financing, no previous paper has quantified to what extent the option to find a partner can help financially constrained agents to create a business. In one extreme, if all the financially constrained startup owners could manage to find a partner and successfully create new businesses, government subsidies aiming at helping those constrained startup owners may not be effective. I show that despite the potential gains from financing, most of the financially constrained potential startup owners do not choose to form partnerships due to the cost of moral hazard and the potential productivity loss associated with working with a partner.

This paper is related to the literature on estimating a one-sided matching model (e.g., Gordon and Knight (2009); Weese (forthcoming)). In attempting to overcome a challenge — we do not observe who is matched with whom in the data, especially with respect to net worth — I exploit the aggregate moments, notably the outcome from the matching. For example, I use the proportion of partners and the first moment of partners’ earning distribution conditional on net worth to identify the extent of financial friction. This is a new approach to estimating a matching model.

The paper is organized as follows. Section 2 describes the data. The baseline model is discussed in Section 3, and its extension into a matching model is presented in Section 4. Section 5 discusses the identification and the estimation of the matching model. The main results are presented in Section 6. Section 7 presents a welfare analysis. Policy experiments are discussed in Section 8. Section 9 concludes.

Following Evans and Jovanovic (1989), many studies indicate that financial constraint is an important factor for starting a business (Evans and Leighton (1989); Holtz-Eakin et al. (1994); Paulson et al. (2006); Cagetti and De Nardi (2006); Adelino et al. (2015); Schmalz et al. (forthcoming)). Hurst and Lusardi (2004) challenge this view by showing a nonlinear relationship between the probability of becoming a business owner and household wealth.

The U.S. Small Business Administration offers various guaranteed loan programs for startup owners.
2 Data

The Survey of Income and Program Participation (SIPP) is used for this study. The SIPP is a nationally representative household-based survey of the United States population designed to collect information for income and program participation. The SIPP was chosen for the following reasons. First, the sample size is large and the survey is nationally representative. The sample size and the time periods of each panel range from approximately 14,000 to 36,000 households and from 3 to 4 years respectively. I use panels including the years 1996, 2001, 2004 and 2008 and hence the time periods spans more than 10 years. Thanks to a large number of respondents, I can observe a relatively large number of individuals at the time they become business owners. Second, the SIPP contains information on the respondent’s equity share of the business. Using this information, I can define partners versus single owners. Finally, the SIPP provides information on the net worth at the household level, so that I can see the relationship between the probability of becoming a partner and household net worth.

Using the SIPP data I construct a two-year panel. To limit the influence of the labor market participation, I focus on males from ages 18 to 65 who are employed before they start their businesses. I drop family businesses since forming a partnership among household members does not increase the total value of household net worth, a mechanism this paper investigates. In the first year, all agents are workers. In the second year, some of the first year workers choose to be a business owner. I call the first year the base year and the second year the subsequent year. A detailed sample construction procedure is found in Appendix A.

2.1 Descriptive Analysis

A business partnership is characterized as a group that divides its profit among its members (equity holders). A similar definition is found in the literature (e.g., Holmstrom (1982); Farrell and Scotchmer (1988); Levin and Tadelis (2005)). A partner is defined as a business owner whose share of business equity is greater than or equal to 25% and less than or equal to 75%. Similarly, a business owner whose equity share is greater than 75% is defined as a single owner.\textsuperscript{9}

Table 1 reports the summary statistics for characteristics of workers, single owners, and

\textsuperscript{9}Given that most business owners have either 50\% or 100\% equity share, the results in this paper are not sensitive to this particular definition of partners versus single owners.
partners. Excluding family businesses, about 15% of startup owners chose to form partnerships. The years of education and the years of experience are similar to both types of owners. Partners are more likely to be white. They are also more likely to be married than single owners.

Figure 1 depicts the density of equity shares for those who have positive but less than 100% equity shares. Most of the equity shares are around 50%. This may suggest that two-owner partnerships with the same equity share are the most common ownership structure among startup partnerships.\footnote{Consistent with Figure 1, in Kauffman Firm Survey (KFS), a panel study of 4,928 businesses founded in 2004, 70% of multi-owner firms are two-owner firms, and 84% of two-owner firms that reported equity share have either 49% to 51% or 50% to 50% equity share.} The same observation is reported in Bitler et al. (2005).

Table 2 reports the summary statistics for incomes of workers, single owners, and partners. Figure 2 depicts the density of log incomes for each group. The median income of startup owners is less than that of workers as is well-documented in the literature (e.g., Hamilton (2000); Pugsley (2013)). Among business owners, median income and mean income of partners are greater than those of single owners. To determine whether the same pattern is observed after controlling for other observables, I conduct a regression analysis for log incomes among startup owners. Table 3 reports the estimates for the regression of log incomes on the partnership dummy and other covariates. I first start with the typical Mincer regression in the regression equation (1) and add various controls for other regressions. For all cases, the partnership dummy is significant and indicates that partners earn about 1.5 times more than the single owners.

Table 4 shows the summary statistics for the net worth of workers, single owners, and partners. Figure 3 depicts the density of net worth for each group. The wealth distribution of partners is more dispersed than that of single owners. Figure 4 plots the proportion of partners among startup owners with respect to net worth. The lower and upper percentile of wealth distribution is where the partnerships are relatively more observed. Among both tails, more partners are observed at the upper percentile of wealth distribution. To control for other observables, I conduct a Probit regression of partnership choice among startup owners including wealth dummies and other covariates. Table 5 reports the estimates from the Probit analysis. Consistent with Figure 4, the predicted probability to form a partnership is higher in the lower and upper percentile of wealth distribution than in the middle.
Table 6 reports the number and the proportion of partners for different industries. Partnerships are observed in all industries.\textsuperscript{11} Figure 5 shows the proportion of partnerships for high and low-starting capital industries. Based on the National Survey of Small Business Finances (NSSBF) in 1987, Hurst and Lusardi (2004) categorize construction and services as low-starting capital industries. They categorize mining; manufacturing; transportation and public utilities; wholesale and retail trade; and finance, insurance and real estate as high-starting capital industries. The partnership proportion for high-starting capital industries is significantly greater than for low-starting capital industries.

Discussion

The fact that the partners earn 1.5 times more than single owners even after controlling for other observables could be explained by two hypotheses. The first hypothesis is that the high earnings stem from productivity gains through the partnership. The second hypothesis is that the high earnings stem from the investment by a wealthy partner in a highly productive, but financially constrained partner.

The fact that relatively more partners are observed in the lowest quintile than in the middle of wealth distribution may support the view that partnerships are driven by startup owners’ financial motive. Furthermore, the fact that relatively more partners are observed in the high-starting capital industries may also support this finance-based hypothesis. One aspect of the partnership is similar to equity financing: financially constrained potential startup owners may give up some portion of equity in exchange for financing from a partner. Therefore, partnerships can be particularly attractive for low-wealth individuals who are more likely to be financially constrained.

Extensive theoretical studies identify moral hazard as a primary concern for working with a partner. The fact that only 15% of startup owners choose to form partnerships, despite the high earnings among the observed partners, may be to some extent explained by moral hazard costs.

Risk sharing is considered the main motive for joining a partnership among lawyers or physicians (Gaynor and Gertler (1995); Lang and Gordon (1995)). However, certain features

\textsuperscript{11}The number of startup owners in the Mining industry is one, and the owner is a single owner.
of the data are not consistent with the risk sharing hypothesis predictions. First of all, most partnerships are formed by two partners. If the reason behind finding a partner is to share risk, the number of partners within a firm would vary depending on the partners’ heterogenous risk preference. For example, if an agent is highly risk-averse, he would find not just one but many partners who have a similar risk preference. Indeed, Gaynor and Gertler (1995) and Lang and Gordon (1995) report sufficient variation on the group size for physicians and law partnerships.\footnote{For example, Gaynor and Gertler (1995) report that the average and standard deviation for the number of partners in law partnerships are 21.13 and 8.76 respectively.}

Second, as shown in Figure 4, the proportion of partners increases with wealth level after the 20th percentile of wealth distribution. The opposite relation is predicted if risk sharing is the main motive for forming partnerships and a decreasing absolute risk aversion preference is assumed.\footnote{Consider a choice between forming a partnership and being a single owner conditional on deciding to start a business. Suppose forming a partnership generates a certain payoff while being a single owner generates a risky payoff. Cressy (2000) shows that the probability of becoming a single owner increases as individual wealth increases if agents’ preference exhibits a decreasing absolute risk aversion.}

Although the data pattern is consistent with several hypotheses regarding partnerships, it is difficult to quantify each one without further explanation of the data generating process. In the next section, I present a simple model of partnership formation. I use this model as a measurement tool to quantify the productivity gain, the financing gain, as well as the cost of moral hazard. The model is also used for welfare and policy analyses.

3 Baseline Model

I first present a baseline model in which the distribution of partnership productivity and the distribution of potential partner’s net worth are fixed. Most intuitions are clearly addressed by using the baseline model. In Section 4, I extend the baseline model into a matching model in which the partnership productivity and the partner’s net worth are endogenously determined as an equilibrium outcome.
3.1 Environment

The model is an extension of Evans and Jovanovic (1989). I incorporate effort and heterogeneous partnership productivity into their model.

Preference

The utility is linear in consumption. The effort is incorporated to capture moral hazard between partners. Under the quadratic form, I can derive an analytic value function, which significantly reduces the computational burden in the estimation. The marginal rate of substitution between consumption and effort is captured by $\kappa$.

\[
u(c, z) = c - \kappa \frac{z^2}{2}, \quad c: \text{consumption}, \ z: \text{effort}
\]

Technology

Income as a worker is determined by worker productivity ($\theta_w$) and efforts. Income as a single owner is determined by solo productivity ($\theta_s$), capital investment ($k$), and efforts.

\[
w = \theta_w z^{1-\alpha} \epsilon_w \quad \text{if} \quad d = 1 \quad (\text{worker}) \]

\[
y = \begin{cases} 
\theta_s k^\alpha z^{1-\alpha} - rk \epsilon_s & \text{if} \quad d = 2 \quad (\text{single owner}) \quad \alpha \in (0, 1) \\
\theta_p k^\alpha (z + z')^{1-\alpha} - rk \epsilon_p & \text{if} \quad d = 3 \quad (\text{partner})
\end{cases}
\]

The partnership output is determined by partnership productivity ($\theta_p$), the joint capital investment, and the efforts from both partners. Under this partnership production function, the efficient outcome from two identical partners is the same as the sum of each partner’s outcome if the solo and partnership productivity are the same (Corollary 1 in Section 3.3). Moreover, the value from single ownership is the same as the value from hiring the potential partner as an employee if we interpret hiring an agent as a situation in which (1) the productivity does not change from the solo productivity, (2) the wage contract is set to eliminate moral hazard of the worker, and (3) the wage is equal to the marginal benefit of hiring the agent. I allow the earning shocks to workers ($\epsilon_w$), to single owners ($\epsilon_s$), and to partners ($\epsilon_p$) with the assumption
that the expectation of the shocks is equal to one.\textsuperscript{14}

**Financial Market**

Without financial friction, agents can borrow any amount of money for the risk-free gross interest rate $r$. A financial friction is modelled as a limited commitment between borrowers and lenders. The fact that lenders cannot force borrowers to repay a loan limits the maximum amount of borrowing. I assume that the maximum borrowing amount depends on the borrowers’ net worth. More specifically, a single owner can borrow up to $(\lambda - 1)A$. $\lambda$ is a constant which is greater than or equal to one. Therefore, the maximum amount of investment by the single owner is $(\lambda - 1)A + A = \lambda A$. If he forms a partnership with another agent with net worth $A'$, then the partnership can invest up to $\lambda(A + A')$.

### 3.2 Value Functions and Earned Incomes

I first consider the value function and earned incomes for partners. Equal sharing is a defining feature of partnerships. For example, Farrell and Scotchmer (1988) define a partnership as “a group that divides its output equally among its members.” Indeed, many data sources, including the sample in this paper, indicate that two-owners with equal shares are the most common partnership structure. Supported by this fact, I assume that a partnership is formed by two agents and that the equity shares are distributed equally among partners.\textsuperscript{15}

Effort in an entrepreneurial team is spent on activities such as generating an idea, validating the business idea, and developing a business model. Effort in these types of activities is hard to observe or monitor. For this reason, I assume that effort is not observable.

Suppose agent $i$ and agent $j$ are matched and their partnership productivity is given as $\theta_p$. The partners choose $\{z_i, z_j, k\}$, their effort levels and the joint investment. Given $\{z_j, k\}$, the

\textsuperscript{14}Business owners in SIPP were asked to report their profits as business earnings. To make the model consistent with the data, I specified a random shock as a shock to a firm’s profit.

\textsuperscript{15}Many studies investigate theoretical implications of equal sharing in partnerships (e.g., Farrell and Scotchmer (1988); Sherstyuk (1998); Levin and Tadelis (2005)). Others provide a rationale for equal sharing among partners (e.g., Bartling and von Siemens (2010); Espino et al. (2015)).
best response function for $z_i$ solves

$$\max_{z_i} \mathbb{E} \left[ \frac{1}{2} \left\{ \theta_p k^\alpha (z_i + z_j)^{1-\alpha} - r k \right\} \epsilon_p \right] - \kappa \frac{z_i^2}{2}$$

Likewise, given $\{z_i, k\}$, the best response of $z_j$ solves

$$\max_{z_j} \mathbb{E} \left[ \frac{1}{2} \left\{ \theta_p k^\alpha (z_i + z_j)^{1-\alpha} - r k \right\} \epsilon_p \right] - \kappa \frac{z_j^2}{2}$$

Finally, given $\{z_i, z_j\}$, $k$ solves

$$\max_k \mathbb{E} \left[ \left\{ \theta_p k^\alpha (z_i + z_j)^{1-\alpha} - r k \right\} \epsilon_p \right]$$

The earned incomes and value function per partner are given as

$$\pi_p(\theta_p, A_p) = \frac{1}{2} \left[ \theta_p k^\alpha (2z^*)^{1-\alpha} - r k^* \right] \epsilon_p$$

$$= \begin{cases} 
  \alpha \theta_p \frac{2}{1+\alpha} \epsilon_p & \text{if } \lambda A_p \geq \hat{a} \theta_p \frac{2}{1+\alpha} \\
  \left\{ a_8 \left( \frac{\lambda A_p}{2} \right) \frac{2}{1+\alpha} \theta_p \frac{2}{1+\alpha} - \frac{r \lambda A_p}{2} \right\} \epsilon_p & \text{if } \lambda A_p < \hat{a} \theta_p \frac{2}{1+\alpha} \end{cases}$$

$$V_p(\theta_p, A_p) = \mathbb{E} [\pi_p(\theta_p, A_p)] - \kappa \frac{z^*}{2}$$

$$= \begin{cases} 
  \alpha \theta_p \frac{2}{1+\alpha} & \text{if } \lambda A_p \geq \hat{a} \theta_p \frac{2}{1+\alpha} \\
  \left\{ a_4 \left( \frac{\lambda A_p}{2} \right) \frac{2}{1+\alpha} \theta_p \frac{2}{1+\alpha} - \frac{r \lambda A_p}{2} \right\} & \text{if } \lambda A_p < \hat{a} \theta_p \frac{2}{1+\alpha} \end{cases}$$

where

$$A_p = A_i + A_j, \quad \hat{a} = \left( \frac{1 - \alpha}{\kappa} \right) \left( \frac{\alpha}{r} \right) \frac{1+\alpha}{\alpha}, \quad a_3 = \hat{a} r \left( \frac{1 - \alpha}{\alpha} \right) \left( \frac{3}{8} \right),$$

$$a_4 = \left( \frac{1 - \alpha}{2\kappa} \right) \frac{1+\alpha}{\alpha} \left( \frac{3 + \alpha}{4} \right), \quad a_1 = \hat{a} r \left( \frac{1 - \alpha}{2\alpha} \right), \quad a_8 = \left( \frac{1 - \alpha}{2\kappa} \right) \frac{1+\alpha}{\alpha}$$
Given the worker productivity $\theta_w$, workers choose effort to maximize their utility. The earned incomes and value function as a worker is given as

$$\pi_w(\theta_w) = a_6 \theta_w^{\frac{2}{1+\alpha}} \epsilon_w$$

$$V_w(\theta_w) = a_2 \theta_w^{\frac{2}{1+\alpha}}$$

where

$$a_2 = \left(1 - \frac{\alpha}{\kappa}\right)^{\frac{1-\alpha}{1+\alpha}} \left(\frac{1 + \alpha}{2}\right), \quad a_6 = \left(1 - \frac{\alpha}{\kappa}\right)^{\frac{1-\alpha}{1+\alpha}}$$

Single owners optimally choose the amount of investment and efforts, and the earned incomes and value function are given as

$$\pi_s(\theta_s, A) = \begin{cases} 
2a_1 \theta_s^{\frac{2}{1+\alpha}} \epsilon_s & \text{if } \lambda A \geq \hat{a} \theta_s^{\frac{2}{1+\alpha}} \\
\{a_6(\lambda A) \frac{2\alpha}{1+\alpha} \theta_s^{\frac{2}{1+\alpha}} - r \lambda A\} \epsilon_s & \text{if } \lambda A < \hat{a} \theta_s^{\frac{2}{1+\alpha}}
\end{cases}$$

$$V_s(\theta_s, A) = \begin{cases} 
a_1 \theta_s^{\frac{2}{1+\alpha}} & \text{if } \lambda A \geq \hat{a} \theta_s^{\frac{2}{1+\alpha}} \\
a_2(\lambda A)^{\frac{2\alpha}{1+\alpha}} \theta_s^{\frac{2}{1+\alpha}} - r \lambda A & \text{if } \lambda A < \hat{a} \theta_s^{\frac{2}{1+\alpha}}
\end{cases}$$

To summarize, the individual choice ($d$) and the corresponding conditional income ($\pi$) are

$$d = \arg \max \{V_w(\theta_w), V_s(\theta_s, A), V_p(\theta_p, A_p)\}$$

$$\begin{align*}
\pi &= \begin{cases} 
\pi_w(\theta_w) & \text{if } d = 1 \text{ (worker)} \\
\pi_s(\theta_s, A) & \text{if } d = 2 \text{ (single owner)} \\
\pi_p(\theta_p, A_p) & \text{if } d = 3 \text{ (partner)}
\end{cases}
\end{align*}$$
3.3 Decomposing Benefits and Costs of Partnerships

Before further discussions, I derive the value for each partner from the first best allocation. In doing so, I assume that the aggregate production from the efficient allocation is distributed equally among partners.

**Proposition 1.** The value function for each partner from the efficient allocation under financial friction is given as

\[
V_p^E(\theta_p, A_p) = \begin{cases} 
  a_1 \theta_p^{\frac{2}{1-\alpha}} & \text{if } \frac{\lambda A_p}{2} \geq \hat{a} \theta_p^{\frac{2}{1-\alpha}} \\
  a_2 \left( \frac{\lambda A_p}{2} \right)^{\frac{2\alpha}{1+\alpha}} \theta_p^{\frac{2}{1+\alpha}} - r \frac{\lambda A_p}{2} & \text{if } \frac{\lambda A_p}{2} < \hat{a} \theta_p^{\frac{2}{1-\alpha}}
\end{cases}
\]

**Proof.** See Appendix C.1.

With Proposition 1 in hand, it is straightforward to prove the following corollaries:

**Corollary 1.** \(V_p^E(\theta_s, 2A) = V_s(\theta_s, A)\)

Suppose two identical agents form a partnership and the solo and partnership productivity are the same. Corollary 1 says that the sum of each partner’s value as a single owner is equal to the aggregate value of the partnership. Three channels make the value from partnerships different from the value from single ownership.

**Corollary 2.** \(V_p^E(\theta_p, 2A) = V_s(\theta_p, A) > V_s(\theta_s, A)\) if and only if \(\theta_p > \theta_s\)

Suppose \(A = A'\). The first best value from partnerships is greater than the value from single ownership if and only if the partnership productivity is greater than the solo productivity. The value difference in Corollary 2 captures the value generated by the partnership productivity.

**Corollary 3.** \(V_p^E(\theta_s, A_p) > V_s(\theta_s, A)\) if and only if \(A' > A\) and \(\lambda A < \hat{a} \theta_s^{\frac{2}{1-\alpha}}\)

Although the partnership productivity and the solo productivity are the same, if an agent is financially constrained as a single owner and his partner’s net worth is greater than his net worth, the value from the partnership is strictly greater than the value from single ownership if
there is no moral hazard. The value difference in Corollary 3 captures the value generated by financing. Notice that if there is no financial friction, the value difference induced by financing would be zero.

**Corollary 4.** $V_p(\theta_s, 2A) < V_p^E(\theta_s, 2A) = V_s(\theta_s, A)$

Finally, the cost of moral hazard is captured by the value difference in Corollary 4.

To summarize,

$$V_p(\theta_p, A_p) - V_s(\theta, A) = \underbrace{V_p^E(\theta_p, 2A) - V_s(\theta_s, A)}_{\Omega_1} + \underbrace{V_p^E(\theta_p, A_p) - V_s(\theta_p, A)}_{\Omega_2} + \underbrace{V_p(\theta_p, A_p) - V_p^E(\theta_p, A_p)}_{\Omega_3}$$

(2)

where $\Omega_1$ represents gains from productivity, $\Omega_2$ represents gains from financing, and $\Omega_3$ captures the cost of moral hazard.

## 4 Matching Model

The baseline model does not provide any prediction on how the partnership productivity and the partner’s net worth are determined. In this section, I extend the baseline model into a matching model in which the partnership productivity and the partner’s net worth are generated as an equilibrium outcome. The model environment is identical to the baseline model. Two additional structures are imposed.

### 4.1 Matching Function and Complementarities

First, instead of treating partnership productivity as fixed, I model it as a function of the solo productivity of each partner. More specifically,

$$\tilde{\theta}_p(i, j) = \sqrt{\tilde{\theta}_i \tilde{\theta}_j} \quad \text{where} \quad \tilde{\theta}_i = g_i \theta_{si}, \quad \tilde{\theta}_j = g_j \theta_{sj}, \quad g_i, g_j > 0$$

(3)

Knowledge transfer has been considered the key mechanism through which teamwork increases productivity (Lazear (1998); Argote and Ingram (2000)). However, knowledge transfer is hard
and it can be costly due to difficulties in coordination and communication (Arrow (1969); Teece (1977); Becker and Murphy (1992); Jensen and Meckling (1992); Haas and Hansen (2005); Jones (2009)). “Collaborative skill” (Hamilton et al., 2003) and “willingness to cooperate” (Kosfeld and von Siemens, 2011) are necessary to facilitate the knowledge transfer. More importantly, such characteristics are heterogenous across individuals independent of the quality of their business idea. \( g \) captures this additional element for teamwork. I call this additional element “collaborative skill.” Depending on individual skills for collaborating, it can enhance \((g > 1)\) or reduce \((g < 1)\) the solo productivity when partners work together. As a result, the productivity as a partner \((\tilde{\theta})\) can be different from the solo productivity \((\theta_s)\).

To model the partnership productivity in a tractable way, I impose a couple of assumptions in (3). First, the collaborative skill \((g)\) interacts with the solo productivity \((\theta_s)\). Second, the contribution to partnership productivity by each partner is the same. Moreover, if two identical partners are matched, the partnership productivity is the same as their individual productivity as a partner. In particular, if we shut down the additional productivity channel by making \(g\) equal to one, the partnership productivity is the same as the solo productivity when two identical agents are matched.

To explain how a partnership can be formed, I introduce the following two definitions.

**Definition 1.** There exists complementarity between two partners’ solo productivities in a partnership by Agent \(i\) and Agent \(j\) if \(\theta_p(i, j) \geq \max\{\theta_{si}, \theta_{sj}\}\).

**Definition 2.** There exists complementarity between one partner’s solo productivity and the other partner’s wealth in a partnership by Agent \(i\) and Agent \(j\) if (1) one partner, say Agent \(i\), is financially constrained as a single owner, and (2) the other partner, Agent \(j\), has more net worth than the financially constrained partner and \(\theta_p(i, j) > \theta_{sj}\).

Suppose there is no borrowing constraint. Then \(V_p^E(\theta_p(i, j), A_p) \geq \max\{V_s(\theta_{si}, A_i), V_s(\theta_{sj}, A_j)\}\) if and only if \(\theta_p(i, j) \geq \max\{\theta_{si}, \theta_{sj}\}\) by Corollary 2 in Section 3.3. This means that without financial friction a partnership is formed only if there exists complementarity between two partners’ solo productivities.

If there is a borrowing constraint, a partnership is formed even without the complementarity between two partners’ solo productivities. The intuition is found in Corollary 3. If a partner
is financially constrained as a single owner, he may be willing to sacrifice his productivity for financing. Definition 2 captures this situation.

The complementarity between two partners’ solo productivities and the complementarity between one partner’s solo productivity and the other partner’s wealth are not mutually exclusive. For example, a financially constrained agent may be matched with a more productive and wealthier agent if his productivity as a partner is large enough. To determine an equilibrium, who is matched with whom, I impose another structure on the matching market.

4.2 Matching Market

I model the matching market as a one-sided, non-transferable utility, frictionless market. Unlike the marriage matching market, I cannot separate market participants into two groups. The assumption that the equity share is fixed as 50% to 50% rules out the possibility that two partners can transfer the outcome. Finally, I assume there is no search friction in the market.\(^{16}\)

There are \(N\) number of agents. Let \(I\) denote the set of agents. An agent \(i \in I\) is characterized by \(\{g_i, \theta_{si}, A_i\}\). I assume that \(\theta_{si}\) is drawn from a continuous distribution and therefore every agent has a different value of \(\theta_s\). A matching \(\Gamma\) is a one-to-one mapping from \(I\) onto itself such that for all \(\{i,j\} \subset I\), \(\Gamma(i) = j\) if and only if \(\Gamma(j) = i\). A matching is stable if (i) no partner would rather prefer to be single owner or worker, and (ii) one cannot find two partners who would both rather form a partnership with each other than remain with their current partners.

**Proposition 2.** There exists a unique stable matching.

**Proof** With 50% to 50% equity share, the value per partner from any partnership is symmetric. Moreover, for any agent \(i \in I\), the value from partnership with \(k \in I\) — including to remain as single — is different for all \(k\) since every agent has a different value of \(\theta_s\). In other words, every agent has a strict preference over \(I\). Under these two conditions, there exists a unique stable matching by Gordon and Knight (2009). \(\blacksquare\)

Without borrowing constraint, characterizing a stable matching is straightforward by the

---

\(^{16}\)As a robustness check, I re-estimated the model under the assumption that two agents in the population are randomly matched to each other, and are not allowed to form a partnership with any other agent except for the randomly matched agent. The result is discussed in Appendix E.
matching function in (3). For expository purpose, assume that \( g \) is so large that every agent wants to form a partnership rather than remain as either worker or single owner. Sort agents in \( I \) with respect to \( \tilde{\theta} \) and denote the agent with the highest \( \tilde{\theta} \) be one and so on. With (3), every agent wants to be matched with Agent 1. Then it is Agent 1’s decision to choose his partner. Again, under (3), Agent 1 wants to form a partnership with Agent 2. By iterating this procedure, we can find a stable matching. The stable matching exhibits an assortative matching with respect to \( \tilde{\theta} \).

With borrowing constraint, however, the stable matching under no borrowing constraint can alter because of the financial motive for financially constrained agents. For example, consider a partnership formed by Agent 1 and Agent 2 in the above example. If financial friction is introduced, Agent 1 and Agent 2 may not be able to finance the optimal amount of capital, and the value from their partnership may decrease. Suppose Agent 3 has a marginally lower productivity but has a much higher net worth than Agent 2. With financial friction, Agent 1 may prefer to be matched with Agent 3 since he can increase financing capacity with Agent 3 without sacrificing much productivity. To illustrate this intuition more clearly, I present a simple example.

**Example**

Suppose there are four agents in the market. The outside option value for workers is assumed to be zero and hence the relevant characteristics for matching is \( \{g_i, \theta_s i, A_i\} \). Let’s assume \( g_i \) is 1.2 so that we can focus on \( \{\theta_s i, A_i\} \). Assume that

\[
\{\theta_{s1}, A_1\} = \{2.1, 0\}, \quad \{\theta_{s2}, A_2\} = \{2, 0\}, \quad \{\theta_{s3}, A_3\} = \{1.1, \infty\}, \quad \{\theta_{s4}, A_4\} = \{1, \infty\}
\]

For convenience, let \( a_1 = 1, \alpha = 0.5 \).

First, I consider a situation in which there is no borrowing constraint. The outside option value as a single owner for each agent is the following.

\[
V_{s1} = \theta^4_{s1} = 19.45, \quad V_{s2} = \theta^4_{s2} = 16, \quad V_{s3} = \theta^4_{s3} = 1.46, \quad V_{s4} = \theta^4_{s4} = 1
\]

It is easy to check that Agent 1 is matched with Agent 2 and Agent 3 is matched with Agent...
4 in the stable matching. The value from the matching is

\[ V_p(1, 2) = \frac{3}{4}(\theta_p(1, 2)^4) = 27.43, \quad V_p(3, 4) = \frac{3}{4}(\theta_p(3, 4)^4) = 1.88 \]

Note that \( \theta_p(1, 2) \geq \max\{\theta_{s1}, \theta_{s2}\} \) and \( \theta_p(3, 4) \geq \max\{\theta_{s3}, \theta_{s4}\} \) and hence both partnerships are driven by the complementarity between two partners’ solo productivities.

Suppose now that there is a borrowing constraint. The outside option value as a single owner changes.

\[ V_{s1} = 0, \quad V_{s2} = 0, \quad V_{s3} = \theta_{s3}^4 = 1.46, \quad V_{s4} = \theta_{s4}^4 = 1 \]

In this situation, it is easy to check that Agent 1 is matched with Agent 3 and Agent 2 is matched with Agent 4 in the stable matching. The value from the matching is

\[ V_p(1, 3) = \frac{3}{4}(\theta_p(1, 3)^4) = 8.3, \quad V_p(2, 4) = \frac{3}{4}(\theta_p(3, 4)^4) = 6.22 \]

Note that \( \theta_p(1, 3) < \max\{\theta_{s1}, \theta_{s2}\} \) and \( \theta_p(2, 4) < \max\{\theta_{s3}, \theta_{s4}\} \) and hence both partnerships are driven by the complementarity between one partner’s solo productivity and the other partner’s wealth, not by the complementarity between two partners’ solo productivities.

5 Estimation

5.1 Specification

Before estimation, I specify worker productivity (\( \theta_w \)), solo productivity (\( \theta_s \)), and the collaborative skill (\( g \)). The specification is similar to the one in Evans and Jovanovic (1989), Xu (1998), Paulson et al. (2006).

Worker Productivity

The worker productivity is determined by education (\( x_1 \)) and experience (\( x_2 \)).

\[ \log \theta_w = \gamma_1 \log x_1 + \gamma_2 \log x_2 \]
To identify $\alpha$ and $\kappa$, I normalize the constant term in worker productivity as zero. A stochastic component in $\theta_w$ in $\theta_s$, or in $g$ cannot be identified. As in Evans and Jovanovic (1989), I assume no unobserved heterogeneity in the worker productivity.

**Solo Productivity**

The solo productivity is assumed to be a function of education, experience, net worth and an unobserved component.

$$\log \theta_s = \beta_0 + \beta_1 \log A + \beta_2 \log x_1 + \beta_3 \log x_2 + \eta, \quad \text{where} \quad \eta \sim N(0, \sigma^2_\eta)$$

Similar to Evans and Jovanovic (1989) and Paulson et al. (2006), I allow a possible relationship between an agent’s net worth and his value as a single owner even without financial friction.

**Collaborative Skill**

I specify $g$ as a function of observable and unobservable component.

$$\log g = g_0 + g_1 \log x_1 + g_2 \log x_2 + u, \quad u \sim N(0, \sigma^2_u) \quad (4)$$

As explained in Section 4.1, $g$ represents an additional ability as a partner such as collaborative skill. Education and work experience are inevitably involved with social interaction, and therefore they may help to develop the collaborative skill.

A critical assumption is that $g$ is not a function of net worth. As explained in Section 5.2, this assumption is required for identification of $\lambda$. Note that even if $g$ is not a function of net worth, the partnership productivity is a function of net worth because it is also affected by $\theta_s$ which is a function of net worth.

With the specification in (4), the matching function — log value of the partnership productivity — can be written as

$$\log \theta_p(i, j) = g_0 + \tilde{g}_1 \log(x_{1i} x_{1j}) + \tilde{g}_2 \log(x_{2i} x_{2j}) + \epsilon_{<i,j>}$$

where $\tilde{g}_1 = \frac{g_1}{2}$, $\tilde{g}_2 = \frac{g_2}{2}$ and $\epsilon_{<i,j>} = \frac{1}{2}(u_i + u_j + \log \theta_{si} + \log \theta_{sj})$. $x_{1i}$ and $x_{1j}$ indicate the year of
education for agent $i$ and $j$ respectively. Similarly, $x_{2i}$ and $x_{2j}$ indicate the year of experience for agent $i$ and $j$ respectively. $\epsilon_{<i,j>}$ represents an unobserved match specific component.

**Earning Shocks**

The expected value of the stochastic components ($\epsilon_w, \epsilon_s, \epsilon_p$) is equal to one, allowing zero income for business owners. As shown in Figure 2, the distribution of conditional income for business owners looks similar to a log-normal distribution with a possible realization of a zero or a negative income. The current specification captures this stylized feature of the data.

$$\log \epsilon_w \sim N(\mu_w, \sigma^2_w), \quad \mathbb{E}[\epsilon_w] = 1$$

$$\epsilon_s = \tilde{\epsilon}_s - P_s, \quad \log \tilde{\epsilon}_s \sim N(\mu_s, \sigma^2_s), \quad \mathbb{E}[\epsilon_s] = 1, \quad P_s \text{ is a positive constant.}$$

$$\epsilon_p = \tilde{\epsilon}_p - P_p, \quad \log \tilde{\epsilon}_p \sim N(\mu_p, \sigma^2_p), \quad \mathbb{E}[\epsilon_p] = 1, \quad P_p \text{ is a positive constant.}$$

### 5.2 Identification

The matching model is fully characterized by a set of parameters: \{\kappa, \gamma_1, \gamma_2, \alpha, \beta_0, \beta_1, \beta_2, \beta_3, \sigma_\eta, g_0, g_1, g_2, \sigma_u, \sigma_w, \sigma_s, \sigma_p, P_s, P_p, \lambda\}. A formal identification argument for the baseline model is shown in Appendix B. I use the intuition developed in the identification of the baseline model, and exploit the aggregate moments, notably the outcome from a matching, to identify the matching model.

The proportion of single owners among those who are not partners identifies $\beta_0$. Given $\lambda$, the proportion of single owners conditional on net worth among those who are not partners identifies $\beta_1$. Likewise, the proportion of single owners conditional on education and on experience identify $\beta_2$ and $\beta_3$ respectively. Among business owners, the proportion of partners identifies $g_0$. Likewise, the proportion of partners among business owners conditional on education and on experience identify $g_1$ and $g_2$ respectively.

The average income for workers conditional on education and on experience identify $\gamma_1$ and $\gamma_2$ respectively. The average income for workers, for single owners, and for partners identify $\kappa$ and $\alpha$. The average income for single owners conditional on education (or experience) identify $\sigma_\eta$. The average income for partners conditional on education (or experience) identify $\sigma_u$.
A couple of data moments identify $\lambda$. First, the proportion of partners among business owners conditional on net worth is only affected by $\lambda$ given $\beta_1$. Consider, for example, business owners with a small net worth in which the measure of constrained single owners is positive. As $\lambda$ becomes smaller, the value of those constrained single owners also decrease. As a result, those whose value as a single owner is marginally higher than the value as a worker or as a partner will change their choice either to workers or to partners. In either case, the proportion of partners among business owners conditional on the small net worth will increase. Second, $\lambda$ also affects the average income both for single owners and for partners conditional on net worth by affecting the income for financially constrained business owners.

The variance of log incomes for workers identifies $\sigma_w$. The variance of log incomes for single owners identifies $\sigma_s$ given $P_s$. Given $\sigma_s$, the proportion of zero or a negative income single owners identifies $P_s$. Likewise, the variance of log incomes for partners and the proportion of zero or a negative income partners identify $\sigma_p$ and $P_p$.

5.3 Defining a Market

I define the entire United States in a particular year as the market for estimation as in Choo and Siow (2006). In existing empirical work on matching markets, a market is defined differently depending on the purpose of the work and the availability of the data. I choose a particular year since there is no way that an agent in a year could have been matched with an agent in another year. I choose the 2004 panel since it has the most observations for both single owners and partners. Location may be a relevant factor for a matching. However, I cannot rule out the possibility that an agent in California may form a partnership with a friend in New York. Therefore, I consider the entire United States as one market. Industry may seem to be a relevant factor as well. However, in the data, the industry as a business owner does not always correspond to the industry as a worker. For example, one partner in the Agricultural sector used to be in the Construction sector as a worker. Similarly, one partner in the Construction sector used to be in the Vocational Rehabilitation Service sector as a worker. Therefore, I did not put a restriction on the matching market based on agents’ industry as a worker. To summarize, I use the entire 2004 panel for estimation.
5.4 Method of Simulated Moments

The analytic derivation of likelihood function is not feasible. I use the method of simulated moments to estimate the model parameters. The criterion function is given in (5).

\[
M(\psi) = \left[ \sum_{i=1}^{n} Z_i (K_i - \tilde{K}_i(\psi)) \right]' \hat{\Sigma}^{-1} \left[ \sum_{i=1}^{n} Z_i (K_i - \tilde{K}_i(\psi)) \right]
\]

(5)

where \( \tilde{K}_i(\psi) = \frac{1}{ns} \sum_{s=1}^{ns} k_i^s(\psi) \)

\( Z_i \) represents a vector of instruments and \( K_i \) represents a vector of data moments. \( \tilde{K}_i(\psi) \) represents a vector of moments simulated by the model given a set of parameters \( \psi \). \( ns \) is the number of simulation.\(^{17} \) \( k_i^s(\psi) \) represents a vector of moments given \( \psi \) derived per simulation.

The moments choice is guided by the identification argument in Section 5.2. In particular, I use the proportion of single owners and partners and the first and second moment of the conditional income. I also include the proportion of zero or a negative income single owners and that of zero or a negative income partners. Net worth, education and experience are used as instruments.\(^{18} \) I choose \( \hat{\Sigma} \) to be a diagonal matrix which contains variances of the data moments. The number of parameters and the number of moments are 19 and 32 respectively.

Log value of net worth is required for estimation and I replace negative net worth as $1. Before estimation, I normalize one unit of net worth and income as $10,000 in 2011.\(^{19} \) The gross risk-free interest rate, \( r \), is assumed to be 1.1 following the literature (e.g., Evans and Jovanovic (1989); Xu (1998); Paulson et al. (2006)). The standard errors are calculated following Gourieroux and Monfort (1996).

5.5 Model Fit

Figure 6 compares the proportion of single owners and partners to the simulated proportion of single owners and partners. The density of data and simulated incomes are shown in Figures

\(^{17}\) Set \( ns = 3 \) for estimation.

\(^{18}\) Since net worth is an important state variable in the model, I use an interaction between net worth and the other covariates as additional instruments.

\(^{19}\) Evans and Jovanovic (1989) used $1,000 in 1976 as one unit.
The model predicts both ownership choice and the distribution of conditional income, especially the former, quite well.

The full set of simulated and targeted moments are shown in Tables 14 and 15 in the Appendix. All moments are matched quite well. Some predicted moments related to single owner income are a little bit higher than actual moments. First investigated by Hamilton (2000), a low earning for business owners has been explained by a non-pecuniary benefit (Hamilton (2000); Pugsley (2013)). Abstracting from such a benefit makes it difficult for the current model to perfectly match the earning differential, especially for single owners.

6 Results

6.1 Estimates

Table 7 shows the estimates for log worker productivity (log $\theta_w$) and log solo productivity (log $\theta_s$). It shows that a 10% increase in the years of education leads to approximately a 4% increase in worker productivity and 3.3% increase in solo productivity. Likewise, a 10% increase in the years of experience leads to approximately a 1% increase in worker productivity and 0.3% increase in solo productivity. The education and experience affect both worker productivity and solo productivity significantly, but their impacts are greater for worker productivity. Table 7 also shows that the relationship between net worth and solo productivity is estimated insignificant.

Table 8 shows the estimates for log value of the collaborative skill. It indicates that a 10% increase in the years of education leads to roughly an additional 2.5% increase in productivity as a partner in addition to solo productivity. The coefficient for the years of experience is estimated insignificant. This implies that the individual skill for collaborating with a partner is fostered more in schools than in workplaces. The constant term is estimated -1.3923, suggesting there is a large productivity loss from working with a partner comparing to working alone.

Table 9 reports the estimates for $\{\lambda, \alpha, \kappa\}$ and the outcome shocks. The collateral constraint parameter ($\lambda$) is estimated as 2.1761, suggesting agents can invest up to about twice their net worth. $\alpha$ is estimated as 0.1958, implying that a 10% increase in investments leads to approximately a 2% increase in outputs for a given effort-level.
To understand the relationship between the solo and partnership productivity, I project a simulated log $\theta$ on a simulated log $\tilde{\theta}$. The regression coefficient for log $\theta$ is 1.22 and its $p$-value is 0.000. This suggests that agents with better solo productivity tend to have better productivity as a partner. However, the collaborative skill, $g$, is estimated so low on average that only a small portion (3.4%) of the agents have the productivity as a partner greater than their solo productivity. This is shown in Figure 10.

Note that I explicitly control for the cost of moral hazard in business partnerships. The cost of moral hazard is not enough to explain the fact that only about 15% of agents choose partnerships conditional on business ownership. This supports the view that the cost of knowledge transfer outweighs the benefit for most individuals. The finding also supports the view that there is an additional productivity loss from “collective decision making” other than the loss from moral hazard in business partnerships. Hansmann (1996) points out that the partnerships are involved when important decisions are made by owners, and this decision process can be costly due to heterogenous interests among the owners.20

6.2 Decomposing Aggregate Gains

With the estimates of structural parameters, I use Equation (2) to measure the benefits and the costs of the partnerships. More specifically, I use the following metric:

$$
\mathbb{E}_{(\eta,u)} \left[ \sum_{i=1}^{n} \{ V_{pi}(\theta_{pi}, A_{pi}) - V_{si}(\theta_{i}, A_{i}) \} I_i(d = 3) \right]
$$

$$
= \mathbb{E}_{(\eta,u)} \left[ \sum_{i=1}^{n} \Omega_{1i} I_i(d = 3) \right] + \mathbb{E}_{(\eta,u)} \left[ \sum_{i=1}^{n} \Omega_{2i} I_i(d = 3) \right] + \mathbb{E}_{(\eta,u)} \left[ \sum_{i=1}^{n} \Omega_{3i} I_i(d = 3) \right]
$$

Equation (6) decomposes the benefits and the costs of business partnerships for those who choose to be partners. For each simulation, I calculate $\Omega_1$ (gains from productivity), $\Omega_2$ (gains from financing) and $\Omega_3$ (losses from moral hazard) in Equation (2) for every partner and then sum them up. After $N$ simulation, I average the aggregate value of $\Omega_1$, $\Omega_2$ and $\Omega_3$ for partners.

20Merges between firms are similar to forming business partnerships in that the mergers are often thought to be driven by the gains in productivity. The low estimated $g$ is also in line with the fact that most mergers – even conditional on being merged – are not successful, and often being divested. (see e.g., Banal-Estañol and Seldeslachts (2011) and references therein.)
A graphical representation is shown in Figure 11. Most aggregate gains — before subtracting the losses from moral hazard — is explained by gains from the partnership productivity: 85% of the aggregate gains are explained by the partnership productivity. The remaining 15% is explained by the gains from financing.\textsuperscript{21}

To investigate the variation between gains from two sources across wealth distribution, I conduct the same exercise for those whose net wealth is below and above the 20th percentile of wealth distribution. The result is reported in Figure 12. Most gains from financing are generated by the low wealth group and it is the major benefit for them: 80% of the aggregate gains for the low wealth group is explained by financing. In contrast, most gains from partnership productivity are generated by the high wealth group. For them, almost all of the gains are generated by the partnership productivity.

Note that a partnership is formed only if there exists the complementarity between two partners’ solo productivities or the complementarity between one partner’s solo productivity and the other partner’s wealth. Looking at Figure 11, we cannot tell whether the large gains in productivity are driven by a pure productivity motive between partners or by the case in which wealthy but unproductive partners are matched with poor but productive partners due to a financial motive.

Figure 13 depicts the decomposition in Equation (6) with respect to two complementarities. The gains generated by partnerships solely driven by the complementarity between two partners’ solo productivities explain 73% of the gains from productivity and 60% of the entire gain from partnerships. This is primarily because 63% of partnerships are solely driven by the complementarity between two partners’ solo productivities. The remaining 37% of partnerships are involved with the complementarity between one partner’s productivity and the other partner’s wealth. Only a small portion (less than 1%) choose a partnership solely due to the complementarity between one partner’s solo productivity and the other partner’s wealth.

The key feature in the data driving this result is the following: relatively more partners

\textsuperscript{21}The gains from financing were calculated conditional on partnership productivity. I could have instead calculated it conditional on solo productivity. The difference between two decompositions matters in the following situation. Suppose an agent is not financially constrained as a single owner with his solo productivity but financially constrained as a single owner with the partnership productivity. Moreover, his partner’s net worth is greater than his. With the current decomposition, the gains of the agent from the partnership is captured both by the productivity gains and the financing gains. In contrast, all the gains from the partnership is attributed to the productivity gains with the alternative decomposition. As a result, the gains from financing with the alternative decomposition is smaller than the current one.
are observed in the upper percentile than in the lower percentile of wealth distribution. For example, 62% of partners are located above the median of wealth distribution and 40% of partners are located above the 70th percentile of wealth distribution. The relationship between solo productivity and wealth is estimated insignificant, and therefore, most of the financially constrained agents are concentrated in the lower percentile of wealth distribution. The fact that we observe more partners in the upper part of wealth distribution implies that there must be partnerships formed by two relatively wealthy agents. Gains generated by them explain major aggregate gains from business partnerships.

### 6.3 Financial Friction vs. Partnership Friction

Evans and Jovanovic (1989) do not allow an option to form business partnerships. This may over-emphasize the negative impact of financial friction on the transition into business ownership. To examine this issue, I simulate four economies and compare the number of workers who would have become business owners without financial friction. I call these workers “constrained workers.” I first start with an economy without financial friction and without the choice to form partnerships (Economy A). This economy corresponds to Evans and Jovanovic (1989) without financial friction. I then impose financial friction as in Evans and Jovanovic (1989) (Economy B). In Economy C, I allow agents to form partnerships but with their productivity as a partner being the same as their solo productivity. I also assume no moral hazard. Any partnership in this economy is solely driven by the complementarity between one partner’s solo productivity and the other partner’s wealth. Finally, I allow the productivity as a partner to be different from the solo productivity and also impose the cost of moral hazard (Benchmark). Note that the benchmark economy corresponds to the estimated economy.

The result is summarized in Figure 14. About 12% of potential business owners are constrained to become a business owner due to financial friction (Economy B). If I allow the constrained workers to find a business partner without collaborative skill and moral hazard, most constrained workers enter entrepreneurship via partnerships (Economy C). The result indicates that there exists a large potential gain in financing through partnerships. However, if I allow the heterogeneous collaborative skill and moral hazard, most constrained workers who entered business ownership in Economy B once again become discouraged from starting
a business (Benchmark). This is because only a small portion of the agents is assigned the productivity as a partner greater than their solo productivity as shown in Figure 10. The cost of moral hazard is also high as shown in the next section. Overall, despite the potential gains from financing, the constrained workers do not choose to form partnerships due to the cost of moral hazard and the potential productivity loss associated with working with a partner.

6.4 Assessing the Cost of Moral Hazard

In this section, I assess the cost of moral hazard in the formation of business partnerships. I first quantify the cost of moral hazard among partners by using Equation (2). Its graphical representation is shown in Figure 11 and 12. The cost of moral hazard among partners is estimated as 39% of the aggregate gains from partnerships (Figure 11). As shown in Figure 12, the cost is less for the low wealth group (29%) than for the high wealth group (41%).

The current framework allows me not only to evaluate the cost of moral hazard for those who actually choose partnerships, but also to examine how many potential partners are discouraged from partnerships due to moral hazard. Figure 15 plots the number of single owners and partners before and after moral hazard is introduced. The cost of moral hazard is large even at the extensive margin: 57% of partners are discouraged from partnerships due to moral hazard.

7 Welfare Analysis

This section analyzes the welfare implication of financial friction. Unlike most of the previous models with financial friction, the current model features an additional distortion due to financial friction: the efficiency loss from mismatch. An intuition is addressed in Section 4.2. Because of financial friction, a productive agent with low-wealth can be matched with a less productive agent than his partner with whom he would have been matched if there were no financial friction.

Table 10A decomposes the welfare losses for each transition group. I first normalize the

---

22 The cost of moral hazard increases as the partnership productivity increases. Some of the low wealth partners sacrifice the gains from productivity for the gains from financing. As a result, the cost of moral hazard is less for the low wealth group than for the high wealth group.
aggregate welfare losses due to financial friction as 100: $\sum_{i=1}^{N}(V_i - V_i^f)\mathbb{I}_{\text{Loss}} = 100$. $V_i^f$ and $V_i$ represent the value of agent $i$ with and without financial friction respectively. $\mathbb{I}_{\text{Loss}}$ represents the indicator function for $V_i - V_i^f$ being positive. Most of the welfare losses come from the single owners who are constrained to become business owners. This accounts for 85.24% of the aggregate welfare losses due to financial friction. The second biggest source of welfare losses is from single owners who remain as single owners but with constrained borrowing. The welfare loss from this group accounts for 11.48% of the aggregate welfare losses due to financial friction.

Before further discussion, let me define mismatch in my model. I define the following three situations as mismatches driven by financial friction: a partner changes his ownership either to a worker or a single owner due to financial friction, a worker or a single owner become a partner due to financial friction, and a partner changes his partner in response to financial friction.

Table 10A shows that the welfare losses from mismatch account for 3.28% of the aggregate welfare losses due to financial friction. Of those 3.28%, single owners who become a partner for a financial motive explain 0.55%. Partners who become a worker or a single owner explain 0.3%. The remaining 2.43% is due to partners who remain as a partner, but are matched with a less productive agent than the partner with whom he would have been matched if there was no financial friction.

The financial friction generates welfare gains for a certain group of people. Table 10B decomposes the welfare gains for each transition group. Workers who become partners benefit from financial friction. The gains for them corresponds to 0.22% of the aggregate welfare losses due to financial friction. There also exist partners and a small portion of single owners who benefit from financial friction. The welfare gains for these business owners correspond to 2.07% of the aggregate welfare losses due to financial friction. Again, an intuition can be found in Section 4.2. Some partners, workers, and single owners with high-wealth can have an opportunity to meet with a more productive agent than the partner with whom he would have been matched if there was no financial friction.

In Table 11, I further investigate welfare losses and gains due to mismatch with respect to wealth level. Most welfare losses due to mismatch are generated by individuals whose net
worth is less than the 20th percentile of wealth distribution. It accounts for 80.23% of the entire welfare loss from mismatch. This is a sharp contrast to the fact that only 3.07% of the gains from mismatch are generated by the low-wealth individuals.

8 Policy Experiments

Financial friction is often a concern for policy makers who try to boost entrepreneurship. The model developed here can be used to evaluate the impact of loan programs not only on startup owners in general but also on the formation of business partnerships.\(^{23}\)

8.1 Policy 1: Loans for All Startup Firms

As a way of reducing inefficiency from financial friction, I experiment with a policy in which a government lends startup businesses up to $25,000 at the market interest rate \(r\). I assign $25,000 per business regardless of the number of owners.

Given that the interest rate is the same as the market interest rate, only financially constrained startup owners will apply to the loan program as long as there exits a small amount of costs associated with the application. The marginal benefit of additional investment is strictly greater than the market interest rate for financially constrained agents, and it is only those agents who apply to the loan program.

The left column of Table 12 shows the welfare gain by Policy 1 for each transition group. About 97.6% of the aggregate welfare losses from financial friction disappear. Workers who could have become single owners without financial friction enter business ownership as a single owner thanks to the government loans. The number of single owners increases more than 10% after the loan program is applied (Table 13). Most financially constrained single owners also benefit from the loan program and increase investments.

Policy 1 also alleviates mismatch by partners. In doing so, some partners — who can be matched with a more productive partner due to financial friction — are worse off by the loan policy. The value of their wealth in the presence of financial friction reduces once other partners can access the loans provided by the government. Despite the gains from alleviating

\(^{23}\)I simulated the economy with each policy 1,000 times and averaged the results from all simulations.
mismatch, the aggregate number of partners is almost identical before and after the loan program is applied (Table 13).

8.2 Policy 2: Loans Targeted for Business Partnerships

I experiment with a loan policy in which a government lends only business partnerships up to $25,000 at the market interest rate $r$. As in Policy 1, only financially constrained partnership firms apply to the program.

The Small Business Technology Transfer (STTR) program sponsored by the U.S government provides funds only for partnerships between small businesses and nonprofit research institutions. The goal is to promote commercialization of innovations in science by encouraging the public/private sector partnership. Although the targeted firms by the STTR program may be quite different from the firms in the sample for this paper, Policy 2 can be helpful to understand an equilibrium impact of the STTR program.

The number of partners remains almost identical before and after the policy is implemented (Table 13). Moreover, although the program specifically targets business partnerships, the welfare gains generated by improving match quality is smaller than Policy 1: the gains from alleviating mismatch by Policy 2 are about 43% of the gains from alleviating mismatch by Policy 1. The key difference is that under Policy 2, partners who would have become single owners without financial friction still remain as partners. In contrast under Policy 1, these partners switch their ownership choice to single ownership and generate welfare gains.

8.3 Discussion

The loan program targeted for general businesses is more effective than the loan program targeted only for business partnerships. The loan program for general businesses increases welfare for workers who would have been single owners without financial friction and for single owners who are constrained to borrow. Moreover, it improves the match quality more than the loan program specifically targeted for business partnerships does.

One caveat of the policy analyses in this section is that I did not take into account applicant’s default decision. Endogenizing agents’ default decision can be valuable for analyzing small business loan programs, but is outside the scope of this paper.
9 Conclusion

I empirically investigate the formation of partnerships among startup owners. I first document that partners earn significantly more than single owners. In addition, relatively more partners are observed in the lowest quintile, than in the middle, of wealth distribution. Despite the high earnings among observed partners, the proportion of partners among startup owners is much smaller than single owners. To further investigate the decision to form a partnership, I develop and estimate a matching model based on Evans and Jovanovic (1989).

Using the estimated model, I first show that the gains from the increase in productivity explain most gains from the observed partnerships. The gains from financing are the major benefit to individuals with a low net worth. I also examine how much the option to form a business partnership can help constrained workers, who could have become business owners without financial friction, to start a business. Despite potential gains from financing, most of the constrained workers do not choose to form a partnership. First of all, the cost of moral hazard is high both at the intensive and at the extensive margin. Moreover, most individuals are estimated as being less productive when working with a partner than working alone. Financial friction generates inefficiency due to mismatch among partners. A government sponsored funding program for all startup firms is effective both in helping financially constrained startup owners to create businesses and in improving the match quality among partners.

This study focuses on the role of partnerships on the business formations. A limitation is that I only use the first year business earnings as a measure of the business’ outcomes. Although the first year’s earning is highly predictive of future earnings from the business (e.g., Hamilton (2000)), it cannot fully capture the dynamics of businesses, especially their failure rate. Many new businesses fail within a few years. A further investigation into whether and how partnerships are related to business dissolution would shed light not only on business partnerships, but also on business failures.

References


**Tables and Figures**

<table>
<thead>
<tr>
<th>Table 1: Summary Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Workers</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Obs.</td>
</tr>
<tr>
<td>Experience (Year)</td>
</tr>
<tr>
<td>Education (Year)</td>
</tr>
<tr>
<td>Race (White)(%)</td>
</tr>
<tr>
<td>Married ( %)</td>
</tr>
</tbody>
</table>

NOTE: This table reports the summary statistics for characteristics of workers, single owners, and partners. Mean is reported unless otherwise indicated. The survey weight is applied.
Table 2: Income

<table>
<thead>
<tr>
<th></th>
<th>Workers</th>
<th>Single owners</th>
<th>Partners</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obs.</td>
<td>47,304</td>
<td>862</td>
<td>150</td>
</tr>
<tr>
<td>Mean</td>
<td>51,858</td>
<td>37,120</td>
<td>55,466</td>
</tr>
<tr>
<td>10%</td>
<td>15,032</td>
<td>2,017</td>
<td>3,149</td>
</tr>
<tr>
<td>50%</td>
<td>40,883</td>
<td>18,561</td>
<td>33,631</td>
</tr>
<tr>
<td>90%</td>
<td>95,612</td>
<td>84,078</td>
<td>114,315</td>
</tr>
</tbody>
</table>

NOTE: Dollar in 2011. This table shows the summary statistics for incomes of workers, single owners, and partners given that they reported a positive income.

Table 3: Income Regression for Business Owners

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>log Income</td>
<td>log Income</td>
<td>log Income</td>
<td>log Income</td>
</tr>
<tr>
<td>Partnership</td>
<td>0.584***</td>
<td>0.553***</td>
<td>0.540***</td>
<td>0.513***</td>
</tr>
<tr>
<td></td>
<td>(0.139)</td>
<td>(0.137)</td>
<td>(0.138)</td>
<td>(0.140)</td>
</tr>
<tr>
<td>Education</td>
<td>0.0527***</td>
<td>0.00200</td>
<td>-0.000945</td>
<td>0.00336</td>
</tr>
<tr>
<td></td>
<td>(0.0188)</td>
<td>(0.0204)</td>
<td>(0.0220)</td>
<td>(0.0218)</td>
</tr>
<tr>
<td>Experience</td>
<td>0.0267</td>
<td>0.0124</td>
<td>0.0136</td>
<td>0.0104</td>
</tr>
<tr>
<td></td>
<td>(0.0165)</td>
<td>(0.0165)</td>
<td>(0.0164)</td>
<td>(0.0167)</td>
</tr>
<tr>
<td>Experience²</td>
<td>-0.000577</td>
<td>-0.000336</td>
<td>-0.000372</td>
<td>-0.000314</td>
</tr>
<tr>
<td></td>
<td>(0.000377)</td>
<td>(0.000376)</td>
<td>(0.000376)</td>
<td>(0.000378)</td>
</tr>
<tr>
<td>Previous income</td>
<td>0.0529***</td>
<td>0.0503***</td>
<td>0.0495***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00995)</td>
<td>(0.0101)</td>
<td>(0.0101)</td>
<td></td>
</tr>
<tr>
<td>Net worth</td>
<td>0.00269*</td>
<td>0.00270*</td>
<td>0.00308**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00141)</td>
<td>(0.00141)</td>
<td>(0.00141)</td>
<td></td>
</tr>
<tr>
<td>Industry dummy</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Other dummies</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>1,012</td>
<td>1,012</td>
<td>1,012</td>
<td>1,012</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.027</td>
<td>0.060</td>
<td>0.086</td>
<td>0.104</td>
</tr>
</tbody>
</table>

NOTE: This table shows the estimates for regression of log incomes on partnership dummy, the years of education, the years of experience and its square, previous income, net worth and industry and other dummies. Partnership means the partnership dummy. Previous income indicates the wage income in the base year. One unit is ten thousand dollars in 2011. Other dummies include race, marital status and year dummies. Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Table 4: Net Worth

<table>
<thead>
<tr>
<th></th>
<th>Workers</th>
<th>Single owners</th>
<th>Partners</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obs.</td>
<td>47,304</td>
<td>941</td>
<td>162</td>
</tr>
<tr>
<td>Mean</td>
<td>180,582</td>
<td>208,958</td>
<td>273,226</td>
</tr>
<tr>
<td>10%</td>
<td>-5,501</td>
<td>-7,815</td>
<td>-5,626</td>
</tr>
<tr>
<td>50%</td>
<td>74,954</td>
<td>78,056</td>
<td>123,865</td>
</tr>
<tr>
<td>90%</td>
<td>493,217</td>
<td>598,411</td>
<td>692,040</td>
</tr>
</tbody>
</table>

NOTE: Dollar in 2011. This table shows the summary statistics for net worth of workers, single owners, and partners.

Table 5: Probit Analysis of Partnership Choice

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>Partnerships</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q0-20</td>
<td>0.356***</td>
</tr>
<tr>
<td></td>
<td>(0.149)</td>
</tr>
<tr>
<td>Q50-100</td>
<td>0.471***</td>
</tr>
<tr>
<td></td>
<td>(0.128)</td>
</tr>
<tr>
<td>Education</td>
<td>-0.0265</td>
</tr>
<tr>
<td></td>
<td>(0.0205)</td>
</tr>
<tr>
<td>Experience</td>
<td>-0.0118**</td>
</tr>
<tr>
<td></td>
<td>(0.0048)</td>
</tr>
<tr>
<td>Previous income</td>
<td>-0.0085</td>
</tr>
<tr>
<td></td>
<td>(0.0105)</td>
</tr>
<tr>
<td>Observations</td>
<td>1,103</td>
</tr>
<tr>
<td>Pseudo R-squared</td>
<td>0.061</td>
</tr>
</tbody>
</table>

NOTE: This tables reports the estimates for Probit regression of the partnership dummy on wealth dummies, the years of education, the years of experience, and previous incomes among business owners. One unit is ten thousand dollars in 2011. Q“x” represents “x”th percentile of wealth distribution of the entire population. Q20: $ 3,998, Q50: $ 75,162. Other dummies include race, marital status and year dummies. Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Table 6: Industry Decomposition

<table>
<thead>
<tr>
<th>Industry</th>
<th>Single owners</th>
<th>Partners</th>
<th>Partners (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture, forestry, fishing, and hunting</td>
<td>53</td>
<td>16</td>
<td>23</td>
</tr>
<tr>
<td>Mining</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Construction</td>
<td>252</td>
<td>40</td>
<td>14</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>33</td>
<td>6</td>
<td>15</td>
</tr>
<tr>
<td>Wholesale trade</td>
<td>19</td>
<td>6</td>
<td>24</td>
</tr>
<tr>
<td>Retail trade</td>
<td>78</td>
<td>21</td>
<td>21</td>
</tr>
<tr>
<td>Transportation, warehousing, and utilities</td>
<td>51</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Information</td>
<td>11</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>Finance, insurance, real estate, and rental and leasing</td>
<td>39</td>
<td>14</td>
<td>26</td>
</tr>
<tr>
<td>Professional, scientific, management, administrative, and waste management</td>
<td>213</td>
<td>27</td>
<td>11</td>
</tr>
<tr>
<td>Services</td>
<td>191</td>
<td>28</td>
<td>13</td>
</tr>
<tr>
<td>Total</td>
<td>941</td>
<td>162</td>
<td>15</td>
</tr>
</tbody>
</table>

NOTE: This table reports the number and the proportion of partners for different industries. Services include Business and repair services; Personal services; Educational, health and social service; Arts, entertainment, recreation, accommodations, and food service; Public administration; and other services.

Table 7: Estimates for log Worker Productivity and log Solo Productivity

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Variables</th>
<th>Estimates</th>
<th>Standard Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_1$</td>
<td>log(Education)</td>
<td>0.3947</td>
<td>(0.0195)</td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>log(Experience)</td>
<td>0.0947</td>
<td>(0.0045)</td>
</tr>
<tr>
<td>$\beta_0$</td>
<td>Constant</td>
<td>0.3104</td>
<td>(0.0620)</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>log(Net worth)</td>
<td>0.0024</td>
<td>(0.0109)</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>log(Education)</td>
<td>0.3245</td>
<td>(0.0887)</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>log(Experience)</td>
<td>0.0359</td>
<td>(0.0188)</td>
</tr>
<tr>
<td>$\sigma_\eta$</td>
<td>Std. of $\eta$</td>
<td>0.1068</td>
<td>(0.0027)</td>
</tr>
</tbody>
</table>

NOTE: This table presents the estimates for the equation: $\log \theta_w = \gamma_1 \log x_1 + \gamma_2 \log x_2$ and $\log \theta = \beta_0 + \beta_1 \log A + \beta_2 \log x_1 + \beta_3 \log x_2 + \eta$, where $\eta \sim N(0, \sigma_\eta^2)$. $A$ is net worth (ten thousands dollars in 2011). $x_1$ is the years of education. $x_2$ is the years of experience. Asymptotic standard errors are in parentheses.
Table 8: Estimates for Matching Function

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Variables</th>
<th>Estimates</th>
<th>Standard Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>(g_0)</td>
<td>Constant</td>
<td>-1.3923</td>
<td>(0.0867)</td>
</tr>
<tr>
<td>(g_1)</td>
<td>log(Education)</td>
<td>0.2487</td>
<td>(0.1171)</td>
</tr>
<tr>
<td>(g_2)</td>
<td>log(Experience)</td>
<td>0.0191</td>
<td>(0.0341)</td>
</tr>
<tr>
<td>(\sigma_u)</td>
<td>Std. of (u)</td>
<td>0.3735</td>
<td>(0.0098)</td>
</tr>
</tbody>
</table>

NOTE: This table presents the estimates for the equation: \(\log g = g_0 + g_1 \log x_1 + g_2 \log x_2 + u, \quad u \sim N(0, \sigma_u^2).\) \(g\) represents the additional ability as a partner. \(x_1\) is the years of education. \(x_2\) is the years of experience. Asymptotic standard errors are in parentheses.

Table 9: Estimates for Other Parameters

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Variables</th>
<th>Estimates</th>
<th>Standard Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\lambda)</td>
<td>Collateral constraint</td>
<td>2.1761</td>
<td>(0.9423)</td>
</tr>
<tr>
<td>(\alpha)</td>
<td>Technology</td>
<td>0.1958</td>
<td>(0.0297)</td>
</tr>
<tr>
<td>(\kappa)</td>
<td>Preference</td>
<td>1.7673</td>
<td>(0.0220)</td>
</tr>
<tr>
<td>(\sigma_w)</td>
<td>Std. of (\log \epsilon_w)</td>
<td>0.7289</td>
<td>(0.0072)</td>
</tr>
<tr>
<td>(\sigma_s)</td>
<td>Std. of (\log \tilde{\epsilon}_s)</td>
<td>1.6143</td>
<td>(0.0240)</td>
</tr>
<tr>
<td>(P_s)</td>
<td></td>
<td>0.0386</td>
<td>(0.0072)</td>
</tr>
<tr>
<td>(\sigma_p)</td>
<td>Std. of (\log \tilde{\epsilon}_p)</td>
<td>1.0245</td>
<td>(0.1111)</td>
</tr>
<tr>
<td>(P_p)</td>
<td></td>
<td>0.1505</td>
<td>(0.0387)</td>
</tr>
</tbody>
</table>

NOTE: This table presents the estimates for \(\{\lambda, \alpha, \kappa\}\) and the outcome shocks. \(\lambda\) captures the extent of collateral constraint. \(\alpha\) governs the marginal productivity of capital. \(\kappa\) captures the marginal rate of substitution between consumption and effort. The outcome shock to workers (\(\epsilon_w\)) is modelled as \(\log \epsilon_w \sim N(\mu_w, \sigma_w^2), \quad \mathbb{E}[\epsilon_w] = 1.\) The outcome shock to single owners (\(\epsilon_s\)) is modelled as \(\epsilon_s = \tilde{\epsilon}_s - P_s, \quad \log \tilde{\epsilon}_s \sim N(\mu_s, \sigma_s^2), \quad \mathbb{E}[\epsilon_s] = 1.\) The outcome shock to partners (\(\epsilon_p\)) is modelled as \(\epsilon_p = \tilde{\epsilon}_p - P_p, \quad \log \tilde{\epsilon}_p \sim N(\mu_p, \sigma_p^2), \quad \mathbb{E}[\epsilon_p] = 1.\) \(P_s\) and \(P_p\) are some positive constants. Asymptotic standard errors are in parentheses.
Table 10: Decomposing Welfare Losses and Gains due to Financial Friction by Each Transition Group (The numbers are normalized by the aggregate welfare losses due to financial friction)

<table>
<thead>
<tr>
<th></th>
<th>(A) Welfare loss</th>
<th>(B) Welfare gain</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Workers</td>
<td>Single owners</td>
</tr>
<tr>
<td>Workers</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Single owners</td>
<td>85.24</td>
<td>11.48</td>
</tr>
<tr>
<td>Partners</td>
<td>0.30</td>
<td>0.003</td>
</tr>
</tbody>
</table>

NOTE: Column A of this table presents welfare losses with respect to each transition group after financial friction is introduced. Column B of this table presents welfare gains for each transition group due to financial friction. The aggregate welfare losses from financial friction is normalized as 100: \( \sum_{i=1}^{N} (V_i - V_i') I_{loss} = 100. \) \( V_i' \) and \( V_i \) represent the value of agent \( i \) with and without financial friction respectively. \( I_{loss} \) represents the indicator function for \( V_i - V_i' \) being positive.

Table 11: Welfare Losses and Gains due to Mismatch by Low-wealth Individuals (The numbers are normalized by the aggregate welfare losses due to financial friction)

<table>
<thead>
<tr>
<th></th>
<th>Welfare loss</th>
<th>Welfare gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>All agents</td>
<td>3.28</td>
<td>2.29</td>
</tr>
<tr>
<td>Agents &lt; 20th</td>
<td>2.63</td>
<td>0.07</td>
</tr>
<tr>
<td>( (b/a) \times 100 )</td>
<td>(80.28%)</td>
<td>(3.07%)</td>
</tr>
</tbody>
</table>

NOTE: This table presents losses and gains due to mismatch by low-wealth individuals. Agents < 20th indicate individuals whose net worth is less than the 20th percentile of wealth distribution. The aggregate welfare losses due to financial friction is normalized as 100: \( \sum_{i=1}^{N} (V_i - V_i') I_{loss} = 100. \) \( V_i' \) and \( V_i \) represent the value of agent \( i \) with and without financial friction respectively. \( I_{loss} \) represents the indicator function for \( V_i - V_i' \) being positive.
Table 12: Decomposing Welfare Gains for Policy 1 & Policy 2 by Each Transition Group (The numbers are normalized by the aggregate welfare losses due to financial friction)

<table>
<thead>
<tr>
<th></th>
<th>After Policy 1</th>
<th>After Policy 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Workers</td>
<td>Single owners</td>
</tr>
<tr>
<td>Workers</td>
<td>0</td>
<td>85.24</td>
</tr>
<tr>
<td>Single owners</td>
<td>0</td>
<td>11.48</td>
</tr>
<tr>
<td>Partners (G+L)</td>
<td>-0.22</td>
<td>0.53</td>
</tr>
<tr>
<td>Group (L)</td>
<td>-0.22</td>
<td>-0.02</td>
</tr>
<tr>
<td>Group (G)</td>
<td>0</td>
<td>0.55</td>
</tr>
</tbody>
</table>

NOTE: The table presents welfare gains generated by Policy 1 and Policy 2 respectively. Policy 1 is a loan program for all businesses with the market interest rate and the maximum lending amount of $25,000. Policy 2 is a loan program targeted for partnership firms with the market interest rate and the maximum lending amount of $25,000. Group (L) indicates ex-partners who incur a loss due to each policy. Group (G) indicates ex-partners who benefit from each policy. A minus figure indicates the welfare losses from each policy. To highlight the extent of welfare improvement by policies, I normalize the gains in both polices with respect to the aggregate welfare losses due to financial friction, which is normalized as 100: $\sum_{i=1}^{N}(V_i - V_i^f)I_{\text{Loss}} = 100$. $V_i^f$ and $V_i$ represent the value of agent $i$ with and without financial friction respectively. $I_{\text{Loss}}$ represents the indicator function for $V_i - V_i^f$ being positive.

Table 13: Number of Business Owners After Policy 1 & Policy 2

<table>
<thead>
<tr>
<th></th>
<th>Single owners (a)</th>
<th>Partners (b)</th>
<th>Business owners (a+b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benchmark</td>
<td>82.04</td>
<td>17.96</td>
<td>100</td>
</tr>
<tr>
<td>Policy 1</td>
<td>93.71</td>
<td>17.92</td>
<td>111.63</td>
</tr>
<tr>
<td>Policy 2</td>
<td>82.04</td>
<td>17.98</td>
<td>100.02</td>
</tr>
</tbody>
</table>

NOTE: This table compares the number of business owners after Policy 1 and Policy 2. Policy 1 is a loan program for all businesses with the market interest rate and the maximum lending amount of $25,000. Policy 2 is a loan program targeted for partnership firms with the market interest rate and the maximum lending amount of $25,000. The benchmark economy is the estimated economy. The number of business owners in the benchmark economy is normalized as 100.
Figure 1: Equity Share for Startup Owners

NOTE: This figure shows the histogram of equity shares for those whose equity share is greater than 1% and less than 100%.

Figure 2: Kernel Density of Log Income

NOTE: This figure shows the kernel density of log income for different groups (Single owners, Partners, Workers).
Figure 3: Kernel Density of Net Worth

Figure 4: Proportion of Partners among Business Owners with respect to Net Worth

NOTE: This figure depicts the proportion of partners among startup owners with respect to net worth. Q“x” represents “x”th percentile of wealth distribution of the entire population. Q20: $3,998, Q50: $75,162
Figure 5: Proportion of Partnerships across Industries

NOTE: This figure shows the proportion of partnerships for high and low starting capital industries. High-starting capital industry includes Mining; Manufacturing; Wholesale and Retail trade; Transportation, warehousing, and utilities; and Finance, insurance, real estate, and rental and leasing. Low-starting capital industry includes Construction and Services. $p$-value of difference is 0.028.
Figure 6: Proportion of Business Owners (Data vs. Model)

NOTE: This figure compares the proportion of single owners and partners to the simulated proportion of single owners and partners.

Figure 7: Kernel Density of Income for Workers (Data vs. Model)

NOTE: This figure compares Kernel density of income for workers and a simulated income for workers given covariates and estimates.
Figure 8: Kernel Density of Income for Single Owners (Data vs. Model)

NOTE: This figure compares Kernel density of income for single owners and a simulated income for single owners given covariates and estimates.

Figure 9: Kernel Density of Income for Partners (Data vs. Model)

NOTE: This figure compares Kernel density of income for partners and a simulated income for partners given covariates and estimates.
Figure 10: Kernel Density of Collaborative Skill

NOTE: This figure shows Kernel density of a simulated value for $g$ given covariates and estimates. $g$ is greater than one means that the productivity as a partner is greater than the solo productivity and $g$ is less than one means that the productivity as a partner is less than the solo productivity.

Figure 11: Decomposing Benefits and Costs of Partnerships 1

NOTE: This figure depicts the decomposition of benefits and costs from partnerships for all partners. The aggregate gains for all partners are normalized as one.
NOTE: This figure depicts the decomposition of benefits and costs from partnerships below and above 20th percentile of wealth distribution (Blue: Partners Below 20th percentile of wealth distribution, Sky-blue: Partners above 20th percentile of wealth distribution). The aggregate gains for each group are normalized as one.
Figure 13: Decomposing Benefits and Costs of Partnerships

NOTE: This figure depicts the decomposition of benefits and costs from partnerships with respect to complementarities (Blue: All partners, Sky-blue: partnerships solely driven by complementarity between two partners’ solo productivities, Green: partnerships driven by complementarity between one partner’s solo productivities and the other partner’s wealth). The aggregate gains for all partners are normalized as one.
NOTE: This figure illustrates the number of single owners, of partners, and of constrained workers in each counterfactual simulation. The constrained worker is a worker who would have become a business owner if there were no financial friction. The Economy A is the economy without financial friction and without an option to form business partnerships. The Economy B is the economy with financial friction but without an option to form business partnerships. The Economy C is the economy in which an option to form business partnerships is introduced from Economy B with the productivity as a partner being the same as the solo productivity and without moral hazard. In this economy, a partnership is formed only if there exists the complementarity between one partner’s solo productivity and the other partner’s wealth. The Benchmark economy is identical to the Economy C except that the collaborative skill and moral hazard are introduced. Note that the benchmark economy corresponds to the estimated economy. The number of single owners in Economy A is normalized as 100.
Figure 15: Counterfactual Exercise 2

NOTE: This figure illustrates the number of single owners and partners before and after moral hazard is introduced into the benchmark economy. The number of business owners in the benchmark economy is normalized as 100.
Appendix

A Sample Construction

In this section, I describe the sample construction in more details. The SIPP was redesigned in 1996, and as a result the variable names, as well as the data editing and imputing procedures, are not consistent between panels before and after the 1996 panel. For this reason, I use panels after 1996 including 1996, 2001, 2004 and 2008. The samples between panels are not overlapping.

In SIPP, the respondents were interviewed every 4 month with questions such as income level and working hours for each month (These questions are labeled as “core modules”). In addition, broader questions ranging from household net worth to child support were asked annually (These questions are labeled as “topical modules”). In particular, for every third wave of interviewing, SIPP provides the household level net worth (THHTNW) and the share of business equity for business owners (EVBOW). Since the household net worth is one of the most important variables for this study and it is recorded annually, I construct a panel where the time unit is a year.

Type of Business Owners A respondent is defined as a business owner if he answered yes to the question “Did you own a business?” (EBIZNOW) during at least one of three previous waves (representing one year) when the household net worth is recorded.\(^{24}\) The equity share for those who have more than 1% and less than 100% equity is concentrated around 50% (Figure 1). I define a partner as a business owner whose share of business equity is greater than or equal to 25% and less than or equal to 75%.\(^{25}\) Business owners whose equity share is greater than 75% are called single owners.

Income SIPP explicitly asks how much each respondent earned for each month in every wave (TBMSUM/TPRFTB for business owners, TPMSUM for workers). In principle, one wave covers four months. I use the monthly income only for the survey month since it is well documented that there is little variation in monthly incomes within the same wave. Thus,

\(^{24}\)Some respondents in the third wave of 2004 panel were wrongly recorded as business owners (APDJBTHN=5). I drop these respondents. For more information, see SIPP user note for Business Feedback Problem.

\(^{25}\)According to this definition, an incorporated firm with multiple owners is considered a business partnership.
I use the term “wave” and “month” interchangeably. For example, suppose a respondent reported earned income for only two months and the total amount of earned income for the two months was $5,000. Then his annual earned income is calculated as $5,000 × \frac{12}{2}.

One issue regarding business income is that negative income is reported only for the panel after 2004 but not for the panel before 2004. This is due to a change in interview questions for business earnings starting in 2004. For example, some portion of earned income from a business is recorded in TPRFTB after 2004 but not before 2004.\footnote{For more detailed information, see SIPP user note for Business Income and Profit/Loss.} The way I handle zero or negative business income is found in the Final Sample construction.

Other Characteristics Other questions regarding characteristics such as age (TAGE), sex (ESEX), race (ERACE), marital status (EMS) and education (EEDUCATE) were asked in the wave after the household net worth is recorded. I use this information as the annual characteristics of each respondent. Industry information is found in the core modules both for workers (EJBIND) and for business owners (TBSIND). Typical hours worked per week is also reported in the core modules (EJBHRS: wage work hours, EHRBSBS: business work hours). The potential experience is calculated as \( \max\{\text{age} - \text{years of education} - 6, 1\} \).

The Final Sample The final sample is constructed in a similar way to the literature (e.g., Evans and Jovanovic (1989); Hamilton (2000); Hurst and Lusardi (2004)). It is a two year panel for males from ages 18 to 65. This sample is chosen to limit the influence of the labor market participation. I call the first year the base year and the second year the subsequent year. I first drop business owners in the base year. Some respondents started their business from the last wave in the base year and I re-categorized these respondents as new business owners in the subsequent year.\footnote{A handful of respondents reported as business owners only in the last wave in the base year but not in the subsequent year. I dropped these respondents.} I also drop unemployed respondents in the base and the subsequent year.\footnote{The unemployed is defined as non-business owners who reported “no job” (RMESR ∈ \{6, 7, 8\}) at least two waves during a given year. To limit measurement error, I also drop respondents who were employed but did not report the wage earnings in the subsequent year.} Those who did not report the household wealth are also dropped. 13 outliers whose net worth is greater than $10 billion are also dropped. Among startup owners in the subsequent year, I drop respondents who answered yes to the question “Was this business owned entirely by members of this household?” (IHHOWN\footnote{This variable is found in the same topical module containing the individual characteristics.}). These individuals are dropped for two reasons.
First, the motivation to form a family business may be different from the motivation to start a business with non-household members. Second, forming a partnership among household members does not increase the total value of household net worth, an important mechanism this paper investigates. I also drop 1 active duty military personnel who are categorized as business owners. Business owners who did not report the equity share or reported equity share which is less than 25% are also dropped. To remove casual business owners, I drop an individual if his business work hours are less than 30 hours and less than wage work hours. I did not exclude business owners who reported their working hours as “Time varies.” The sample consists of 48,407 individuals. 1,103 out of 48,407 became business owners in the subsequent year and among them 162 are categorized as partners. The number of observations at each stage of sample construction is summarized in Table 16.

Out of 1,103 startup owners, 161 reported zero income and 30 reported a negative income. Note that a negative income is reported in panels after 2004. During the same period, zero income is also reported (70 observations). Some business owners who reported zero income before the 2004 panel may realize a negative income. To make variables consistent across survey years, I first replace negative income to zero. Then, for startup owners with zero income, I impute business incomes to the one reported one year after the subsequent year if it is available (89 observations imputed). If not, and if the business owners started their business from the last wave in the base year, I impute business incomes to the one reported in the base year (11 observations imputed). The resulting number of business owners with zero income is 79 for single owners and 12 for partners. Finally, I impute the years of education for those who did not respond with 13, the average years of education for the whole population conditional on reporting.30

SIPP Oversampling For SIPP sample design, the Census Bureau assigned the universe of addresses into two strata, one with a higher proportion of poverty than the other. The Census Bureau select more samples in the high-poverty stratum. Most of the high-poverty regions coincide with areas in which more African-American, Hispanic, and female headed family are populated. As a result, most over-sampled low income individuals or families are African-American, Hispanic, or female headed family (Huggins and King, 1997). To assess

30The years of education for 234 workers, 9 single owners and one partner were imputed.
whether the final sample contains relatively more of these low-income individuals, I compare racial composition of the final sample to that of Consumer Expenditure Survey (CEX) in 2011.

Table 17 compares the proportion of African-Americans with respect to income both for the final sample and for CEX. The proportion of low-income African-Americans in the final sample is smaller than that of CEX. This suggests that most of the over-sampled low income individuals are more likely to be dropped in the process of sample construction.

B Identification of the Baseline Model

The estimation of the baseline model requires a specification of the partnership productivity and the potential partner’s net worth. Theories predict that the partnership productivity can be different from the solo productivity. I assume that the proportional difference in productivity is determined by each agent’s observable and unobservable characteristics. More specifically,

$$\log \theta_p - \log \theta = \hat{g}_0 + \hat{g}_1 \log A + \hat{g}_2 \log x_1 + \hat{g}_3 \log x_2 + \hat{u}, \quad \text{where} \quad \hat{u} \sim N(0, \sigma_{\hat{u}}^2)$$

Education and experience may have an additional value on a team’s work. Having more wealth may increase the opportunity to meet with a partner with a higher solo productivity. $u$ captures other unobserved components affecting partnerships productivity. I assume that the partner’s net worth to be randomly assigned from the whole population.

With the specification above, the model is fully characterized by a set of parameters: \{\kappa, \gamma_1, \gamma_2, \alpha, \beta_0, \beta_1, \beta_2, \beta_3, \sigma_\eta, \hat{g}_0, \hat{g}_1, \hat{g}_2, \hat{g}_3, \sigma_\hat{u}, \sigma_w, \sigma_s, \sigma_p, \lambda\}. 31 \psi denotes the set of all parameters. $X$ denotes \{\{A, x_1, x_2\}\}, the set of all covariates. $Y$ denotes \{\{d, \pi \cdot d\}\}, the ownership choice and income conditional on each choice.

For identification of the baseline model, I exploit the following idea. The effect of financial friction can be different across wealth distribution: agents in the lower percentile are more affected than agents in the upper percentile by the financial friction. If an agent’s net worth is large enough, he may not be financially constrained, and thus the primary reason for him to choose a partnership is because of the productivity gains. Therefore, the ownership choice and the conditional income for those wealthy agents can be informative on the joint distribution

31 For expositional convenience, I normalize $P_s$ and $P_p$ as zero.
of solo and partnership productivity. Given the joint distribution of solo and partnership productivity, the probability to become a partner for financially constrained agents is only affected by the collateral constraint $\lambda$. Thus, the ownership choice of agents in the lowest quintile of wealth distribution can be informative on $\lambda$.

To formally address the above argument, I first clarify what I mean by identification in this context. The notations and definitions are similar to Cameron and Trivedi (2005) and Dufour and Hsiao (2008). A structure, indexed by $\psi$, is defined as the joint probability distribution $\Pr(Y, X; \psi)$ where $\psi \in \Psi$ and $(Y, X) \in W = \{1, 2, 3\} \times [0, \infty) \times [0, \infty) \times [0, \infty) \times [0, \infty)$. Two structures are observational equivalent if $\Pr(Y, X; \psi^1) = \Pr(Y, X; \psi^2)$ for all $(Y, X) \in W$. $\psi$ is identified if there is no other observational equivalent structure in $\Psi$.

To identify $\psi$, I impose the following assumption.

Assumption 1.

$$\frac{1 - \alpha}{2} > \max\{\beta_1, \beta_1 + \hat{g}_1\}$$

The solo and partnership productivity can be allowed to increase positively with respect to net worth, but the rate must not be so large to identify the model. Under Assumption 1,

$$\lim_{A \to \infty} F(\eta \leq G_s\bar{X}, \eta + \hat{u} \leq G_p\bar{X}) = 1$$

(7)

where

$$\begin{bmatrix} G_s \\ G_p \end{bmatrix} = \begin{bmatrix} -\frac{(1-\alpha)}{2} \log \left( \frac{\hat{a}}{\bar{X}} \right) - \beta_0 & \frac{1-\alpha}{2} - \beta_1 & -\beta_2 & -\beta_3 \\ -\frac{(1-\alpha)}{2} \log \left( \frac{\hat{a}}{\bar{X}} \right) - \beta_0 - \hat{g}_0 & \frac{1-\alpha}{2} - \beta_1 - \hat{g}_1 & -\beta_2 - \hat{g}_2 & -\beta_3 - \hat{g}_3 \end{bmatrix}$$

and $\bar{X} = [1 \log X]'$. $F$ is a bivariate normal distribution. Equation (7) says that, for any $(x_1, x_2)$, one can find a value of $A^*$ such that the probability for agents whose net worth is greater than $A^*$ not to be financially constrained both as a single owner and as a partner is arbitrarily close to one. In other words, the measure of financially unconstrained agents both as a single owner and as a partner is one at the infinity of wealth distribution. The first part of
the identification argument is to show the ownership choice and the first moment of conditional
income of these unconstrained agents identify \( \{ \gamma_1, \gamma_2, \alpha, \beta_0, \beta_1, \beta_2, \beta_3, \sigma_n, \hat{g}_0, \hat{g}_1, \hat{g}_2, \hat{g}_3, \sigma_u, \kappa \} \) for a given \( \{ \lambda, \sigma_w, \sigma_s, \sigma_p \} \).

The ownership choice for unconstrained agents is:

\[
\begin{align*}
\text{if } u_1^* = \max\{u_1^*, u_2^*, u_3^*\}, & \quad d = 1, \\
\text{if } u_2^* = \max\{u_1^*, u_2^*, u_3^*\}, & \quad d = 2, \\
\text{if } u_3^* = \max\{u_1^*, u_2^*, u_3^*\}, & \quad d = 3,
\end{align*}
\]

(8)

\[
\begin{pmatrix} e_2 \\ e_3 \end{pmatrix} \sim N(0, \Sigma), \quad \Sigma = 
\begin{bmatrix} 1 & 1 \\ 1 & 1 + \rho^2 \end{bmatrix}
\]

where

\[
\begin{bmatrix} G_1 \\ G_2 \\ G_3 \end{bmatrix} = 
\begin{bmatrix} \frac{(1-\alpha) \log a_2}{2\sigma_n} & 0 & \frac{\hat{\gamma}_1}{\sigma_n} & \frac{\hat{\gamma}_2}{\sigma_n} \\ \frac{(1-\alpha) \log a_1}{2\sigma_n} + \frac{\beta_0}{\sigma_n} & \frac{\beta_1}{\sigma_n} & \frac{\beta_2}{\sigma_n} & \frac{\beta_3}{\sigma_n} \\ \frac{(1-\alpha) \log a_3}{2\sigma_n} + \frac{\hat{g}_0 + \hat{g}_1}{\sigma_n} & \frac{\hat{g}_1 + \hat{g}_2}{\sigma_n} & \frac{\hat{g}_2 + \hat{g}_3}{\sigma_n} \end{bmatrix},
\]

\( \rho = \frac{\sigma_u}{\sigma_n}, \quad \hat{\gamma}_1 = \frac{1-\alpha}{1+\alpha} \gamma_1, \quad \text{and} \quad \hat{\gamma}_2 = \frac{1-\alpha}{1+\alpha} \gamma_2. \) Notice that (8) is equivalent to the Multinomial probit
model and hence \( \tilde{G}_2 := G_2 - G_1, \tilde{G}_3 := G_3 - G_1 \) and \( \rho \) are identified with the same argument
by which parameters in the Multinomial probit model are identified (Bunch, 1991).

The first moment of the conditional log income in each ownership is:

\[
\begin{align*}
\mathbb{E}[\log \pi_w | d = 1, X] &= G_4 X \\
\mathbb{E}[\log \pi_s | d = 2, X] &= G_5 X + \frac{2\sigma_n}{1-\alpha} \mathbb{E}[e_2] - e_2 < \tilde{G}_2^* X, e_3 - e_2 < (\tilde{G}_2^* - \tilde{G}_3^*) X \\
\mathbb{E}[\log \pi_p | d = 3, X] &= G_6 X + \frac{2\sigma_n}{1-\alpha} \mathbb{E}[e_3] - e_3 < \tilde{G}_3^* X, e_2 - e_3 < (\tilde{G}_3^* - \tilde{G}_2^*) X
\end{align*}
\]

(9)
where

\[
\begin{bmatrix}
G_4 \\
G_5 \\
G_6
\end{bmatrix} =
\begin{bmatrix}
\log a_6 & 0 & \frac{2\gamma_1}{1+\alpha} & \frac{2\gamma_2}{1+\alpha} \\
\log(2a_1) + \frac{2\beta_0}{1-\alpha} & \frac{2\beta_1}{1-\alpha} & \frac{2\beta_2}{1-\alpha} & \frac{2\beta_3}{1-\alpha} \\
\log a_1 + \frac{2(\beta_0+\hat{g}_0)}{1-\alpha} & \frac{2(\beta_1+\hat{g}_1)}{1-\alpha} & \frac{2(\beta_2+\hat{g}_2)}{1-\alpha} & \frac{2(\beta_3+\hat{g}_3)}{1-\alpha}
\end{bmatrix}
\]

The expectation terms in (9) can be calculated by using the estimates from (8). Once the selection effect is controlled, \(G_4, G_5\) are identified by a regression. Finally, I can recover the structural parameters by using the estimates \(\tilde{G}_2^*, \tilde{G}_3^*, \tilde{G}_4^*, \tilde{G}_5^*, \tilde{G}_6^*, \) and \(\rho^*\)\(^{32}\).

Given the parameters identified above, \(\lambda\) is identified by the ownership choice of agents whose net worth is near zero where the measure of financially constrained agents is positive. For example, consider \(\Pr(d = 3|X, A \leq 0)\).

\[
\Pr(d = 3|X, A \leq 0) = \int_0^\infty \left\{ \Pr(d = 3, \hat{\theta}_p^{\frac{2}{1-\alpha}} \leq \lambda A'|X, A \leq 0, A') + \Pr(d = 3, \hat{\theta}_p^{\frac{2}{1-\alpha}} > \lambda A'|X, A \leq 0, A') \right\} f(A')dA'
\]

where \(f(A')\) is the population distribution of net worth. Given that \(A \leq 0\), the probability to become a partner for those who are assigned a negative \(A'\) is zero. The equation (10) is a function of \(\lambda\). I show that it is strictly increasing with respect to \(\lambda\) and hence one to one mapping with \(\lambda\).

**Proposition 3.** Given \(\{\gamma_1, \gamma_2, \alpha, \beta_0, \beta_1, \beta_2, \beta_3, \sigma_\eta, \hat{g}_0, \hat{g}_1, \hat{g}_2, \sigma_\xi, \kappa\}\), \(\Pr(d = 3|X, A \leq 0)\) is strictly increasing with respect to \(\lambda\).

**Proof** See Appendix C.2.

\(^{32}\)Some parameters are over-identified. See Appendix C.3.
Finally, \( \{\sigma_w, \sigma_s, \sigma_p\} \) are identified by the following moments.

\[
\text{var}(\log y_w|X, d = 1) = \mathbb{E}[(\log y_w - \mathbb{E}(\log y_w|X, d = 1))^2|X, d = 1] = \sigma_w^2
\]

\[
\text{var}(\log y_s|X, d = 2) = \mathbb{E}[(\log y_s - \mathbb{E}(\log y_s|X, d = 2))^2|X, d = 2] = \mathbb{E}[(\eta - \mathbb{E}(\eta|X, d = 2))^2|X, d = 2] + \sigma_s^2
\]

\[
\text{var}(\log y_p|X, d = 3) = \mathbb{E}[(\log y_p - \mathbb{E}(\log y_p|X, d = 3))^2|X, d = 3] = \mathbb{E}[(\eta + \hat{\eta} - \mathbb{E}(\eta + \hat{\eta}|X, d = 3))^2|X, d = 3] + \sigma_p^2
\]

Therefore, \( \psi \) is identified.

C Proofs

C.1 Proof of Proposition 1

The planner solves the following problem:

\[
\max_{\{z_i, z_j, k\}} F(z_i, z_j, k) = \mathbb{E}[\theta_p k^\alpha (z_i + z_j)^{1-\alpha} - rk\epsilon_p - \frac{K}{2}(z_i^2 + z_j^2)]
\]

\[
= \mathbb{E}[\theta_p k^\alpha (z_i + z_j)^{1-\alpha} - \underbrace{rk + \frac{K}{2}(z_i^2 + z_j^2)}_{\text{Aggregate production}}]
\]

subject to \((z_i, z_j, k) > 0, \quad \alpha \in (0, 1), \quad \kappa > 0, \quad A_p \geq 0, \quad k \leq \lambda A_p \)

(Case 1) \( \lambda A_p \geq 2\hat{\alpha} \theta_p^{\frac{2}{1-\alpha}} \)

I first consider the case in which \( k \leq \lambda A_p \) is not binding. I first show that \( F(z_i, z_j, k) \) is a concave function on \( U = \{x \in \mathbb{R}^3|(z_i, z_j, k) > 0\} \). Since \( F: U \to \mathbb{R}^1 \) is \( C^2 \) function and \( U \) is a convex open subset of \( \mathbb{R}^3 \), \( F \) is a concave function on \( U \) if and only if \( D^2 F(x) \) is negative semidefinite for all \( x \in U \). By deriving leading principal minors, one can easily show that \( D^2 F(x) \) is negative definite for all \( x \in U \) and hence negative semidefinite for all \( x \in U \). Given that \( F(z_i, z_j, k) \) is a concave function on \( U \), it is sufficient to show that there exists the unique \( x^* \in U \) satisfying \( DF(x^*) = 0 \) to guarantee that \( x^* \) is the unique global max of \( F \) on \( U \).
The first order conditions are:

\[(1 - \alpha) \theta_p \left( \frac{k}{z_i + z_j} \right)^\alpha = \kappa z_i \quad (11)\]

\[(1 - \alpha) \theta_p \left( \frac{k}{z_i + z_j} \right)^\alpha = \kappa z_j \quad (12)\]

\[\alpha \theta_p \left( \frac{k}{z_i + z_j} \right)^{\alpha - 1} = r \quad (13)\]

(11) and (12) imply \(z^* = z_i = z_j\) for any given \(k\). The first order conditions reduce

\[(1 - \alpha) \theta_p \left( \frac{k}{2z^*} \right)^\alpha = \kappa z^* \quad (14)\]

\[\alpha \theta_p \left( \frac{k}{2z^*} \right)^{\alpha - 1} = r \quad (15)\]

By putting (15) into (14), I get

\[z^* = \left( \frac{1 - \alpha}{\kappa} \right) \theta_p^{\frac{1}{1-\alpha}} \left( \frac{\alpha}{r} \right)^{\frac{\alpha}{1-\alpha}} \] and hence \(k^* = 2 \left( \frac{1 - \alpha}{\kappa} \right) \left( \frac{\alpha}{r} \right)^{\frac{\alpha}{1-\alpha}} \theta_p^{\frac{2}{1-\alpha}} = 2 \hat{a} \theta_p^{\frac{2}{1-\alpha}}\)

Therefore, I derived the unique \(x^* = (z^*, z^*, k^*)\) such that \(DF(x^*) = 0\).

Given that the aggregate production is distributed equally, the value function per parter is given

\[V_p^E(\theta_p, A_p) = \frac{1}{2} \left[ \theta_p k^{\alpha} (2z^*)^{1-\alpha} - rk^* \right] - \frac{\kappa}{2} z^*^2 = a_1 \theta_p^{\frac{2}{1-\alpha}}\]

(Case 2) \( \lambda A_p < 2 \hat{a} \theta_p^{\frac{2}{1-\alpha}} \)

In this case, \(k \leq \lambda A_p\) is binding and \(k^* = \lambda A_p\). The problem reduces

\[\max_{\{z_i, z_j\}} F(z_i, z_j) = \left[ \theta_p k^{\alpha} (z_i + z_j)^{1-\alpha} - \left( rk + \frac{\kappa}{2} (z_i^2 + z_j^2) \right) \right] \]

Aggregate production Aggregate cost

subject to \((z_i, z_j) > 0, \quad \alpha \in (0, 1), \quad \kappa > 0, \quad A_p \geq 0, \quad k = \lambda A_p\)
Following the same argument as in (Case 1), there exists the unique maximum and

\[ z_i = z_j = z^* = \left( \frac{1 - \alpha}{\kappa} \right)^{\frac{1}{1 + \alpha}} \left( \frac{\lambda A_p}{2} \right)^{\frac{\alpha}{1 + \alpha}} \theta_p^{\frac{1}{1 + \alpha}} \]

Therefore,

\[
V_p^E(\theta_p, A_p) = \frac{1}{2} \left[ \theta_p k^\alpha (2z^*)^{1-\alpha} - r k^\alpha \right] - \frac{\kappa}{2} z^*^2 = a_2 \left( \frac{\lambda A_p}{2} \right)^{\frac{\alpha}{1 + \alpha}} \theta_p^{\frac{2}{1 + \alpha}} - \frac{r \lambda A_p}{2}.
\]

### C.2 Proof of Proposition 3

If \( \{\gamma_1, \gamma_2, \alpha, \beta_0, \beta_1, \beta_2, \beta_3, \sigma_\eta, \hat{\gamma}_0, \hat{\gamma}_1, \hat{\gamma}_2, \hat{\gamma}_3, \sigma_\tilde{a}, \kappa \} \) is given, \( \lambda \) is the only variable which governs \( \Pr(d = 3 | X, A \leq 0) \). The equation (10) is rewritten as \( \Pr(d = 3 | X, A \leq 0) = \int_0^\infty \left( \Gamma_1 + \Gamma_2 \right) f(A') dA' \) where

\[
\Gamma_1 = \Pr \left[ \left( \frac{W(X)}{a_3} \right)^{\frac{1-\alpha}{2}} < \theta_p \leq \left( \frac{\lambda A'}{\hat{a}} \right)^{\frac{1-\alpha}{2}} \right]
\]

\[
\Gamma_2 = \Pr \left[ \max \left\{ \left( \frac{\lambda A'}{\hat{a}} \right)^{\frac{1-\alpha}{2}}, \left( \frac{W(X)}{a_4} + \frac{r \lambda A'}{2a_4} \right)^{\frac{1+\alpha}{2}} \left( \frac{\lambda A'}{2} \right)^{-\alpha} \right\} \leq \theta_p \right]
\]

\( W(X) \) is the wage earnings given \( X \).

\[
\frac{\partial \Pr(d = 3 | X, A \leq 0)}{\partial \lambda} = \int_0^\infty \left( \frac{\partial \Gamma_1}{\partial \lambda} + \frac{\partial \Gamma_2}{\partial \lambda} \right) f(A') dA'
\]

Fix \( A' \).

**Case 1** \( \left( \frac{\lambda A'}{\hat{a}} \right)^{\frac{1-\alpha}{2}} \geq \left( \frac{W(X)}{a_3} + \frac{r \lambda A'}{2a_4} \right)^{\frac{1+\alpha}{2}} \left( \frac{\lambda A'}{2} \right)^{-\alpha} \)

Under Case 1, \( \Gamma_1 + \Gamma_2 = \Pr \left[ \left( \frac{W(X)}{a_3} \right)^{\frac{1-\alpha}{2}} < \theta_p \right] \) and hence \( \left( \frac{\partial \Gamma_1}{\partial \lambda} + \frac{\partial \Gamma_2}{\partial \lambda} \right) = 0 \).

**Case 2** \( \left( \frac{\lambda A'}{\hat{a}} \right)^{\frac{1-\alpha}{2}} < \left( \frac{W(X)}{a_3} + \frac{r \lambda A'}{2a_4} \right)^{\frac{1+\alpha}{2}} \left( \frac{\lambda A'}{2} \right)^{-\alpha} \)

It is obvious that \( \frac{\partial \Gamma_1}{\partial \lambda} > 0 \). It is also easy to show that \( \frac{\partial \Gamma_2}{\partial \lambda} > 0 \). Therefore, \( \left( \frac{\partial \Gamma_1}{\partial \lambda} + \frac{\partial \Gamma_2}{\partial \lambda} \right) > 0 \).

It is now sufficient to show that the measure of \( A' \) satisfying Case 2 is positive for a given \( \lambda \). Case 2 is equivalent to \( \left( \frac{2W(X)}{a_4 \lambda A'} + \frac{r}{a'_4} \right)^{\frac{1+\alpha}{2}} > 1 \). For any \( \lambda \), one can always find \( A'_\lambda > 0 \) below which Case 2 is satisfied. Therefore, the measure of \( A' \) satisfying Case 2 is positive for
any given $\lambda$. This completes the proof. ■

C.3 Recovering Structural Parameters

There are many ways to recover structural parameters from $\tilde{G}_2$, $\tilde{G}_3$, $G_4$, $G_5$, $G_6$ and $\rho$. I present one example. I start with recovering $\alpha$ by using the following two equations.

$$ G_6(1) = \log a_1 + \frac{2(\beta_0 + \hat{g}_0)}{1 - \alpha} $$

$$ \tilde{G}_3(1) = \frac{1 - \alpha}{2\sigma_\eta}(\log a_3 - \log a_2) + \frac{\beta_0 + \hat{g}_0}{\sigma_\eta} $$

Combining two equations, I get

$$ G_6(1) = \frac{2\sigma_\eta}{1 - \alpha} \tilde{G}_3(1) + \log a_2 - \log \frac{3}{4} $$

$$ = \frac{2\sigma_\eta}{1 - \alpha} \tilde{G}_3(1) + \log a_6 + \log \left(\frac{1 + \alpha}{2}\right) - \log \frac{3}{4} $$

Since $\frac{2\sigma_\eta}{1 - \alpha}$ and $\log a_6$ can be recovered by $\frac{G_6(2)}{G_3(2)}$ and $G_4(1)$ respectively, $\alpha$ can be recovered. Once $\alpha$ is known, $\kappa$ can be recovered by the equation $G_4(1) = \log a_6$. The remaining parameters can be recovered as follows.

$$ \gamma_1 = \frac{(1+\alpha)G_4(3)}{2} \quad \gamma_2 = \frac{(1+\alpha)G_4(4)}{2} \quad \sigma_\eta = \frac{1 - \alpha}{2} \left(\frac{G_6(2)}{G_3(2)}\right) $$

$$ \sigma_\hat{u} = \rho\sigma_\eta \quad \beta_0 = \frac{1 - \alpha}{2}(G_5(1) - \log(2a_1)) \quad \beta_1 = \frac{(1-\alpha)G_5(2)}{2} $$

$$ \beta_2 = \frac{(1-\alpha)G_5(3)}{2} \quad \beta_3 = \frac{(1-\alpha)G_5(4)}{2} \quad \hat{g}_0 = \frac{1 - \alpha}{2}(G_6(1) - \log a_1) - \beta_0 $$

$$ \hat{g}_1 = \frac{1 - \alpha}{2}(G_6(2) - G_5(2)) \quad \hat{g}_2 = \frac{1 - \alpha}{2}(G_6(3) - G_5(3)) \quad \hat{g}_3 = \frac{1 - \alpha}{2}(G_6(4) - G_5(4)) $$

64
D Additional Tables

Table 14: Moments for Ownership Choice

<table>
<thead>
<tr>
<th>$K_i$</th>
<th>$Z_i$</th>
<th>$\frac{1}{n} \sum^n_{i=1} K_i Z_i$</th>
<th>$\frac{1}{n} \sum^n_{i=1} \tilde{K}_i(\psi) Z_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathbb{I}_{si}$</td>
<td>1</td>
<td>0.0218</td>
<td>0.0208</td>
</tr>
<tr>
<td>$A_i$</td>
<td>0.5231</td>
<td>0.5442</td>
<td></td>
</tr>
<tr>
<td>$x_{1i}$</td>
<td>0.3042</td>
<td>0.2992</td>
<td></td>
</tr>
<tr>
<td>$x_{2i}$</td>
<td>0.4644</td>
<td>0.3560</td>
<td></td>
</tr>
<tr>
<td>$A_i x_{1i}$</td>
<td>7.9762</td>
<td>8.2319</td>
<td></td>
</tr>
<tr>
<td>$A_i x_{2i}$</td>
<td>12.8884</td>
<td>10.1497</td>
<td></td>
</tr>
<tr>
<td>$\mathbb{I}_{pi}$</td>
<td>1</td>
<td>0.0045</td>
<td>0.0043</td>
</tr>
<tr>
<td>$A_i$</td>
<td>0.1320</td>
<td>0.1216</td>
<td></td>
</tr>
<tr>
<td>$x_{1i}$</td>
<td>0.0638</td>
<td>0.0639</td>
<td></td>
</tr>
<tr>
<td>$x_{2i}$</td>
<td>0.0876</td>
<td>0.0838</td>
<td></td>
</tr>
<tr>
<td>$A_i x_{1i}$</td>
<td>1.9135</td>
<td>1.8931</td>
<td></td>
</tr>
<tr>
<td>$A_i x_{2i}$</td>
<td>3.0964</td>
<td>2.9320</td>
<td></td>
</tr>
</tbody>
</table>

NOTE: This table compares the actual and the simulated moments for ownership choice. $\mathbb{I}_{si}$ is the indicator function for single owners. $\mathbb{I}_{pi}$ is the indicator functions for partners. $A$ is net worth (ten thousands dollars in 2011). $x_1$ is the years of education. $x_2$ is the years of experience.
### Table 15: Moments for log Conditional Incomes

<table>
<thead>
<tr>
<th>$K_i$</th>
<th>$Z_i$</th>
<th>Observed $\frac{1}{n} \sum_{i=1}^{n} K_i Z_i$</th>
<th>Simulated $\frac{1}{n} \sum_{i=1}^{n} \tilde{K}_i(\psi) Z_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi_{wi} I_{wi}$</td>
<td>1</td>
<td>1.3339</td>
<td>1.3495</td>
</tr>
<tr>
<td></td>
<td>$x_{1i}$</td>
<td>19.2330</td>
<td>18.9771</td>
</tr>
<tr>
<td></td>
<td>$x_{2i}$</td>
<td>30.6172</td>
<td>30.3514</td>
</tr>
<tr>
<td></td>
<td>$(\pi_{wi} - \frac{1}{\sum_{i=1}^{n} \pi_{wi} I_{wi}}) I_{wi}$</td>
<td>1</td>
<td>0.6153</td>
</tr>
<tr>
<td>$\pi_{si} I_{si}$</td>
<td>1</td>
<td>0.0037</td>
<td>0.0113</td>
</tr>
<tr>
<td></td>
<td>$A_i$</td>
<td>0.2576</td>
<td>0.2876</td>
</tr>
<tr>
<td></td>
<td>$x_{1i}$</td>
<td>0.0583</td>
<td>0.1679</td>
</tr>
<tr>
<td></td>
<td>$x_{2i}$</td>
<td>0.1042</td>
<td>0.2537</td>
</tr>
<tr>
<td></td>
<td>$A_i x_{1i}$</td>
<td>3.7701</td>
<td>4.4577</td>
</tr>
<tr>
<td></td>
<td>$A_i x_{2i}$</td>
<td>7.4794</td>
<td>7.2806</td>
</tr>
<tr>
<td></td>
<td>$(\pi_{si} - \frac{1}{\sum_{i=1}^{n} \pi_{si} I_{si}}) I_{si}$</td>
<td>1</td>
<td>0.0583</td>
</tr>
<tr>
<td>$I_{si}$</td>
<td>1</td>
<td>0.0019</td>
<td>0.0021</td>
</tr>
<tr>
<td>$\pi_{pi} I_{pi}$</td>
<td>1</td>
<td>0.0044</td>
<td>0.0044</td>
</tr>
<tr>
<td></td>
<td>$A_i$</td>
<td>0.1623</td>
<td>0.1463</td>
</tr>
<tr>
<td></td>
<td>$x_{1i}$</td>
<td>0.0671</td>
<td>0.0668</td>
</tr>
<tr>
<td></td>
<td>$x_{2i}$</td>
<td>0.0865</td>
<td>0.0938</td>
</tr>
<tr>
<td></td>
<td>$A_i x_{1i}$</td>
<td>2.4830</td>
<td>2.2951</td>
</tr>
<tr>
<td></td>
<td>$A_i x_{2i}$</td>
<td>3.7125</td>
<td>3.3380</td>
</tr>
<tr>
<td></td>
<td>$(\pi_{pi} - \frac{1}{\sum_{i=1}^{n} \pi_{pi} I_{pi}}) I_{pi}$</td>
<td>1</td>
<td>0.0069</td>
</tr>
<tr>
<td>$I_{pi}$</td>
<td>1</td>
<td>0.0004</td>
<td>0.0004</td>
</tr>
</tbody>
</table>

NOTE: This table compares the actual and the simulated moments for log conditional incomes. $\pi_{wi}$ is log income conditional on $i$ is a worker. $\pi_{si}$ is log income conditional on $i$ is a single owner and $i$ reports a positive income. $I_{si}$ is the indicator function for single owners who report a positive income. $I_{wi}$ is the indicator function for single owners who report a negative or zero income. $\pi_{pi}$ is log income conditional on $i$ is a partner and $i$ reports a positive income. $I_{pi}$ is the indicator function for partners who report a positive income. $I_{pi}$ is the indicator function for partners who report a negative or zero income. $A$ is net worth (ten thousands dollars in 2011). $x_1$ is the years of education. $x_2$ is the years of experience.
Table 16: Number of Observations at Each Stage of Sample Construction

<table>
<thead>
<tr>
<th>Drop if</th>
<th>Workers</th>
<th>Single owners</th>
<th>Partners</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>247,828</td>
<td>16,078</td>
<td>3,580</td>
</tr>
<tr>
<td>Female</td>
<td>114,745</td>
<td>9,888</td>
<td>2,385</td>
</tr>
<tr>
<td>Age &lt; 18 or &gt; 65</td>
<td>65,777</td>
<td>8,856</td>
<td>2,189</td>
</tr>
<tr>
<td>Unemployed in either one of periods</td>
<td>49,224</td>
<td>8,519</td>
<td>2,154</td>
</tr>
<tr>
<td>Business owners in the base year</td>
<td>48,182</td>
<td>1,564</td>
<td>288</td>
</tr>
<tr>
<td>No info. on net worth</td>
<td>47,466</td>
<td>1,541</td>
<td>285</td>
</tr>
<tr>
<td>Family business</td>
<td>47,317</td>
<td>1,491</td>
<td>199</td>
</tr>
<tr>
<td>Net worth &gt; $10 billion</td>
<td>47,304</td>
<td>1,491</td>
<td>199</td>
</tr>
<tr>
<td>Active duty military</td>
<td>47,304</td>
<td>1,490</td>
<td>199</td>
</tr>
<tr>
<td>No info. on equity share or</td>
<td>47,304</td>
<td>1,135</td>
<td>199</td>
</tr>
<tr>
<td>equity share &lt; 25%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wage work hours &gt; business work hours</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&amp; business work hours &lt; 30</td>
<td></td>
<td>941</td>
<td>162</td>
</tr>
</tbody>
</table>

NOTE: Business owners who did not report equity share or whose equity share is less than 25% were categorized as single owners before being deleted.

Table 17: Proportion of African-American with respect to Income Before Tax

<table>
<thead>
<tr>
<th></th>
<th>$18,559</th>
<th>$35,645</th>
<th>$58,272</th>
<th>$93,837</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEX 2011 (%)</td>
<td>20</td>
<td>14</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>Final sample (%)</td>
<td>12</td>
<td>12</td>
<td>8</td>
<td>6</td>
</tr>
</tbody>
</table>

NOTE: Dollar in 2011. This table compares the proportion of African-American in terms of income levels both for the final sample and for CEX 2011. For the final sample, incomes in the base year are used.
E  Random Matching

As a robustness check, I introduce a search friction into the model. I estimated the parameters under the assumption that the matching market is frictionless. The frictionless assumption has an implication for the estimates of the collaborative skill. Without a search friction, the collaborative skill may be estimated low to match the small proportion of business partners in the data.

A search friction can be implemented as follows. The frictionless matching market allows an agent to be matched with any other agent in the market. We can limit this possibility by dividing the market into, for example, two groups and not allowing an agent in one group to be matched with any agent in the other group. As an extreme, two agents in the market are randomly matched to each other and not allowed to form a partnership with any other agent except for the randomly matched agent. I re-estimate the model under this random matching assumption. For estimation, I assign each individual in the final sample a matched agent, who is drawn from the uniform distribution over the final sample with replacement.

The parameter estimates under the random matching is similar to those under the frictionless matching. Notably, $g_0$, the constant term for log collaborative skill, is still estimated negative; it was estimated $-1.3923$ under the frictionless matching, and is estimated $-1.3518$ under the random matching. I also simulate the collaborative skill with the new estimates and compare it with Figure 10. The result is shown in Figure 16. The distribution of $g$ is shifted toward the right. Hence the proportion of agents who have productivity as a partner greater than their solo productivity increases from 3.4% to 4.3%. Even so, the majority of individuals are still estimated to have productivity as a partner less than their solo productivity.
Figure 16: Kernel Density of Collaborative Skill: Frictionless Matching vs. Random Matching

NOTE: This figure shows Kernel density of a simulated value for $g$ given covariates and estimates under different matching markets. $g$ is greater than one means that productivity as a partner is greater than the solo productivity. $g$ is less than one means that productivity as a partner is less than the solo productivity.