Spouses and Entrepreneurship

João Alfredo Galindo da Fonseca[1] and Charles Berubé[2]

[1]Université de Montréal and CIREQ
[2]Innovation, Science and Economic Development Canada

May 2, 2019

(First Draft : 25 April 2019)

Abstract

Does having a spouse influence an individual's decision to start a firm and what firms they create? The answer to this question is crucial for our understanding of how recent changes to family composition influence firm creation. We develop a model of endogenous entrepreneurship with spousal labour supply decisions and endogenous marriage. Married individuals have three channels, that go in opposite directions, which influence their choice to start a firm relative to the unmarried. Firstly, spouses work less when the business is more profitable partially offsetting the benefit of higher profits (spousal substitution effect). Secondly, if the business fails, the spouse works more hours (spousal insurance effect). Finally, a married individual shares their income with their spouse which decreases their income as a worker, their cost to entrepreneurship (spousal opportunity cost effect). We proceed to test empirically the relative strength of these channels. The model is informative of the components of the error term and the conditions for validity of our instrumental variable strategy. Using city level variation in the past composition of immigrants we show higher marriage rates are associated to more entry and lower average size of startups.

JEL-Codes : E24, E23, J63, J64

Keywords : Firm dynamics, Macroeconomics, Labor Markets, Family Economics

1 Introduction

How does having a spouse affect the decision to start a firm? How does it affect the type of firms created by individuals? The answer to this question is crucial if we wish to understand how changes in household composition affect job creation. We know there have been secular changes to household composition in the US and Canada. This is best portrayed by the decline in marriage rates (from 59.71% in 1984 to 45.58% in 2013 for Canada and from 62% in 1990 to 55% in 2013 for the US) and the rise of one person households (from 20.3% in 1981 to 27.6% in 2011 for Canada and from 9% to 12% for the US). To understand how these changes affected net job creation and productivity, we need to know how they affected firm formation and the composition of new firms.[1] Despite the huge literature on entrepreneurship and firm formation[2], there has been no research on how marriage can affect both entry into entrepreneurship and firm size. This paper fills in this gap.

To study the role of spouses on the entrepreneurship of the main earner we propose a tractable model of endogenous entrepreneurship, endogneous marriage, endogenous spousal labour supply, heterogeneity in innate entrepreneurial ability and business project specific productivity. In the model, main earners choose ex-ante whether to marry based on a taste for marriage, their innate entrepreneurial ability and the overall productivity of the economy. Households can either be two persons or single person households. Main earners in the household draw business opportunities from an exogenous distribution. In equilibrium, individuals exhibit a threshold policy rule. For all business opportunities with productivity above the optiomal threshold, the individual starts a firm. In two person households, main earner and spouse share their income, such that the total consumption of each individual depends on both their income and of their partner. Spouses maximize their

[1] Haltiwanger (2011), Haltiwanger et al. (2015), Diez et al. (2014) and Clementi and Palazzo (2016) study the contribution of firm creation to productivity. Decker et al. (2014) highlight the strong contribution of startups to job creation.

[2] See Lucas Jr (1978), Holtz-Eakin et al. (1994), Fonseca et al. (2001), Hurst and Lusardi (2004), Cagetti et al. (2006), Quadrini (2000), Beaudry et al. (2018), Hamilton (2000) and Haltiwanger et al. (2013).

individual utility by choosing how many hours to work, given their wage and the income of the main earner. Main earners make the individually optimal decision of whether or not to start a firm taking into account their spouse's optimal behaviour.

This tractable framework delivers three channels via which being a couple influences the decision to start a firm. First, spouses work less when the main earner's business is more profitable partially offsetting the benefit of high profits (spousal substitution effect). This channel makes entrepreneurship relatively less attractive to married individuals which in turn become more selective on which business projects to implement.

Secondly, if the business fails, the spouse works more hours (spousal insurance effect). This effect is a natural application of the concept of added worker effect vastly studied in labor economics (See Lundberg (1985), Maloney (1987), Hyslop (2001), Stephens (2002), Juhn and Potter (2007), Gallipoli and Turner (2009), Blundell et al. (2016), Blundell et al. (2018), Wu and Krueger (2018)). This insurance channel decreases the cost of failure for married individuals. This pushes a married person to become less selective in which business projects to implement.

Finally, a married individual shares their income with their spouse which decreases their income as a worker, their cost to entrepreneurship (spousal opportunity cost effect). This channel is consistent with a growing literature studying the importance of the opportunity costs to entrepreneurship. This channel further pushes married individuals to accepting less productive business opportunities. As a result, the total effect of having a spouse on entry into entrepreneurship is ambiguous. The model predicts that if the spousal insurance effect and the spousal opportunity cost effect dominate over the spousal substitution effect than higher marriage rates induce higher entry rates into entrepreneurship and lower average number of employees of new startups. If on the other hand, the spousal substitution effect dominates over the other two effects, we expect higher marriage rates to induce lower entry rates into entrepreneurship and higher average number of employees of new startups.

This tells us that regardless of the which effect dominates the response of entry into entrepreneurship and average startup size to marriage should be the inverse of each other. Although the model predicts ambiguity concerning the response of entry rates into entrepreneurship to marriage it has a sharp prediction that marriage rates affect average size of startups in the opposite direction it affects entry rates. This is a key empirical implication as it directly relates to the selection mechanism behind the differences between married and unmarried individuals in the model.

Next, we proceed to an empirical analysis with the goal of investigating the relative strenghts of these different channels and testing the model's empirical prediction. The empirical approach we use is a structural-IV approach. We use the model to derive our instrument and the conditions under which it provides consistent estimates of the coefficients of interest. We estimate straightforward linear regressions using an instrument indicated from the theory. This strategy fits with a literature that puts together structural modelling and instrumental variable estimation (Blundell et al. (1998), Beaudry et al. (2012), Beaudry et al. (2018), Tschopp (2015), Sand et al. (2016) and Green et al. (2017))). We focus on estimating first order implications of the theory as implied by its linear approximation. This offers a simple, clear and intuitive exposition of results. Our instrument derived from the theory is the sum of changes in marriage of a immigrant group at the national level weighted by the share of the population in the city of those immigrant groups. We find that a 1% point increase in marriage rates is associated to 15.7% increase in the entry rate into entrepreneurship and a 15.46% drop in the average size of startups.

These estimates allows us to perform a back of the enveloppe calculation of the response of job creation to marriage rates. The results imply that a 1% increase in the marriage induces a 0.24% increase in job creation. This number hides the importance of taking into account the selection mechansim that imposes a negative relationship between average size of startups and entry rates into entrepreneurship. In particular, if a researcher does not taken into account the selection mechansim (ignores the response of the average size of startups), they would wrongly conclude that a 1% point increase in the marriage rate is associated to 15.7% increase in job creation. This serves as a word of caution for policy analysis that ignores the impact on average size and focuses on the response of firm entry. The result also highlights the importance of being explicit of the source of endogeneity and how to address it. Ignoring the endogeneity problem would also lead to a small response of job creation. However, this would derive from the wrong conclusion that the selection mechanism is not present. In reality, job creation does not respond exactly due to the strong negative relationship between entry and average size generated by the selection mechanism.

One concern is that our results are being driven by borrowing constraints (more easily overcome by two people households) and/or spouses working in the business. We verify our results are robust in both magnitude and significance to excluding high capital industries, firms created jointly by spouses and firms for which the spouse is listed as an employee. This is consistent with our results not being driven by these alternative channels.

The model relates to papers studying portofolio allocation of households composed of an entrepreneur and a worker (Panousi (2008), Angeletos and Panousi (2009) and Angeletos and Panousi (2011)). The key difference is that in these papers there is a fixed composition of one entrepreneur and one worker. In our model, on the other hand, one of the two agents beloging to the household is making endogenously the decision to enter entrepreneurship. As a result, there are moments where both agents in the household are workers while other moments where one is an entrepreneur and the other is a worker.

2 Model

In this section I go over the main model of the paper. The main objective is to derive the key intuition of the different forces that come into play when an individual decides to start a firm as a function of having a spouse or not. We start by describing individual choices conditional on marriage and individual entrepreneurial ability θ. After having done that, we describe the endogenous decision to marry as a function of entrepreneurial ability θ and individual taste for marriage.

In the model we abstract from the possibility of the spouse helping out in the business started by their partner. This assumption is consistent with our results in the empirical section. In particular, we verify our empirical results are not being driven by firms started jointly by the couple or for which the spouse is listed as an employee of the firm. [3] We also abstract from borrowing constraints. This is consistent with the evidence provided in the empirical section where we show our results are robust to excluding industries that are capital intensive.[4]

In the economy there are two types of households. Single individual households are composed of only one individual who is also the main earner. The second type of household is that of couples. These are composed of one main earner and one spouse. There is no savings and each household consumes their current income. We consider log utility. There are no search frictions. All individuals searching for a job, find one instantaneously. The utility derived from income I for the single person household is given by

$$log(I). \tag{1}$$

For households composed of two individuals, there are two sources of income, one is the income of the main earner I and the other is the income of the spouse $w^s h$. The income of the spouse is a function of how many hours the spouse chooses to optimally work h and the wage paid to do so w^s. To make the model tractable we consider the wage of the spouse is paid by an outside sector not explicitly modelled.[5] We consider a parametrization where the wage of the main earner is larger than that of the spouse, $w^s < w$. Although not modelled here to keep the model tractable, this is consistent with our notion of the spouse being the individual in the household that specializes on non-market activities.

[3]We also abstract from the benefit of entrepreneurs receiving employer provided health insurance via their spouse. This choice is consistent with our data where such considerations are likely of second order due to Canada's public health care system.

[4]This is not to say that borrowing constraints are not important for decisions to start a firm. Rather, it says that the differences in entrepreneurship caused by marriage do not seem to be driven by borrowing constraint considerations.

[5]This can be thought of as not modelling the specific sectors that spouses tend to concentrate more relative to main earners. This allows us to avoid the discussion on the source of lower wages among women relative to men.

Let χ denote the share of spousal income that is left for the spouse to consume and γ be the share of main earner income that is left for the main earner to consume. Then the utility of the spouse working h hours with a main earner of income I is given by

$$log(\gamma I + \chi w^s h) \tag{2}$$

and the utility of the main earner in the couple is given by

$$log((1-\gamma)I + (1-\chi)w^s h). \tag{3}$$

where γ and μ are such that $0 \leq \gamma \leq 1, 0 \leq \chi \leq 1$.[6] Furthermore, let us assume $\frac{1}{2-\chi} > \gamma$.[7] We abstract from the choice to start a firm or not for the spouse, so as to concentrate on the main earner. The spouse faces the static problem of how many hours to work every period. Given the income I of the main earner, the spouse maximizes his/her utility by choosing how many hours h to work which are paid according to w^s. There is also a disutility associated to working for the spouse of ϕh. As a result, the spouse solves

$$\max_{0 \leq h \leq 1} log(\gamma I + \chi w^s h) - \phi h \tag{4}$$

This gives us

$$h = \frac{1}{\phi} - \frac{\gamma I}{\chi w^s} \tag{5}$$

when h is an interior solution and 0 or 1 when we get corner solutions. We focus on parameter specifications where h always has an interior solution. Hence, it follows that

[6]Note that χ and γ imply that income from the spouse and from the main earner might not be perfect substitutes in the utility function of an individual in the couple.

[7]This is a necessary condition to guarantee the value of being an entrepreneur is always increasing in productivity of the business z.

the hours worked by the spouse are a decreasing function of the income of the main earner:

$$\frac{\partial h}{\partial I} < 0 \tag{6}$$

Let us consider the problem of a currently operating entrepreneur. Each main earner has innate entrepreneurial ability of θ. The distribution of θ in for main earners of both single and couple households is exogenous and equal to $G(\theta)$. Let also the productivity of the firm be a function of a economy-wide productivity component y. In other words, y is the same for all entrepreneurs in a same economy while θ is a individual ability component that varies across entrepreneurs. The individual entrepreneur takes wages w and firm producitivity components z, θ and y as given and choose how many individuals n to hire. Hence they solve

$$\pi(z) = \max_n y\theta e^z n^\alpha - wn \tag{7}$$

which gives

$$n(z, w) = (\frac{\alpha y\theta e^z}{w})^{\frac{1}{1-\alpha}} \tag{8}$$

and

$$\pi(z) = (1-\alpha)(\frac{\alpha}{w})^{\frac{\alpha}{1-\alpha}} e^{\frac{z}{1-\alpha}} (y\theta)^{\frac{1}{1-\alpha}}. \tag{9}$$

Once the entrepreneur start operating, firm productivity is fixed at θe^z. For the remainder of this section we omit the dependance of value functions on θ to make the notation lighter. With probability λ the firm fails and the individual goes bankrupt. What follows next should be understood as conditional on a entrepreneurial ability θ. Let $J^m(z)$ represent the value of being an entrepreneur with a spouse, B^m the value of being bankrupt with a spouse, then,

$$rJ^m(z) = log[(1-\gamma)\pi(z) + (1-\chi)w^s h(\pi(z), w^s)] + \lambda(B^m - J^m(z)). \tag{10}$$

where the flow utility of an entrepreneur depends on how much they get from their income, $(1-\gamma)\pi(z)$, and how much they get from their spouse's income, $(1-\chi)w^s h(\pi(z), w^s)$. Similarly, let $J^u(z)$ represent the value of being an entrepreneur without a spouse, B^u the value of being bankrupt without a spouse, then,

$$rJ^u(z) = log[\pi(z)] + \lambda(B^u - J^u(z)). \tag{11}$$

The difference between the value of being an entrepreneur between the two groups depends on any differences in the value of being bankrupt B^m versus B^u and differences in their flow utility. In particular, while the unmarried individual gets to keep all of profits, the married individual needs to share some with the spouse. But, the married individual also has the advantage of getting some of the income from the spouse. Note that the income as an entrepreneur when unmarried is increasing in z.[8] On the other hand, for the married individual as their entrepreneurial profits, $\pi(z)$ increase in z, hours worked by the spouse, h, decrease. Hence, spousal income decreases partially offsetting some of the increase in income due to higher profits. This is an important effect which we revisit later. We call it the **spousal substitution effect**.

Once bankrupt the individual works as a wage worker but is forced to pay a cost c and is not allowed to entrepreneurship. Hence, the income of the main earner in this case is $w - c$. The individual exits bankruptcy back to the start of wage work with probability p. Let W^m be the value of being a married wage worker, then the value of being bankrupt married is

$$rB^m = log[(1-\gamma)(w-c) + (1-\chi)w^s h((w-c), w^s)] + p(W^m - B^m) \tag{12}$$

where the flow utility of the bankrupt married individual depends on how much they get from their own income $(1-\gamma)(w-c)$ and how much they get from their spouse income $(1-\chi)w^s h((w-c), w^s)$. Similarly, let W^u denote the value of being a unmarried wage worker, the the value of being bankrupt and unmarried, B^u is

$$rB^u = log[w-c] + p(W^u - B^u) \tag{13}$$

[8]Recall that profits, $\pi(z)$ is increasing in z.

The difference in the value of being bankrupt between the two groups, depends on the difference in continution values W^m versus W^u and in differences in flow utility. In particular, while for unmarried individuals, income falls by the cost of bankruptcy c, married individuals benefit from the income of the spouse. When the main earner is bankrupt, with a lower income, spouses increase working hours h. The result is an increase in spousal earning which partially offsets the decrease in income suffered by married bankrupt individuals. We call this effect the **spousal insurance effect**. We will revist it later.

Next, consider the value of being a wage worker when married W^m versus unmarried W^u. For both types of individuals, business projects arrive a rate ψ. Each project is associated to a firm productivity z drawn from an exogenous distribution $F(z)$. Individuals choose optimally which projects to implement comparing the value of opening a firm ($J^m(z)$ if married and $J^u(z)$ if unmarried) to the value of being a wage worker. Let $\underline{z}^m$ represent the firm productivity that makes the married individual between opening a firm and continuing to work. Then,

$$J^m(\underline{z}^m) = W^m \tag{14}$$

It follows the value of being a married wage worker is given by

$$rW^m = log[(1-\gamma)w + (1-\chi)w^s h(w, w^s)] + \psi \int_{\underline{z}^m} (J^m(z) - W^m)dF(z). \tag{15}$$

The flow utility a married individual when working is given by how much they get from their own income, $(1-\gamma)w$, and how much they get from their spouse's income, $(1-\chi)w^s h(w, w^s)$. Similarly, let $\underline{z}^u$ represent the firm productivity that makes the married individual indifferent between opening a firm and continuing to work. Then,

$$J^u(\underline{z}^u) = W^u \tag{16}$$

It follows the value of being a unmarried wage worker is given by

$$rW^u = log[w] + \psi \int_{\underline{z}^u} (J^m(z) - W^m)dF(z). \tag{17}$$

The difference in the value of being a wage worker between married and unmarried main earners, depends on the difference in continuation values $J^m(z)$ versus $J^u(z)$ and in the differences in flow utility. Furthermore, this flow utility as a worker represent the opportunity cost to starting a firm. Intuitively, when an individual starts a firm they must forgo of their current flow utility. For an unumarried individual the flow utility is given exclusively by their income as a worker w. For the married main earner their utility depends on how much they get from their income, $(1-\gamma)w$, and how much they get from their spouse's income, $(1-\chi)w^s h(w, w^s)$. Note that since $0 \leq h \leq 1$ and $w^s < w$ it follows that

$$log[(1-\gamma)w + (1-\chi)w^s h(w, w^s)] < log[w]. \tag{18}$$

Hence, the opportunity cost of entrepreneurship is larger for unmarried individuals relative to married individuals. For all else equal, this increases the cost to enter entrepreneurship for unmarried individuals relative to married individuals. We call this the **spousal opportunity cost effect**. This effect just comes from the fact that a married individual must share their income with their spouse. In the case of a main earner who is a wage worker this means a lower consumption than they would have otherwise had when unmarried. However, since this represents the opportunity cost to entrepreneurship, it implies there is a smaller cost for married individuals.

2.1 Equilibrium Measure of Entrepreneurs

Next, we solve for the equilibrium measure of entrepreneurs conditional on a share o married individuals. Let M be the exogenous measure of married individuals in the economy. Let $\eta_m(\theta)$ be share of married individuals with entrepreneurial ability θ that are entrepreneurs, $e_m(\theta)$, be the share of married individuals with entrepreneurial ability θ that are workers and $b_m(\theta)$ the share of married individuals with entrepreneurial ability θ

that are bankrupt.[9] Similarly define $\eta_u(\theta)$, $e_u(\theta)$ and $b_u(\theta)$ as the same corresponding share for unmarried individuals. The share of entrepreneurs for both groups is characterized by

$$\dot{\eta}_i(\theta) = \psi(1 - F(\underline{z}^i(\theta)))e_i(\theta) - \lambda\eta_i(\theta), \quad \forall i \in \{m, u\}. \tag{19}$$

Similarly, the share of individuals bankrupt in both groups is characterized by

$$\dot{b}_i(\theta) = \lambda\eta_i(\theta) - pb_i(\theta), \quad \forall i \in \{m, u\}. \tag{20}$$

Finally, the share of workers in both groups is characterized by

$$\eta_i(\theta) + b_i(\theta) + e_i(\theta) = 1, \quad \forall i \in \{m, u\}. \tag{21}$$

Setting $\dot{\eta}(\theta) = 0$ and $\dot{b}_i(\theta) = 0$ and solving for these shares gives us

$$e_i(\theta) = \frac{\lambda p}{\lambda p + \psi(1 - F(\underline{z}^i(\theta)))(p + \lambda)}, \quad \forall i \in \{m, u\}. \tag{22}$$

$$b_i(\theta) = \frac{\psi(1 - F(\underline{z}^i(\theta)))\lambda}{\lambda p + \psi(1 - F(\underline{z}^i(\theta)))(p + \lambda)}, \quad \forall i \in \{m, u\}. \tag{23}$$

$$\eta_i(\theta) = \frac{\psi(1 - F(\underline{z}^i(\theta)))p}{\lambda p + \psi(1 - F(\underline{z}^i(\theta)))(p + \lambda)}, \quad \forall i \in \{m, u\}. \tag{24}$$

Futhermore, the fraction of individuals $i \in \{m, u\}$ (m = married, u = unmarried) of productivity, e^z with entrepreneurial ability θ, $\Gamma_i(z, \theta)$ is characterized by

$$\dot{\Gamma}_i(z, \theta) = \psi f(z)e_i(\theta) - \lambda\Gamma_i(z, \theta), \quad \forall z \geq \underline{z}^i(\theta) \quad i \in \{m, u\} \tag{25}$$

[9] In other words, $\eta_m(\theta) + e_m(\theta) + b_m(\theta) = 1$.

Using the expression for $e_i(\theta)$ and setting $\dot{\Gamma}_i(z,\theta) = 0$ we get

$$\Gamma_i(z,\theta) = \frac{\psi f(z)p}{\lambda p + \psi(1 - F(\underline{z}^i(\theta)))(p+\lambda)} \tag{26}$$

This implies the measure of firms of productivity θe^z, $\Lambda(z,\theta)$, of this economy is given by

$$\Lambda(z,\theta) = \frac{M\psi f(z)pG(\theta)}{\lambda p + \psi(1 - F(\underline{z}^m(\theta)))(p+\lambda)} \mathbb{1}\{z \geq \underline{z}^m(\theta)\} + \frac{(1-M)\psi f(z)pG(\theta)}{\lambda p + \psi(1 - F(\underline{z}^u(\theta)))(p+\lambda)} \mathbb{1}\{z \geq \underline{z}^u(\theta)\} \tag{27}$$

where $\mathbb{1}\{z > a\}$ is an indicator function equal to 1 if $z > a$ and 0 otherwise, $\forall a$.[10]Hence, the average firm productivity $E[z]$ in this economy is given by

$$E[z] = ME[z]^m + (1-M)E[z]^u = M \int \int_{\underline{z}^m(\theta)} \frac{e^z f(z)}{1 - F(\underline{z}^m(\theta))} dzG(\theta)d\theta + (1-M) \int \int_{\underline{z}^u(\theta)} \frac{e^z f(z)}{1 - F(\underline{z}^u(\theta))} dzG(\theta)d\theta.$$

where $E[z]^m$ is the average productivity among firms created by married individuals and $E[z]^u$ is the average productivity among firms created by unmarried individuals. From this expression we can see that $E[z]$ is increasing in M if $\underline{z}^m > \underline{z}^u$ and decreasing otherwise. It follows that relative selection of both groups ($\underline{z}^m$ versus $\underline{z}^u$) is crucial for our understanding of how changes in the marriage rate affect the firm productivity distribution.

2.2 Outcomes and Spouses

Given the three channels we discussed (**spousal insurance effect**, **spousal substitution effect** and **spousal opportunity cost effect**), we are interested in how each of them affect the entry into entrepreneurship and the size of new firms created among married versus un-

[10]In other words, $\Lambda(z,\theta)$ is defined such that $\int \Lambda(z,\theta)dz = MG(\theta)\eta_m(\theta) + (1-M)G(\theta)\eta_u(\theta)$.

married individuals. In our model these objects are both captured by the selection thresholds, $\underline{z}^m$ for the married, and $\underline{z}^u$ for the unmarried. The discussion that follows should be understood as conditional on a value for θ. Once, we bring model to the data we discuss how our identification strategy deals with the endogeneity problem associated to the presence of θ.

The entry rate into entrepreneurship for both groups is determined respectively by

$$\psi(1 - F(\underline{z}^m)) \tag{28}$$

and

$$\psi(1 - F(\underline{z}^u)). \tag{29}$$

From the expressions above we see that a higher treshold decreases the entry rate of the corresponding group. Let us define as startups firms that have just started. Then, given the exogenous distribution $F(z)$ from where business projects are drawn, the average size of startups among married entrepreneurs, $E[s]^m$, is given by

$$E[s]^m = \int_{\underline{z}^m} n(z,w)\frac{f(z)}{1 - F(\underline{z}^w)}dz = \int_{\underline{z}^m} (\frac{\alpha y \theta e^z}{w})^{\frac{1}{1-\alpha}} \frac{f(z)}{(1 - F(\underline{z}^m))}dz. \tag{30}$$

Similarly average size of startups among unmarried individuals, $E[s]^u$, is given by

$$E[s]^u = \int_{\underline{z}^u} n(z,w)\frac{f(z)}{1 - F(\underline{z}^u)}dz = \int_{\underline{z}^u} (\frac{\alpha y \theta e^z}{w})^{\frac{1}{1-\alpha}} \frac{f(z)}{(1 - F(\underline{z}^u))}dz. \tag{31}$$

From these expressions we see that

$$\frac{\partial E[s]^m}{\partial \underline{z}^m} > 0 \tag{32}$$

and

$$\frac{\partial E[s]^u}{\partial \underline{z}^u} > 0. \tag{33}$$

Hence, we conclude that average size of startups for a group (married versus unmarried) is increasing in the treshold chosen by that group. To summarize, if we want to know whether married individuals enter more or less and create smaller or larger firms than unmarried individual then we must uncover the relationship between their productivity tresholds :

$$\underline{z}^m \lesseqgtr \underline{z}^u? \tag{34}$$

To answer this question let us look separately at the effect of each difference in value function across married and unmarried individuals. In particular, we are going to consider separately, the effect of having the spouse work when the individual is a worker, when the individual is an entrepreneur and when the individual is bankrupt. As we shall see these will all have separate effects on the decision of an individual to start a firm relative to when unmarried.

With this purpose let us first consider the extreme case where the spouse of the married individual only works when the individual is an operating entrepreneur. In this case the value of being an entrepreneur for the married M and unmarried U are given by

$$rJ^m(z) = log[(1-\gamma)\pi(z) + (1-\chi)w^s h(\pi(z))] + \lambda(B^m - J^m(z)) \tag{35}$$

and

$$rJ^u(z) = log[\pi(z)] + \lambda(B^u - J^u(z)). \tag{36}$$

The value of being bankrupt for the married and unmarried would be given by

$$rB^m(z) = log[w-c] + p(W^m - B^m(z)) \tag{37}$$

and

$$rB^u(z) = log[w-c] + p(W^u - B^u(z)). \tag{38}$$

Finally the value of being an entrepreneur for the married and unmarried would be given by

$$rW^u = log[w] + \psi \int_{\underline{z}^u} (J^u(z) - W^u)dF(z) \tag{39}$$

and

$$rW^m = log[w] + \psi \int_{\underline{z}^m} (J^m(z) - W^m) dF(z). \tag{40}$$

In this case, we see that the only difference between married and unmarried individuals is that for married individuals they get an additional income when an entrepreneur, however, this additional income of the spouse is decreasing in the productivity of the firm. To see this, recall that the hours worked by the spouse are a decreasing function of the income of the main earner. Hence, a higher profit of the firm of the main earner, decreases the hours worked by the spouse. If we take the derivative with respect to firm productivity z of the value of being an entrepreneur $J^m(z, h(\pi(z))$ for the married individual we get

$$\frac{\partial J^m(z, h(\pi(z)))}{\partial z} = \frac{u'((1-\gamma)\pi(z) + (1-\chi)w^s h)[(1-\gamma)\frac{\partial \pi(z)}{\partial z} + (1-\chi)w^s \frac{\partial h}{\partial z}]}{r+\lambda} \tag{41}$$

Recall that

$$\frac{\partial h}{\partial I} = -\frac{\gamma}{w^s} \tag{42}$$

Hence

$$\frac{\partial h}{\partial z} = \frac{\partial h}{\partial \pi(z)} \frac{\partial \pi(z)}{\partial z} = -\frac{\gamma(1-\chi)}{w^s} \frac{\partial \pi(z)}{\partial z} \tag{43}$$

which implies

$$\frac{\partial J^m(z, h(\pi(z)))}{\partial z} = \frac{u'((1-\gamma)\pi(z) + (1-\chi)w^s h)[(1-\gamma) - \gamma(1-\chi)]\frac{\partial \pi(z)}{\partial z}}{r+\lambda} \tag{44}$$

$$\frac{\partial J^m(z, h(\pi(z)))}{\partial z} = \frac{u'((1-\gamma)\pi(z) + (1-\chi)w^s h)[1 - 2\gamma + \chi\gamma]\frac{\partial \pi(z)}{\partial z}}{r+\lambda} > 0. \tag{45}$$

In other words, the value of being an entrepreneur is increasing in the productivity of the firm for married individuals.[11] For the unmarried we have

$$\frac{\partial J^u(z)}{\partial z} = \frac{u'(\pi(z))\frac{\partial \pi(z)}{\partial z}}{r+\lambda} > 0 \tag{46}$$

[11]To see that the derivative is positive note that $\frac{1}{2-\chi} > \gamma \Rightarrow 1 - 2\gamma + \chi\gamma > 0$.

It follows that the derivative of the value of being an entrepreneur is increasing with productivity for both married and unmarried individuals, however, the slope is different between both individuals. Crucially whether or not the slope for a married entrepreneur is larger or not depends on whether

$$u'((1-\gamma)\pi(z) + (1-\chi)w^s h)[1 - 2\gamma + \chi\gamma] \lesseqgtr u'(\pi(z)). \tag{47}$$

which is equivalent to

$$\frac{1 - 2\gamma + \chi\gamma}{(1-\gamma)\pi(z) + (1-\chi)w^s h} \lesseqgtr \frac{1}{\pi(z)} \tag{48}$$

given $u(c) = log(c)$. This simplifies to

$$-\pi(z)\gamma + \chi\gamma\pi(z) \lesseqgtr (1-\chi)w^s h \tag{49}$$

$$-\pi(z)\gamma(1-\chi) < w^s h(1-\chi) \tag{50}$$

It follows married individuals have a smaller slope for the value of being an entrepreneur, $J(z)$, to productivity z relative to unmarried individuals. Intuitively, this comes from the fact that as the individual becomes richer (by starting a higher productivity business project z) the spouse works less. This partially offsetts the increase in household income. Hence, the benefit of a marginal increase in productivity for a married entrepreneur is smaller than for a unmarried entrepreneur. This is the **spousal substitution effect** we talked about before. All else equal, this pushes married persons to be more selective on which business projects to implement (higher $\underline{z}_w^m$) relative to unmarried individuals. The result is a lower entry rate, higher average productivity and higher average size among firms created by married individuals.

Next, let us consider the situation where the spouse of the married individual only works when the main earner is bankrupt. In this case the value of being an entrepreneur for the married M and unmarried U are given by

$$rJ^m(z) = u(\pi(z)) + \lambda(B^m - J^m(z)) \tag{51}$$

and

$$rJ^u(z) = u(\pi(z)) + \lambda(B^u - J^u(z)). \tag{52}$$

The value of being bankrupt for the married and unmarried would be given by

$$rB^m = u(w - c) + p(W^m - B^m(z)) \tag{53}$$

and

$$rB^u = u((1 - \gamma)(w - c) + (1 - \chi)w^s h) + p(W^u - B^u(z)). \tag{54}$$

Finally the value of being an entrepreneur for the married and unmarried would be given by

$$rW^u = u(w) + \psi \int_{\underline{z}^u} (J^u(z) - W^u)dF(z) \tag{55}$$

and

$$rW^m = u(w) + \psi \int_{\underline{z}^m} (J^m(z) - W^m)dF(z). \tag{56}$$

In this case we see that the only difference between the two individuals is their value of being bankrupt. In particular, for the married individual, bankruptcy is a less painful, because they can rely on the income of their spouse, as long as c is sufficiently large.[12] As a result, when the spouse of the individual only works when the individual is bankrupt the effect is additional insurance for the married individual that decides to become an entrepreneur. This makes entrepreneurship more attractive for married individuals for any productivity level z. This is the **spousal insurance effect** that we talked about before. This channel makes married main earners less selective in which business project to implement

[12]Formally, we need

$$c > w - \frac{w^s(1 - \chi)}{\chi}. \tag{57}$$

This condition is found using the optimal solution for h as a function of the income of the main earner I, here equal to $w - c$. We then find the condition for c to guarantee that

$$(1 - \gamma)(w - c) + (1 - \chi)(w^s h(w^s, w - c) > w - c. \tag{58}$$

(decrease $\underline{z}^m$ relative to $\underline{z}^u$). The result is a higher entry rate and a lower average productivity among firms created by married individuals. Already now we see that there are two effects going in opposite directions. Hence, whether married individuals enter more or less relative to unmarried individuals depends crucially on the strength of these two effects.

Finally let us consider the case where the spouse in a couple only works when the main earner is a wage worker. In this case, the value of being an entrepreneur and being bankrupt between married and unmarried individuals $J(z)$, B would be the same conditional on the value of being a worker. Then the value of being an entrepreneur for the married and unmarried would be written as

$$rJ^m(z) = log[\pi(z)] + \lambda(B^m - J^m(z)) \tag{59}$$

and

$$rJ^u(z) = log[\pi(z)] + \lambda(B^u - J^u(z)). \tag{60}$$

The value of being bankrupt for the married and unmarried is given by

$$rB^m = log[w - c] + p(W^m - B^m(z)) \tag{61}$$

and

$$rB^u = log[w - c] + p(W^u - B^u(z)). \tag{62}$$

Finally the value of being an entrepreneur for the married and unmarried is given by

$$rW^u = log[w] + \psi \int_{\underline{z}^u} (J^u(z) - W^u) dF(z) \tag{63}$$

and

$$rW^m = log[(1-\gamma)w + (1-\chi)w^s h(w, w^s)] + \psi \int_{\underline{z}^m} (J^m(z) - W^m) dF(z). \tag{64}$$

From this set of six equations we see that the only difference between the two type of individuals is that the married individual has a different household income when a worker relative to the unmarried. Any difference in the value of being an entrepreneur or bankrupt must come from this difference in income when working. Furthermore, note that since, $w^s < w$ and $0 \leq h \leq 1$, it follows that $w > w^s h$, $\forall h$. Hence, the married worker has a smaller income relative to the unmarried worker. The reason for this is that the married worker shares a part of their income with their spouse which earns less. Since the income earned as a worker respresents the outside option of the entrepreneur, this represents a higher opportunity cost to entrepreneurship for the unmarried individual. As a consequence, if for married individuals, their spouses only work when they were workers this would imply they would be less selective than unmarried individuals since their opportunity cost to entrepreneurship would be lower. This is our **spousal opportunity cost effect** which we talked about previously. It pushes the married individual to be less selective in which business projects to implement (decreases $\underline{z}_w$). The result is a higher entry rate and a lower average size among firms created by married individuals.

We conclude that relative to unmarried individuals, married people face three additional effects that influence their decision to start a firm. The first is the **spousal opportunity cost effect**. This comes from the fact that when starting a firm, married persons sacrifice a lower income (lower opportunity cost). This comes from the fact that married individuals need to share some of their income with their spouse. In the case where they are earning more then their spouse such as when a wage worker this implies a smaller income for the individual. This in turn represents a lower opportunity cost to entrepreneurship which pushes married pesons to be less selective on which business projects to implement (higher $\underline{z}^m$ relative to unmarried individuals, $\underline{z}^u$).The second effect is what we call the **spousal insurance effect**. This comes from the insurance married individuals have due to the possibility of their spouse working more hours when they go bankrupt. This makes entrepreneurship relatively more attractive to married individuals relative to unmarried persons. This pushes down the selection in business projects among the married (lower $\underline{z}^m$) relative to the unmarried ($\underline{z}^u$). Finally, the **spousal substitution effect** refers to how the slope of the value of being an entrepreneur as a function of productivity z is smaller for married individuals. This is coming from the fact that as a main earner becomes richer the spouse decreases their labor supply partially offsetting the total increase in household income. This effect decreases the average value of entrepreneurship for married persons.

As a consequence they become more selective on which business projects to implement (higher z_w^m).

This tractable framework tells us that whether married individuals enter more or less into entrepreneurship and whether they open larger or smaller firms crucially depends on which of these effects dominates. In the next section I go over the data and the empirical strategy used to test the strength of these different channels. It is important to note that although the model implies an ambiguos response of entry rates and average size of startups to changes in marriage rates and spousal wages it imposes the restriction that any variable that changes the entry rate into entrepreneurship must change the average size of startups in the opposite direction. This comes from the model restriction that we are identifying changes in average startup size due to changes in the selection of business projects upon entry. As shall become clear, this restriction holds in the data.

Before proceeding to the data analysis we consider the endogenous decision to marry. This is important since it will make clear the same source of endogeneity that needs to be overcome to estimate the effect of marriage on entrepreneurship.

2.3 Endogenous decision to marry

Up until now we have considered the marital status of an individual as exogenous. In this section, we formalize the decision of individuals to marry or not. The objective is not to provide a full detailed theory of the formation and dissolution of marriage but rather to inform us how entrepreneurial ability θ and economy wide productivity y affect both marriage and entrepreneurial outcomes. This is an important point given our desire to empirically estimate the effect of marriage on entrepreneurship. As shall be clear in this section any empirical analysis must take into account the endogeneity problem caused by entrepreneurial ability and economy wide productivity.

With this objective in mind we consider a simple form of endogenous marriage formation. Ex-ante, individuals choose to marry based on the expected value of being married and unmarried. In particular, individuals make this choice under the veil of ignorance, before knowledge of whether they enter or not entrepreneurship. They weigh each state by the measure of individuals of same entrepreneurial ability as themselves in each state.

Furthermore, each individual belongs to a group g.[13] Each group g has a idyossincratic utility value of v_g associated to being married. Let $v_g \sim H(v)$. Hence, the ex-ante value of marriage $V^M(\theta, g)$ for an individual of ability θ from group g is

$$V^M \equiv e_m(\theta)W^M(\theta) + b_m(\theta)B^M(\theta) + \int_{\underline{z}^m(\theta)} J(z,\theta)\Gamma_M(z,\theta)dz + v_g. \tag{65}$$

The ex-ante value of being unmarried, $V^U(\theta, g)$ for an individual of ability θ from group g is

$$V^U \equiv e_u(\theta)W^U(\theta) + b_u(\theta)B^U(\theta) + \int_{\underline{z}^u(\theta)} J(z,\theta)\Gamma_U(z,\theta)dz. \tag{66}$$

It follows that an individual of ability θ from group g marries if

$$\begin{aligned} e_m(\theta)W^M(\theta) + b_m(\theta)B^M(\theta) + \int_{\underline{z}^m(\theta)} J(z,\theta)\Gamma_M(z,\theta)dz + v_g > \\ e_u(\theta)W^U(\theta) + b_u(\theta)B^U(\theta) + \int_{\underline{z}^u(\theta)} J(z,\theta)\Gamma_U(z,\theta)dz. \end{aligned} \tag{67}$$

Let $Pr(M = 1)$ be the probability an individual chooses to marry then

$$Pr(M = 1) = Pr(V^M > V^U). \tag{68}$$

Taking a first order log-linearization of the set of value functions for θ and y both around 1 gives us the probability as a linear function of individual entrepreneurial ability, θ, economy specific productivity y and tast for marriage v_g. The proposition below states this formally.

Proposition 1. *The probability to marry $Pr(M = 1)$ is characterized by*

$$Pr(M = 1) = \gamma_0 + \gamma_1 log(\theta) + \gamma_2 log(y) + \gamma_v v_g. \tag{69}$$

[13] Later when we go to the data this will be individuals of different countries of origin.

The proposition above makes explicit that the decision to marry is a function of the entrepreneurial ability of an individual, the economy specific productivity and their taste for marriage. Intuitively, a higher entrepreneurial ability increases the individual's incentive to start a firm which in turn makes the effect of marriage on the value of entrepreneurship more important for these individuals. This intuition also makes clear that entrepreneurial ability affects both the decision to start a firm conditional on marital status as well as the decision to marry. Similarly, in productive economies (with higher y), individuals are more prone to start a firm. But this in turn makes the differences in the value of entrepreneurship between married and unmarried all the more important for this individual. As a result, a change in y affects both the decision to start a firm conditional on marriage but also the decision to marry. It follows that a naive regression of entry into entrepreneurship on marriage suffers from endogeneity due to both y and θ. When we bring the model to the data we implement a strategy to overcome this endogeneity problem. Crucially, we propose an instrument that captures variation in v_g to identify changes in $Pr(M = 1)$ uncorrelated to the decision to start a firm.

When we bring the model to the data we will consider variation across local economies. In particular, we consider each local economy-time period pair to be described by the model layed out in this section. In line with this logic, Proposition 2 below derives the marriage rate of an economy as a function of the weighted sum of the marriage rate of different groups g in the economy. It makes clear that in fact v_g can be captured by this weighted sum of the marriage rate of different groups g in the economy.

Proposition 2. *Suppose there exists a large number of economies c all of which are characterized by the model described in this section. Let $\eta_{j,c,t}$ be the share of individuals of group j in the economy c at year t, $G(\theta)$ be the unconditional distribution of entrepreneurial ability θ in economy c, $G_c(\theta|g)$ be the distribution of entrepreneurial ability θ conditional on being from group g in economy c and $\pi_{c,j,t}$ be the share of individuals of group j located in the economy c at year t. Then, the aggregate marriage rate MR in the* *economy can be written as*

$$MR_{c,t} = \frac{\gamma_1}{\gamma_v}[\gamma_v \int log(\theta)dG_c(\theta) - \sum_{\forall g} \eta_{g,c,t} \int log(\theta)dG_c(\theta|g)] + \frac{\gamma_2}{\gamma_v}[\gamma_v log(y_{c,t}) - \sum_{\forall g} \eta_{c,g,t} \sum_{\forall c} \pi_{c,j,t} log(y_{c,t})] + \sum_{\forall g} \eta_{g,c,t} MR_{g,t} \quad (70)$$

2.4 Empirical Analysis

In this section I go over the main empirical strategies used to disentangle the relative strengths of each channel spouses affect the individual's decision to start a firm. The strategy uses variation across cities. Intuitively, consider each city as a local economy described by our model in the previous section. Then Proposition 2 informs us that a natural candidate for an instrument for marriage rates is any variation in

$$\sum_{\forall g} \eta_{g,c,t} MR_{g,t} \quad (71)$$

which is uncorrelated to heterogeneity in unobserved entrepreneurial ability $\int log(\theta)dG_c(\theta)$ and the economy wide productivity $log(y_{c,t})$.

Furthemore, our source of variation allow us to test the model prediction that the effect of marriage on the entry rate into entrepreneurship must have the opposite sign of the effect of marriage on the average size of startups. This restriction comes directly from our selection mechanism, in which firm heterogeneity is being driven by entry decision of entrepreneurs. To our knowledge this is the first paper to empirically test this restriction of firm selection mechanisms.

We focus on men between 25 and 65 years old to focus on individuals with high labor market attachment. Our main objective is to verify which effect from the theory (**spousal insurance effect**, **spousal opportunity cost effect** and **spousal substitution effect**) is strongest. To verify this we use variation across different cities in Canada. Let c denote city, t year then the entry into entrepreneurship $ER_{c,t}$ can be written as a function of the marriage rate of that city $MR_{c,t}$

Proposition 3. *Suppose there exists a large number of economies c all of which are characterized by the model described in the previous section. Then, the entry rate into entrepreneurship and the average size of startups in each of these c economies can be written as*

$$ER_{c,t} = \gamma_{0,1} + \gamma_{1,1} MR_{c,t} + \epsilon^{e}_{c,t}. \tag{72}$$

and

$$SY_{c,t} = \gamma_{0,2} + \gamma_{1,2} MR_{c,t} + \epsilon^{s}_{c,t}. \tag{73}$$

where

$$\epsilon^{e}_{c,t} = \varepsilon^{f,e}_{c} + \varepsilon^{e}_{c,t} \tag{74}$$

and

$$\epsilon^{s}_{c,t} = \varepsilon^{f,s}_{c} + \varepsilon^{s}_{c,t}. \tag{75}$$

The components $\varepsilon^{f,s}_{c}$ and $\varepsilon^{f,e}_{c}$ represent city specific comparative advantages that induce permanent differences across cities in entrepreneurship. In our model these come from different distributions in entrepreneurial ability across cities $G_c(\theta)$ but fixed over time. $\varepsilon^{e}_{c,t}$ and $\varepsilon^{s}_{c,t}$ represent city specific characteristics not fixed over time that influence entrepreneurship. In our model, these are city specific productivities $y_{c,t}$ that varies over time. The possible presence of different entrepreneurial ability distributions across cities $G_c(\theta)$ (captured by $\varepsilon^{f,s}_{c}$ and $\varepsilon^{f,e}_{c}$) create an endogeneity problem. In order to eliminate this source of endogeneity we take first differences. Finally, to allow for national level trends in the change in entry into entrepreneurship and the average size of startups we include year dummies $\mathbb{1}\{year = t\}_t$. It follows our specifications become

$$\Delta ER_{c,t} = \gamma_{0,1} + \gamma_{1,1} \Delta MR_{c,t} + \mathbb{1}\{year = t\}_t + \Delta\varepsilon^{e}_{c,t} \tag{76}$$

and

$$\Delta SY_{c,t} = \gamma_{0,2} + \gamma_{1,2} \Delta MR_{c,t} + \mathbb{1}\{year = t\}_t + \Delta\varepsilon^{s}_{c,t}. \tag{77}$$

2.4.1 Endogeneity

The specifications above could be estimated with OLS, however they still suffer from endogeneity. In particular, if $\varepsilon^e_{c,t}$ and $\varepsilon^s_{c,t}$ represent city specific productivity $y_{c,t}$ then according to our model we expect this to influence both marriage decisions and decisions to start a firm. Intuitively, individuals that expect to start a firm (due to high $y_{c,t}$) are more likely to marry so as to take advantage of having spousal insurance in case of firm failure. This implies

$$CORR(\Delta MR_{c,t}, \Delta\varepsilon^e_{c,t}) \neq 0 \tag{78}$$

and

$$CORR(\Delta MR_{c,t}, \Delta\varepsilon^s_{c,t}) \neq 0. \tag{79}$$

To resolve this issue we use an instrumental variable strategy for $\Delta MR_{c,t}$ based on our model. In what follows we explain our strategy.

2.4.2 Instrument for Marriage Rates

The instrument we consider uses variation in the distribution of immigrant groups across cities and the variation in marriage rate across these different groups[14]. If we use the expression found for marriage rates $MR_{c,t}$ in Proposition 2 and take first differences we get

$$\Delta MR_{c,t} = \frac{\gamma_2}{\gamma_v}\Delta[\gamma_v log(y_{c,t}) - \sum_{\forall g}\eta_{c,g,t}\sum_{\forall c}\pi_{c,j,t}log(y_{c,t})] + \Delta\sum_{\forall g}\eta_{g,c,t}MR_{g,t}. \tag{80}$$

Now note that

$$\Delta\sum_{\forall g}\eta_{g,c,t}MR_{g,t} \approx \sum_{\forall g}\eta_{g,c,1}\Delta MR_{g,t} + \sum_{\forall g}\Delta\eta_{g,c,t}MR_{g,1} \tag{81}$$

[14]Groups are defined by country of birth of the individuals. Individuals born in Canada are included as a separate group

where $\eta_{g,c,1}$ is the share of individuals of group j in the economy c at the first year of the sample and $MR_{g,1}$ is the marriage rate of group g at the first year of the sample. For our instrument we consider the term

$$\sum_{\forall g} \eta_{g,c,1} \Delta MR_{g,t}. \tag{82}$$

Intuitively, in our model the variation in this instrument comes from variation in the taste for marriage across groups over time, $\Delta v_{g,t}$, and the composition of each city in terms of each group. This instrument is valid as long as

$$\lim_{c\to\infty} \frac{1}{C} \sum_{\forall c} \sum_{\forall g} \eta_{g,c,1} \Delta MR_{g,t} \Delta\varepsilon^{e}_{c,t} = \sum_{\forall g} \Delta MR_{g,t} \lim_{c\to\infty} \frac{1}{C} \sum_{\forall c} \eta_{g,c,1} \Delta\varepsilon^{e}_{c,t} = 0 \tag{83}$$

and

$$\lim_{c\to\infty} \frac{1}{C} \sum_{\forall c} \sum_{\forall g} \eta_{g,c,1} \Delta MR_{g,t} \Delta\varepsilon^{s}_{c,t} = \sum_{\forall g} \Delta MR_{g,t} \lim_{c\to\infty} \frac{1}{C} \sum_{\forall c} \eta_{g,c,1} \Delta\varepsilon^{s}_{c,t} = 0. \tag{84}$$

where we can take out $\Delta MR_{g,t}$ from the sum because it is at the group level.

Sufficient conditions for these are

$$\lim_{c\to\infty} \frac{1}{C} \sum_{\forall c} \eta_{g,c,1} \Delta\varepsilon^{e}_{c,t} = 0 \tag{85}$$

and

$$\lim_{c\to\infty} \frac{1}{C} \sum_{\forall c} \eta_{g,c,1} \Delta\varepsilon^{s}_{c,t} = 0. \tag{86}$$

This happens if city specific productivity changes $\Delta log(y_{c,t})$ are uncorrelated to the group composition of cities at the first year of the sample.[15]

[15] Although not present in our model, we could imagine different groups g have different time varying taste for entrepreneurship. Our identification strategy here allows for the possibility that this taste for entrepreneurship correlates with the taste for marriage. What it doesn't allow is for the change in the group g specific taste for entrepreneurship to be correlated to changes in group g specific taste for marriage.

Finally, as robustness to our main specifications we consider controlling for changes in the share of men $25-40$ years old, $\Delta Share_{c,t}^{25-40}$, the share of men $41-55$ years old, $\Delta Share_{c,t}^{41-55}$, the share of men $56-65$ years old, $Share_{c,t}^{56-65}$, change in the share of children[16], $Share_{c,t}^{children}$ and change in the share of total employment in the current employment in the oil, gas and mining sector $Share_{c,t}^{oil}$.[17] The point estimates are unchanged when adding these controls.[18]

3 Data and Empirical Results

In this section we go over the data we use and our empirical results.

3.1 Data and Measurement

The data used for the empirical analysis is the Canadian Employer-Employee Dynamics Database (CEEDD). It contains the entire universe of Canadian tax filers, and privately owned incorporated firms. The dataset links employees to firms and firms to their corresponding owners across space and time. This is achieved by linking individual tax information (T1 files, individual tax returns), with linked employer-employee information (T4 files)[19] and firm ownership and structure information (T2 files).[20] The data is annual and is available from 2001 to 2010. This constitutes an advantage relative to employer-employee

[16]Individuals with less than 15 years old

[17]This last variable is important, since during the period Canada was experiencing an important oil boom.

[18]Results with the controls are available upon request from the author.

[19]According to Canadian law, each employer must file a T4 file for each of her employees. The equivalent in the US is the W-2, Wage and Tax Statement. In this form, the employer identifies herself, identifies the employee and reports the labour earnings of the employee.

[20]T2 forms are the Canadian Corporate Income Tax forms. In the T2 files there is the schedule 50 in which each corporation must list all owners with at least 10% of ownership. This allows me to link each firm to individual entrepreneurs. The equivalent in the US to the schedule 50 of the T2 form is the schedule G of 1120 form (Corporate Income Tax Form in the US)

firm population data from the US, which does not allow the researcher to identify the owners of the firm.

The data is annual with information on all employers and any businesses an individual owned in a given year. Using this database, I can examine the characteristics of both the business owner and the firm. I concentrate on firms that contribute to job creation by hiring employees. This is done by focusing on employers instead of self-employed individuals.

Business owners are identified as individuals present in the schedule 50 files from the T2 that have employees. Wage workers are identified as those who are not entrepreneurs and report a positive employment income on their T4. I use the information in the T1 files to control for characteristics such as gender, age and marital status. Finally, the dataset is also linked to immigrant landing files allowing me to observe country of birth of the individual.

The linkage between each firm and its corresponding owner is only available for privately owned incorporated firms. Incorporated firms have two key characteristics which correspond closely to how economists typically think about firms : limited liability and separate legal identity. Furthermore, there is a growing literature showing that incorporated firms tend to be larger and that they are more likely to contribute to aggregate employment.[21] There is also evidence that there is little transition from unincorporated to incorporated status.[22] These facts, highlight how incorporated firms with employees are the most appropriate measure of firms to consider if we are interested in the interplay

[21] Glover and Short (2010) document that incorporated entrepreneurs operate larger businesses, accumulate more wealth, and are on average more productive than unincorporated entrepreneurs. Chandler (1977) and Harris (2000) argue that over time the incorporated business structure was created with the explicit goal of fostering investment in large, long gestation, innovative and risky activities.

[22] Levine and Rubinstein (2017) show that there is little transition from unincorporated to incorporated status. They also show that the observed earnings increase for incorporated business owners does not take place before opening the business, indicating that incorporation is not just a result of higher earnings, rather, people choose the firm structure based on their planned business activity. The authors demonstrate how the often cited puzzle, that entrepreneurs earn less than they would have as salaried workers, is no longer true once we consider incorporated business owners. Together with other patterns of income dynamics and observable characteristics of owners, the authors highlight how incorporated businesses are closer to firms in traditional macro models.

between entrepreneurship and the aggregate economy.[23]

For the remainder of the paper, the empirical definition of an entrepreneur is an owner and founder of a privately owned incorporated firm with employees.

Besides the CEEDD we also use data from the Labour Force Survey (LFS) for Canada. We use the LFS to construct industry premia at the national level using hourly wages. These industry premia are used for the construction of our second instrument for marriage rates at the city level. From these datasets we construct city year level objects for our regressions.

3.1.1 Results

In this section I present the main results of our specifications. Column 1 of Table 1 presents the OLS results for our entry rate specification. We see that marriage rates are insignificant if we don't instrument for it. Column 2 for our entry rate specification once we instrument marriage rates by our instrument $IV_{c,t}$. The results indicate that a 1% point increase in the marriage rate is associated to a 0.15% point increase in the entry probability into entrepreneurship. Given the benchmark entry rate of 1%, this corresponds to a 15% increase in the entry probability. The results are robust to including additional controls (changes in the share of men $25-40$ years old, $\Delta Share_{c,t}^{25-40}$, the share of men $41-55$ years old, $\Delta Share_{c,t}^{41-55}$, the share of men $56-65$ years old, $Share_{c,t}^{56-65}$, change in the share of children[24], $Share_{c,t}^{children}$ and change in the share of total employment in the current employment in the oil, gas and mining sector $Share_{c,t}^{oil}$). The point estimate and the significance are unchanged when adding these controls.[25] Finally, Row 2 indicates the results with are robust to including additional controls.

[23]Another reason to focus on incorporated firms with employees is Canadian corporte law. In Canada there are significant tax advantages for incorporating as a higher earner. So to exclude from my analysis high-earning workers that incorporate exclusively due to tax purposes, I focus on incorporated firms with employees.

[24]Individuals with less than 15 years old

[25]Results with the controls are available upon request from the author.

Next, we turn to the results on average size of firms. Column 1 of Table 2 indicates marriage, $\Delta MR_{c,t}$ has no effect on the average size of startups in the city if not instrumented for. Column 2 of Table 2 shows results for average size of startups once we instrument marriage rates by our instrument $IV_{c,t}$. The results indicate that a 1% point increase in the marriage rate is associated to a 15% decrease in the average size of startups. The results are robust to including additional controls (changes in the share of men $25 - 40$ years old, $\Delta Share_{c,t}^{25-40}$, the share of men $41 - 55$ years old, $\Delta Share_{c,t}^{41-55}$, the share of men $56-65$ years old, $Share_{c,t}^{56-65}$, change in the share of children[26], $Share_{c,t}^{children}$ and change in the share of total employment in the current employment in the oil, gas and mining sector $Share_{c,t}^{oil}$). The point estimate and the significance are unchanged when adding these controls.[27] Finally, Row 2 indicates the results with are robust to including additional controls.

These results are consistent with the prediction of our model that the group (married vs unmarried) with the highest entry rate into entrepreneurship has the lower average size of startups.[28]

[26]Individuals with less than 15 years old

[27]Results with the controls are available upon request from the author.

[28]These opposing predictions for entry and average size are an immediate consequence of the selection mechanism in our model. Hence, the empirical results are consistent with difference between married and unmarried individuals being driven by difference in selection upon entry into entrepreneurship.

Table 1: Main specifications

	OLS	IV
$\Delta MR_{c,t}$	0.002	0.157^{***}
	(0.007)	(0.059)
Significance IV for $\Delta MR_{c,t}$	-	Yes
Observations	836	836

Notes: Regressions of changes in the entry rate into entrepreneurship in economic region c and year t, $\Delta ER_{c,t}$, on the change in the marriage rate in economic region c year t, $\Delta MR_{c,t}$. Column 1 reports OLS results. Column 2 reports results when using our instrument for marriage rates. Our instrument is the sum of the share of individuals from group g in a economic region c in the first year of the sample multiplied by the change in marriage rate of individual at group g at the national level in year t. Both specifications include year dummies to capture national trends. * represents 10% significance, ** represents 5% significance and *** represents 1% significance. Standard errors are clustered at the economic region c level.

Table 2: Main specifications

	OLS	IV
$\Delta MR_{c,t}$	-1.93	-15.46^{**}
	(2)	(7.51)
Significance IV for $\Delta MR_{c,t}$	-	Yes
Observations	836	836

Notes: Regressions of changes in the average number of employees of startups in economic region c and year t, $\Delta ER_{c,t}$, on the change in the marriage rate in economic region c year t, $\Delta MR_{c,t}$. Startups are defined as firms that are at most 1 year old. Column 1 reports OLS results. Column 2 reports results when using our instrument for marriage rates. Our instrument is the sum of the share of individuals from group g in a economic region c in the first year of the sample multiplied by the change in marriage rate of individual at group g at the national level in year t. Both specifications include year dummies to capture national trends. * represents 10% significance, ** represents 5% significance and *** represents 1% significance. Standard errors are clustered at the economic region c level.

4 Discussion on Alternative Mechanisms

The theoretical model is purposely tractable to keep the intuition clear and concise. However, there are other economic mechanisms affecting entrepreneurship in real life not present in the model. In particular, spouse might help individuals overcome borrowing constraints via wealth sharing or faster wealth accumulation.

In this line of thought, there are two ways borrowing constraints affects outcomes for an individual considering starting a firm. The first is that individuals are constrained in the scale of the firm they create. If married individuals on average are wealthier we expect higher marriage rates to be associated to higher average size of startups. This is the opposite of what we find. The second way borrowing constraints might matter is if there are substantial startup costs to opening a firm. This is an alternative narrative for our results that marriage rates induce higher entry rates into entrepreneurship. To verify our findings are not being driven by this channel we check that our results are robust to excluding business starts in high capital demanding sectors.[29] This robustness check alleviates the concern that our results are being driven by borrowing constraint concerns.

Another potential mechanism ignored by the model is the possibility of joint entrepreneurship by both main earner and spouse. In particular, married individuals can be more likely to start a firm because they have the option of starting with their spouse which also contributes to running the business. Of course this is inherently hard to measure. But to the extent that a couple running together a firm are both listed as owners of the business or has the spouse as employee of the firm, we can verify our results are robust to excluding these types of businesses. Indeed, we verify that our results continue to hold, in both magnitude and significance once we exclude firms started jointly by both main earner and spouse and firms for which the spouse is listed an employee.

Both of these robustness checks are available upon request from the author.

5 Implications for Job Creation

Given our results for the response of average size of startups and entry rate into entrepreneurship to marriage rates, we can calculate the resulting change in job creation due to a change in marriage rates. In particular, let $E^s_{c,t}$ represent total employment among startups at city c year t, then

$$log(E^s_{c,t}) = log(ER_{c,t}) + log(SY_{c,t}). \tag{87}$$

[29]In particular, we exclude entries into Oil, Mining, Gas, Utilities, Manufacturing, Wholesale Trade, Retail Trade, Transportation and Wharehousing, Information and Cultural Industries and Real Estate.

It follows that

$$\frac{\Delta log(E^s_{c,t})}{\Delta MR_{c,t}} = \frac{\Delta log(ER_{c,t})}{\Delta MR_{c,t}} + \frac{\Delta logSY_{c,t}}{\Delta MR_{c,t}} \quad . \tag{88}$$

Finally, note that $\frac{\Delta logSY_{c,t}}{\Delta MR_{c,t}}$ has been estimated (equal to 15.46 according to Table 2) and

$$\frac{\Delta log(ER_{c,t})}{\Delta MR_{c,t}} = \frac{\Delta ER_{c,t}}{\Delta MR_{c,t}} \cdot \frac{1}{ER_{c,t}} \quad . \tag{89}$$

Using our estimate of 0.157 for $\frac{\Delta ER_{c,t}}{\Delta MR_{c,t}}$ and the average entry rate of 0.01 to calculate $\frac{1}{ER_{c,t}} = \frac{1}{0.01}$ gives

$$\frac{\Delta log(ER_{c,t})}{\Delta MR_{c,t}} = 15.7 \quad . \tag{90}$$

Hence, we conclude that for a 1% point increase in the marriage rate total employment among startups increases by 0.24% points.[30] They key take away from this exercise is the importance of taking into account the selection mechanism when talking about job creation. If a researcher had ignored the selection mechanism, inducing a fall in average size of startups, they might conclude wrongly that a 1% increase in the marriage rate induces a 15.7% increase in employment. This point is more general than the particular mechanism of this paper of marriage and entrepreneurship. Futhermore, this highlights as well the importance of being explicit in the form of endogeneity and the solution to it. If a researcher ignored the endogeneity problem, they might wrongly infer that job creation responds little to changes in the marriage rate as a result of small responses of both entry and average size of startups.[31] Here, on the other hand, the small response of job creation comes exactly because of the strong negative relationship between average size and entry generated by the selection mechanism.

[30]This just comes from $15.7\% - 15.46\%$.

[31]Tables 1 and 2 tell us that if we ignore the endogeneity problem we would conclude that the effect of marriage on both entry and average size is not stastifcally different than zero.

6 Conclusion

In this paper we explore for the first time the importance of spouses for an individdual's decision to start a firm. Through a tractable model we show that spouses affect the decision through different opposing channels. Using this same model we derive empirical specifications and a instrumental variable that allow us to bring the model to the data. We show that marriage is associated to larger entry into entrepreneurship and lower average size of startup. This is consistent with the importance of the insurance provided by the spouse in case of business failure. We also verify our results are robust to alternative mechanisms such as borrowing constraints or spouses working in the business.

References

Alon, Titan, David Berger, Robert Dent, and Benjamin Pugsley, "Older and slower: The startup deficit's lasting effects on aggregate productivity growth," *Journal of Monetary Economics*, 2018, *93*, 68–85.

Angeletos, George-Marios and Vasia Panousi, "Revisiting the supply side effects of government spending," *Journal of Monetary Economics*, 2009, *56* (2), 137–153.

_ and _ , "Financial integration, entrepreneurial risk and global dynamics," *Journal of Economic Theory*, 2011, *146* (3), 863–896.

Beaudry, Paul, David A Green, and Benjamin M Sand, "In search of labor demand," *American Economic Review*, 2018, *108* (9), 2714–57.

_ , _ , and Benjamin Sand, "Does industrial composition matter for wages? a test of search and bargaining theory," *Econometrica*, 2012, *80* (3), 1063–1104.

Blundell, Richard, Alan Duncan, and Costas Meghir, "Estimating labor supply responses using tax reforms," *Econometrica*, 1998, pp. 827–861.

_ , Luigi Pistaferri, and Itay Saporta-Eksten, "Consumption inequality and family labor supply," *American Economic Review*, 2016, *106* (2), 387–435.

_ , _ , and _ , "Children, time allocation, and consumption insurance," *Journal of Political Economy*, 2018, *126* (S1), S73–S115.

Cagetti, Marco, Mariacristina De Nardi et al., "Entrepreneurship, frictions, and wealth," *Journal of Political Economy*, 2006, *114* (5), 835–870.

Cao, Shutao, Mohanad Salameh, Mai Seki, Pierre St-Amant et al., "Trends in Firm Entry and New Entrepreneurship in Canada," Technical Report 2015.

Chandler, Alfred Dupont, *The visible hand: The Managerial Revolution in American Business*, Cambridge, MA : Harvard University Press, 1977.

Clementi, Gian Luca and Berardino Palazzo, "Entry, exit, firm dynamics, and aggregate fluctuations," *American Economic Journal: Macroeconomics*, 2016, *8* (3), 1–41.

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda, "The role of entrepreneurship in US job creation and economic dynamism," *The Journal of Economic Perspectives*, 2014, pp. 3–24.

Diez, Federico J et al., "The Great Recession, entrepreneurship, and productivity performance," Technical Report, Federal Reserve Bank of Boston 2014.

Fonseca, Raquel, Paloma Lopez-Garcia, and Christopher A Pissarides, "Entrepreneurship, start-up costs and employment," *European Economic Review*, 2001, *45* (4), 692–705.

Gallipoli, Giovanni and Laura Turner, "Household responses to individual shocks: Disability and labor supply," Technical Report 2009.

Glover, Andy and Jacob Short, "Bankruptcy, Incorporation, and the Nature of Entrepreneurial Risk," 2010.

Green, David A, Rene Morissette, Benjamin M Sand et al., "Economy Wide Spillovers From Booms: Long Distance Commuting and the Spread of Wage Effects," Technical Report 2017.

Haltiwanger, John, "Job Creation and firm dynamics in the US," *Unpublished manuscript, May*, 2011.

_ , Ron S Jarmin, and Javier Miranda, "Who creates jobs? Small versus large versus young," *Review of Economics and Statistics*, 2013, *95* (2), 347–361.

_ , Ryan Decker, and Ron Jarmin, "Top ten signs of declining business dynamism and entrepreneurship in the US," in "Kauffman Foundation New Entrepreneurial Growth Conference" Citeseer 2015.

Hamilton, Barton H, "Does entrepreneurship pay? An empirical analysis of the returns to self-employment," *Journal of Political Economy*, 2000, *108* (3), 604–631.

Harris, Ron, *Industrializing English law: entrepreneurship and business organization, 1720-1844*, Cambridge University Press, 2000.

Hathaway, Ian and Robert E Litan, "Declining business dynamism in the United States: A look at states and metros," *The Brookings Institution*, 2014.

Holtz-Eakin, Douglas, David Joulfaian, and Harvey S Rosen, "Sticking it Out: Entrepreneurial Survival and Liquidity Constraints," *Journal of Political Economy*, 1994, pp. 53–75.

Hurst, Erik and Annamaria Lusardi, "Liquidity constraints, household wealth, and entrepreneurship," *Journal of Political Economy*, 2004, *112* (2), 319–347.

Hyslop, Dean R, "Rising US earnings inequality and family labor supply: The covariance structure of intrafamily earnings," *American Economic Review*, 2001, *91* (4), 755–777.

Jr, Melvin Stephens, "Worker displacement and the added worker effect," *Journal of Labor Economics*, 2002, *20* (3), 504–537.

Jr, Robert E Lucas, "On the size distribution of business firms," *The Bell Journal of Economics*, 1978, pp. 508–523.

Juhn, Chinhui and Simon Potter, "Is there still an added-worker effect?," 2007.

Levine, Ross and Yona Rubinstein, "Smart and illicit: who becomes an entrepreneur and do they earn more?," *The Quarterly Journal of Economics*, 2017, *132* (2), 963–1018.

Lundberg, Shelly, "The added worker effect," *Journal of Labor Economics*, 1985, *3* (1, Part 1), 11–37.

Maloney, Tim, "Employment constraints and the labor supply of married women: A reexamination of the added worker effect," *Journal of Human Resources*, 1987, pp. 51–61.

Panousi, Vasia, "Capital Taxation with Entrepreneurial Risk," 2008.

Pugsley, Benjamin Wild and Aysegül Sahin, "Grown-up business cycles," *The Review of Financial Studies*, 2018, *32* (3), 1102–1147.

Quadrini, Vincenzo, "Entrepreneurship, saving, and social mobility," *Review of Economic Dynamics*, 2000, *3* (1), 1–40.

Sand, Ben, Chris Bidner et al., "Job Prospects and Pay Gaps: Theory and Evidence on the Gender Gap from US Cities," Technical Report 2016.

Tschopp, Jeanne, "The Wage Response to Shocks: The Role of Inter-Occupational Labour Adjustment," *Labour Economics*, 2015, *37*, 28–37.

Wu, Chunzan and Dirk Krueger, "How Much Consumption Insurance in Bewley Models with Endogenous Family Labor Supply?," Technical Report, National Bureau of Economic Research 2018.

Proof of Proposition 1.

For the first part of this proof I show how to obtain an equation defining thresholds $\underline{z}^M$ and $\underline{z}^U$.

Define u_S^x as the flow utility an individual receives when married or unmarried ($x \in \{u, m\}$) and in State S, where $S \in \{W, B, J)$ and $u_J^x(z)$ represents the value of being $i \in \{u, m\}$ when running a firm of productivity z. Now note the fact that, in equilibrium, $\underline{z}^M$ and $\underline{z}^U$ are defined by

$$J^U(\underline{z}^U) = W^U \tag{91}$$

$$J^W(\underline{z}^W) = W^M. \tag{92}$$

We can rewrite value functions $W^x, J^x(z), B^x$ as

$$J^x(z) = \frac{u_J^i(z) + \lambda B^x}{r + \lambda} \quad \forall x \in \{u, m\}. \tag{93}$$

$$B^x = \frac{u_B^i + pW^x}{r + p} \quad \forall x \in \{u, m\}. \tag{94}$$

$$W^x = \frac{u_W^x + \psi \int_{\underline{z}^x} J^x(z) dF(z)}{r + \psi(1 - F(\underline{z}^x))} \tag{95}$$

Using the expression for $J^x(z)$ and using $J^x(\underline{z}^x) = W^x$ we get

$$0 = J^x(\underline{z}^x) - W^x = \frac{u_J^x(\underline{z}^x)}{r + \lambda} + \frac{\lambda u_B^x}{(r + \lambda)(r + p)} - \frac{rW^x(r + p + \lambda)}{(r + p)(r + \lambda)} \tag{96}$$

Now after some algebra we find

$$\begin{aligned} rW^x &= \frac{(r+p)(r+\lambda)u_W^x}{(r+\lambda)(r+p) + r(r+\lambda+p)\psi(1-F(\underline{z}^x))} \\ &+\psi \int_{\underline{z}^x} \frac{u_J^x(z)(r+p)dF(z)}{(r+\lambda)(r+p) + r(r+p+\lambda)\psi(1-F(\underline{z}^x))} + \frac{\psi(1-F(\underline{z}^x))\lambda u_B^x}{r(r+p)(r+\lambda) + r(r+\lambda+p)\psi(1-F(\underline{z}^x))}. \end{aligned} \tag{97}$$

If we replace this expression for rW^x in Equation 109 we find equations defining each treshold $\underline{z}^x$ as a function of parameters and y and θ. Using these optimal expression for each threshold we can log-linearize $\underline{z}^M$ and $\underline{z}^U$, y and θ around the point $(z^*, z^*, 1, 1)$ which gives

$$log(\underline{z}^M) = \zeta_0^M + \zeta_1^M log(y) + \zeta_2^M log(\theta) \tag{98}$$

$$log(\underline{z}^U) = \zeta_0^U + \zeta_1^U log(y) + \zeta_2^U log(\theta). \tag{99}$$

Next, we log-linearize $Pr(M = 1)$ for $\underline{z}^M, \underline{z}^U, y, \theta, v_g$ around the point $(z^*, z^*, 1, 1, 0)$. Finally, we replace $log(\underline{z}^M)$ and $log(\underline{z}^U)$ by their expressions given by Equations (98) and (99) to arrive at the desired result.

$$Pr(M = 1) = \gamma_0 + \gamma_1 log(\theta) + \gamma_2 log(y) + \gamma_v v_g. \tag{100}$$

Proof of Proposition 2.

Let there be a large number of economies c characterized by the model described in the paper. Let $\eta_{j,c,t}$ be the share of individuals of group j in the economy c at year t, $G(\theta)$ be the unconditional distribution of entrepreneurial ability θ in economy c, $G_c(\theta|g)$ be the distribution of entrepreneurial ability θ conditional on being from group g in economy c and $\pi_{c,j,t}$ be the share of individuals of group j located in the economy c at year t. Then if we aggregate Equation (69) at the economy wide level we obtain

$$MR_{c,t} = \gamma_0 + \gamma_1 \int log(\theta) dG_c(\theta) + \gamma_2 log(y_{c,t}) + \gamma_v \sum_{\forall g} \eta_{g,c,t} v_{g,t} \tag{101}$$

If we aggregate Equation (69) at the group g level we obtain

$$MR_{g,t} = \gamma_0 + \gamma_1 \int log(\theta) dG_c(\theta|g) + \gamma_2 \sum_{\forall c} \pi_{c,g,t} log(y_{c,t}) + \gamma_v v_{g,t} \tag{102}$$

Next, use Equation (102) to isolate an expression for $v_{g,t}$ and replace it on Equation (101) to obtain

$$MR_{c,t} = \frac{\gamma_1}{\gamma_v}[\gamma_v \int log(\theta)dG_c(\theta) - \sum_{\forall g} \eta_{g,c,t} \int log(\theta)dG_c(\theta|g)] + \frac{\gamma_2}{\gamma_v}[\gamma_v log(y_{c,t}) - \sum_{\forall g} \eta_{c,g,t} \sum_{\forall c} \pi_{c,j,t} log(y_{c,t})] + \sum_{\forall g} \eta_{g,c,t} MR_{g,t}. \quad (103)$$

Proof of Proposition 3.

For the first part of this proof I show how to obtain an equation defining thresholds $\underline{z}^M$ and $\underline{z}^U$. For this first part I omit subscripts c for economy c, t for time and i for individual to avoid a heavy notation. After this first part, I use the appropriate subscripts to denote what varies at individual level and what varies at the economy level. Each economy c differs exclusively in their aggregate productivity component y and the distribution of entrepreneurial ability $G(\theta)$.

Define u^x_S as the flow utility an individual receives when married or unmarried ($x \in \{u, m\}$) and in State S, where $S \in \{W, B, J\}$ and $u^x_J(z)$ represents the value of being $i \in \{u, m\}$ when running a firm of productivity z. Now note the fact that, in equilibrium, $\underline{z}^M$ and $\underline{z}^U$ are defined by

$$J^U(\underline{z}^U) = W^U \quad (104)$$

$$J^W(\underline{z}^W) = W^M. \quad (105)$$

We can rewrite value functions W^x, $J^x(z)$, B^x as

$$J^x(z) = \frac{u^i_J(z) + \lambda B^x}{r + \lambda} \quad \forall x \in \{u, m\}. \quad (106)$$

$$B^x = \frac{u^i_B + pW^x}{r + p} \quad \forall x \in \{u, m\}. \quad (107)$$

$$W^x = \frac{u^x_W + \psi \int_{\underline{z}^x} J^x(z)dF(z)}{r + \psi(1 - F(\underline{z}^x))} \quad (108)$$

Using the expression for $J^x(z)$ and using $J^x(\underline{z}^x) = W^x$ we get

$$0 = J^x(\underline{z}^x) - W^x = \frac{u_J^x(\underline{z}^x)}{r+\lambda} + \frac{\lambda u_B^x}{(r+\lambda)(r+p)} - \frac{rW^x(r+p+\lambda)}{(r+p)(r+\lambda)} \tag{109}$$

Now after some algebra we find

$$\begin{aligned} rW^x &= \frac{(r+p)(r+\lambda)u_w^x}{(r+\lambda)(r+p) + r(r+\lambda+p)\psi(1-F(\underline{z}^x))} \\ &+\psi \int_{\underline{z}^x} \frac{u_J^x(z)(r+p)dF(z)}{(r+\lambda)(r+p) + r(r+p+\lambda)\psi(1-F(\underline{z}^x))} + \frac{\psi(1-F(\underline{z}^x))\lambda u_B^x}{r(r+p)(r+\lambda) + r(r+\lambda+p)\psi(1-F(\underline{z}^x))}. \end{aligned} \tag{110}$$

If we replace this expression for rW^x in Equation 109 we find equations defining each treshold $\underline{z}^x$ as a function of parameters and y and θ. Using these optimal expression for each threshold, for each household i in economy c we can log-linearize $\underline{z}_{i,c,t}^M$, $\underline{z}_{i,c,t}^U$, $y_{c,t}$ and θ_i around the point $(z^*, z^*, 1, 1)$ which gives

$$log(\underline{z}_{i,c,t}^M) = \zeta_0^M + \zeta_1^M log(y_{c,t}) + \zeta_2^M log(\theta_i) \tag{111}$$

$$log(\underline{z}_{i,c,t}^U) = \zeta_0^U + \zeta_1^U log(y_{c,t}) + \zeta_2^U log(\theta_i). \tag{112}$$

Now recall the expression for the entry rate in the economy is given by

$$ER_{c,t} = MR_{c,t} \int \psi(1-F(\underline{z}^M(\theta)_{i,c,t}))dG_c(\theta) + (1-MR_{c,t}) \int \psi(1-F(\underline{z}^U(\theta)_{i,c,t}))dG_c(\theta) \tag{113}$$

and the one for average size of startups is given by

$$\begin{aligned} SY_{c,t} = MR_{c,t} \int \int_{\underline{z}^M(\theta)_{i,c,t}} (\frac{\alpha y_{c,t}\theta_i e^z}{w})^{\frac{1}{1-\alpha}} \frac{f(z)}{1-F(\underline{z}^M(\theta)_{i,c,t})} dz dG_c(\theta) \\ + (1-MR_{c,t}) \int \int_{\underline{z}^U(\theta)_{i,c,t}} (\frac{\alpha y_{c,t}\theta_i e^z}{w})^{\frac{1}{1-\alpha}} \frac{f(z)}{1-F(\underline{z}^U(\theta)_{i,c,t})} dz dG_c(\theta). \end{aligned} \tag{114}$$

Next, log linearize both these expression for $(\underline{z}_{i,c,t}^{M}, \underline{z}_{i,c,t}^{U}, y_{c,t}, \theta_i, MR_{c,t})$ around $(z^*, z^*, 1, 1, \hat{M})$. Then replace $log(\underline{z}_{i,c,t}^{M})$ and $log(\underline{z}_{i,c,t}^{U})$ by their expressions given by Equations (111) and (112). Finally, aggregate the expressions to the city level to obtain

$$ER_{c,t} = \gamma_{0,1} + \gamma_{1,1} MR_{c,t} + \epsilon_{c,t}^{e}. \tag{115}$$

and

$$SY_{c,t} = \gamma_{0,2} + \gamma_{1,2} MR_{c,t} + \epsilon_{c,t}^{s}. \tag{116}$$

where

$$\epsilon_{c,t}^{e} = \varepsilon_{c}^{f,e} + \varepsilon_{c,t}^{e} \tag{117}$$

and

$$\epsilon_{c,t}^{s} = \varepsilon_{c}^{f,s} + \varepsilon_{c,t}^{s}. \tag{118}$$

In particular,

$$\varepsilon_{c,t}^{e} = \gamma_{1,y} log(y_{c,t}) \tag{119}$$

$$\varepsilon_{c,t}^{2} = \gamma_{2,y} log(y_{c,t}) \tag{120}$$

and

$$\varepsilon_{c}^{f,e} = \gamma_{\theta,1} \int log(\theta) dG_c(\theta) \tag{121}$$

$$\varepsilon_{c}^{f,2} = \gamma_{\theta,2} \int log(\theta) dG_c(\theta). \tag{122}$$