

(WHY) DO SELF-EMPLOYED PARENTS HAVE MORE CHILDREN?

NZINGA BROUSSARD
RALPH CHAMI
GREGORY D. HESS

CESIFO WORKING PAPER NO. 1103
CATEGORY 4: LABOUR MARKETS
DECEMBER 2003

An electronic version of the paper may be downloaded

- *from the SSRN website:* www.SSRN.com
- *from the CESifo website:* www.CESifo.de

(WHY) DO SELF-EMPLOYED PARENTS HAVE MORE CHILDREN?

Abstract

We provide a theory whereby non-benevolent, self-employed households increase their expected family size to raise the likelihood that an inside family member will be a good match at running the business. Hence, having larger family sizes raises the self-employed household's expected return to their business. Using data from the General Social Survey, we find that respondents have approximately .2 to .4 more actual and expected number of children if they are self-employed as compared to if they are not self-employed. This empirical relationship is established across a broad array of sub-samples using a simple differences in means test. As well, the empirical relationship holds using a regression framework, including the use of instrumental variables estimation to allow for the possibility of endogeneity of the respondent's self-employment status and whether the respondent's spouse stays at home.

JEL Classification: D10, D13

Keywords: self-employed, children, family business, matching.

Nzinga Broussard
University of Michigan
Department of Economics
Ann Arbor, MI 48109
USA

Ralph Chami
International Monetary Fund Institute
IMF – Middle East Division
Washington, D.C. 20431
USA

Gregory D. Hess
Department of Economics
Claremont McKenna College
500 E. 9th Street
Bauer Center
Claremont, CA 91711
USA
ghess@mckenna.edu

We thank Bryan Engelhardt for excellent research assistance, and the Reed Institute for Decision Sciences at Claremont McKenna College for Broussard's funding. We also thank Heather Antecol, Kelly Bedard, Joshua Rosett, Kwanho Shin and seminar participants at Claremont McKenna College and Korea University for their comments. Part of this paper was written while Hess was a visiting scholar at the Federal Reserve Bank of Cleveland and the IMF Institute. The opinions expressed are those of the authors and do not necessarily reflect views of the Federal Reserve Bank of Cleveland, the Federal Reserve System, or the IMF.

1 Introduction

It is well known that agents facing risk and uncertainty on their own find it optimal to seek arrangements to mitigate that risk. These arrangements could be market based, such as insurance schemes, which could be privately or publicly-provided. But, it is also well known that such schemes are never complete, as agency problems and other types of frictions preclude the provision of full insurance. As a result, agents continue to seek other forms of insurance, despite the exclusivity clauses in many of the formal insurance contracts. These other forms of arrangements, which are typically referred to as non-market arrangements, fall outside the market-based ones, and are used to complete or to substitute for the presence or lack thereof of market based insurance schemes. For example, family co-insurance arrangements, which include *inter-vivos* transfers as well as bequests, familial loans and other forms of financial assistance, are typically geared to assist family members in need – for example, see Cox (1987), Arnott and Stiglitz (1991), Altig and Davis (1992), and Bernheim and Kolkoff (2001).

Implicit in these non-market arrangements, however, is the assumption that somehow the agents involved in the co-insurance arrangements find it optimal to adhere to their promises, which are normally made *ex-ante*. That is, it is possible that promises of help made before an accident, are reneged on after the fact. Or, the presence of such transfers may induce moral hazard type problems, leading to the failure of such arrangements. Thus, the question is what mitigates such situations? Arnott and Stiglitz (1991) suggest ‘peer monitoring’ as one mechanism through which agency problems among co-insuring agents can be mitigated. Chami and Fischer (1996) show that trust may preclude the need for costly monitoring, as trust does away with the gains from such type of externalities. But trust or mutual caring are only important here in the sense that they insure that agents through such sentiments can find a match, that is a person who shares their views, ideals, or their preferences, which would help reduce the agency problem, and preclude costly negotiations or the need for monitoring. But, the question now is how can one be sure of finding that agent who shares similar preferences?

In this paper, we explore this insight further in the context of self-employed agents. Such individuals present a natural example of individuals facing risk on their own, and who are in need of

finding ways of mitigating such needs. One way could be through marriage. Recently, Chami and Hess (2002) and Hess (2004) explore theoretical and empirical aspects of individuals who attempt to offset idiosyncratic risk to their incomes by marrying hedges.

Another way in which risks to a family could be offset would be through procreation! The latter is easier to see in the context of a farmer facing risk, and where having children help mitigate that risk. Typically, the explanation given is that farmers view children as assets. They help in farming, among the other tasks that are not easily market substitutable. Indeed, Dasgupta (1995) provides a number of examples linking the relationship between population growth and economic incentives and activity in developing economies.

There is, however, another compelling reason as to why farmers, like other self-employed agents may decide to have more children. Farmers, like other family-business owners, share a very important concern, unlike other economic agents, who are employed in the market place. They are concerned with insuring that the farm or family-owned business continues and prospers, even after they are gone. Thus, succession is a major factor that motivates founders to choose a single child to inherit the business. Having more children, despite its costs, would go a long way to finding that person who shares the parents love for the business, or in other words, is as close as possible to a perfect match. It is estimated that in the US and most western economies, more than 75 % of all family firms are transferred to one child (see Gersick et. al. 1997), and that number is much higher in developing economies. Thus the business is not looked on as wealth to be shared equally among the siblings. In many cases, the parent chooses one child to run the business, where the other members are given side payments, but are not involved in managing the business – e.g., see the famous Vanderbilt case described in Clark (1966). Moreover, this motive for choosing an heir is not altruistically motivated; rather it is purely predicated on the presumption that the child with the closest match to the parent, would ensure the continued success of the business. Such a move would, in turn, maximize the value of the firm to the parent/owner, and perhaps even to his or her children if a bequest is made.

The outline of the paper is as follows. In section 2 we present a theory whereby self-employed households choose their consumption plans as well as their family size, and we empirically analyze

the predictions in Section 3. More specifically, in sub-Section 3.1 we describe the data used in this study, and in sub-Section 3.2 we provide simple tests of the differences in means and regression analysis to test whether self-employed parents have more children. As well, we explore in sub-Section 3.3 some additional issues that might affect our results on self-employment status and family size.

2 Theory

We now consider a model of a self-employed individual's decision to have children. In particular, we consider a theory based on a parent's non-benevolent behavior.¹ In the model, every individual born is endowed with an idea in period 0. Some are promising, while others are not. Let s be the percentage of the population that is endowed with a good idea, and $1 - s$ be the fraction of the population that is not. Assume that those endowed without promising ideas can derive a living of \bar{V} from a firm during their period 1 working lives. Such income is, for simplicity, assumed to be certain. Consumption in period 1 is supported by income less net-savings and the costs of child rearing, while net-savings supports consumption in period 2. More specifically, consumption in periods 1, c_1 , and 2, c_2 , equal:

$$c_1 = \bar{V} - a - \alpha \cdot n \quad \text{and} \quad c_2 = a$$

where n is the number of children and α is the per-child cost of raising a child. For simplicity, the rate of return on net-savings (a) is set equal to zero.

Alternatively, those born with a promising idea can choose to start their own firms and reap a potential surplus. Unfortunately, the potential rewards from a start-up business with promising ideas do not come free – it requires a start-up expense to be paid in period one, and the value of

¹ The assumption of non-benevolence is made for simplicity. The model's main prediction, namely that self-employed parents have more children, can be shown to work with a standard model of benevolence to their children. This is so because the key mechanism, whereby having more children raises the parent's expected value of the business through an increased likelihood of a good match, remains in place even with benevolence. Indeed, benevolence may magnify the effect as the enhanced financial prospects will raise the parents per-child gift. Mulligan (1997) examines the interaction between the parent's altruism and fertility choice, in a model where both fertility and altruism are endogenous – also see the earlier work by Barro and Becker (1989).

the good idea is not known until period 2 when the business must be sold to provide the individual with retirement income. Let $\omega > 0$ be the start-up cost of a new business, and let \tilde{V} be the uncertain value of his promising idea in period 1 when of working age. To make matters simple, let the business yield \bar{V} in period 1 with a potential surplus in period 2.² Period 1 consumption is supported by business income less any net savings and the cost of having children, namely:

$$c_1 = \bar{V} - \omega - a - \alpha \cdot n$$

In the second period, the self-employed business person attempts to support his second period consumption by selling the potential surplus created from his business in the second period to a member of the next generation – e.g. a child. But at what price? Let ϕ be the true value of the business’s surplus worth if the business owner finds the right person who sees the value. However, as noted above, the key to a family business is that with the insider’s valuation of the business being private information, it is unlikely that an outsider’s valuation of the business’s surplus is worth more than 0. Why? Outsiders know that the family has private information about the business that may not be currently observed or easily quantified.³ As a consequence, outsiders will not want to pay for the business’s surplus – a premium for the business. Of course, an inside family member may also not see the true value of the firm. To capture this phenomenon, let us assume that with probability $1 - p$ the self-employed business owner does not find an insider or outsider match that sees the full value of the business. Hence, if the business owner does not find a match, the value of his firm’s surplus is 0 in period 2. Alternatively, with probability p the business owner may find an insider or an outsider who has the skills needed to see the business’s full potential or true value. Consistent with observed facts, we assume that the business owner sells it to just one individual and that the owner and purchaser share the surplus. Since the individual who buys the business will also be endowed with some ideas of his own, the sale price of the business is $\phi\lambda$, where λ is the fraction of the surplus that the firm owner keeps.

²Note that it must be incentive compatible for the individual with a promising idea to be willing to undertake a start-up. This can be achieved by lowering the start-up cost, ω relative to the expected surplus, $\lambda\phi$, described below.

³While the policymakers are currently enthralled with accounting irregularities in large publicly traded companies, the problem could be much worse in privately-held, family businesses – e.g. chaebols in South Korea. Tax evasion and under-the-table payments may also be a formidable factor for self-employed businesses. As such, the market value of these privately held firms may be extremely difficult for outsiders to determine, though less so for family insiders.

To recap, there are two possible outcomes for the business owner in period 2. If he has a promising idea but does not match with someone who sees the value, then $c_2^{NM} = a$. Finally, if he has a promising idea and finds a match then $c_2^M = a + \phi\lambda$. Taken together, the self-employed households decision is to choose the number of children and the amount of net savings to maximize welfare:

$$\max_{\{n,a\}} W = \theta \cdot n + U(c_1) + (1 - p) \cdot U(c_2^{NM}) + p \cdot U(c_2^M) \quad (1)$$

where θ represents a household's exogenous and random preference for children, $U'(\cdot) > 0$, $U''(\cdot) < 0$ and $c_2^M > c_2^{NM}$.

As mentioned above, these potential matches and the sale of a business are available only to the fraction of the population, s , that are born with promising ideas. By comparison, the fraction of the population $1 - s$ who are born without promising ideas can be thought of as individuals where the probability p is zero of ever finding a promising match for expression (1). Notice that since starting a new business involves a start-up cost, and since those who start out working lives without promising ideas do not have a chance to reap any surplus, only those born with promising ideas will attempt to start a business on their own. Of course, to insure that individuals with promising ideas are willing to start their own businesses, it must be that the expected return to doing so ($\phi \cdot \lambda$) is sufficiently large relative to the start-up cost ω .

The key to the model is that the probability of finding a match, p , may be affected by family size. Consider the following components that go into the probability of an owner of a family business finding a match. Let γ_0 , $0 \leq \gamma_0 \leq 1$, be the probability that the business owner is approached by an outsider, who has enough expertise and knowledge to see the true inside value of the firm. Furthermore, let γ_1 , $0 \leq \gamma_1 \leq 1$, be the per-child probability that a business owner's progeny sees the true value of the firm. For example, if a business owner has a good business but no children, his chance of finding a match is γ_0 . If he has one child it is $1 - (1 - \gamma_0) \cdot (1 - \gamma_1)$. With two children, the probability of a match is $1 - (1 - \gamma_0) \cdot (1 - \gamma_1)^2$. More generally, for "n" number of children, the probability of finding a match is:

$$p(n) = 1 - (1 - \gamma_0) \cdot (1 - \gamma_1)^n$$

Note that $p(n)$ is increasing in γ_0 , γ_1 , and n , where $p'(n) = -(1 - \gamma_0) \cdot (1 - \gamma_1)^n \cdot \log(1 - \gamma_1) \geq 0$, where the equation holds with equality only if $\gamma_1 = 0$ or $\gamma_0 = 1$.⁴ Note that if $\gamma_1 = 0$, then the probability of a match is constant and independent of n .

The optimality condition for a turns out to be the standard relationship where the business owner equates the marginal utility from consumption in period 1 with the expected marginal utility from consuming in period 2: namely,

$$U'(c_1) - (1 - p(n)) \cdot U'(c_2^{NM}) - p(n) \cdot U'(c_2^M) = 0 \quad (2)$$

The optimality condition for the number of children, n , is the following:

$$\theta - \alpha \cdot U'(c_1) + p'(n) \cdot \{U(c_2^M) - U(c_2^{NM})\} = 0 \quad (3)$$

The first two terms of equation (3) represent the standard marginal benefit and marginal cost of having children. The final term, reflects the positive impact that having more children has on making a positive business match. Note that if having more children does not help in matching, then $\gamma_1 = 0$ and $p'(n) = 0$, so that the last term vanishes.

The key proposition that we want to show is that if having more children increases your ability to obtain a good match, then self-employed people will have more of them. There are two distinct effects. The first is the direct effect of how γ_1 affects the optimality condition for the optimal choice of the number of children, expression (3). It is straightforward to see that this direct effect will be positive: from the optimality condition for the number of children, expression (3), the marginal benefit for having more children rises as long as $p'(n) > 0$. In other words, if having more children makes it easier to find a successful match, this will raise the desire to have more children. The second effect is to see how the choice of n is indirectly affected by a change in γ_1 that affects the optimal choice of net-savings in the intertemporal consumption smoothing

⁴Note that: $dp(n)/d\gamma_1 = +(1 - \gamma_0) \cdot n \cdot (1 - \gamma_1)^{n-1} > 0$.

decision, expression (2). Namely,

$$\left[\frac{dn}{d\gamma_1} \right] = \left[\frac{\partial n}{\partial \gamma_1} \right] + \left[\frac{\partial n}{\partial a} \right] \cdot \left[\frac{\partial a}{\partial \gamma_1} \right]$$

Linearizing the system of two unknowns, n and a , and two equations, (2) and (3), and differentiating with respect to γ_1 , in the neighborhood of $\gamma_1 = 0$, the total effect is:

$$\begin{aligned} \left[\frac{dn}{d\gamma_1} \right] \Big|_{\gamma_1=0} &= \left\{ - \left[U''(c_1) + pU''(c_2^M) + (1-p)U''(c_2^{NM}) \right] \cdot \left[(\partial p'(n)/\partial \gamma_1) \left\{ U(c_2^M) - U(c_2^{NM}) \right\} \right] \right. \\ &\quad + \left[(\partial p(n)/\partial \gamma_1) \left\{ U'(c_2^M) - U'(c_2^{NM}) \right\} \right] \cdot \left[\alpha U''(c_1) + p'(n) \left\{ U'(c_2^M) - U'(c_2^{NM}) \right\} \right] \left. \right\} / \\ &\quad \left\{ \left[pU''(c_2^M) + (1-p)U''(c_2^{NM}) \right] \cdot \left[\alpha^2 U''(c_1) \right] \right\} > 0 \end{aligned} \quad (4)$$

Critical to signing the effect is to note that $\partial p'(n)/\partial \gamma_1 > 0$ for $1 > \gamma_1 \geq 0$.⁵ It is straightforward to show that the numerator and denominator are both positive, so that $dn/d\gamma_1|_{\gamma_1=0} > 0$.

While the above model provides the important prediction that individuals who are self-employed will want to have more children than they would if they were not self-employed, it did so by assuming non-benevolence to their children. As alluded to in footnote 1, this assumption allows us to present the prediction in the simplest model possible. Indeed, while the optimal number of children would take into account all the standard costs and benefits to the decision maker of having an additional child (i.e. child rearing costs, the non-pecuniary utility benefit from having children, the cost an extra child on watering down the per-child bequest, etc...), the fact remains that as long as self-employed individuals have an additional matching benefit for having children, that is increasing the likelihood of the business's succession, they will choose to have more children. As such, the theory's prediction can be shown to hold for model's with explicit bequests and benevolent parents.⁶

⁵Note that

$$\begin{aligned} p''(n) &= -(1-\gamma_0) \cdot (\log(1-\gamma_1))^2 \cdot (1-\gamma_1)^n < 0, \\ \partial p'(n)/\partial \gamma_1 &= (1-\gamma_0) \cdot (1-\gamma_1)^{n-1} \cdot (1-n \cdot \log(1-\gamma_1)) > 0. \end{aligned}$$

⁶In a related paper, Bernheim, Schleifer and Summers (1985) explore the theoretical and empirical aspects of a strategic bequest motive where parents condition their bequests on the actions of their children's actions. Our strategic motive in this paper, endogenizing the number of children, is complementary to their strategic bequest motive. See Chami (2001) for an extensive theoretical treatment of the issues surrounding strategic bequests and family businesses.

3 Empirical Analysis

The theoretical model presented above provides a strong prediction: namely, that individuals who own businesses will have larger preferred family sizes. However, we are confronted with a number of practical issues when implementing an empirical test of this prediction. First, most data sets do not comprehensively ask questions about whether an individual is part of a family business.⁷ However, a type of family business that is consistently and comprehensively recorded in the annual General Social Survey (GSS) is whether the respondent is self-employed. From the perspective of our model, self-employment captures the main features we are attempting to proxy in a potential family business: namely, an individual who owns a business and who at some point would benefit from selling the business off to either an outsider or an inside family member.⁸ Hence, we use self-employment status throughout the empirical work below to proxy the type of individual who owns a business and whose preferred family size may be influenced by the factors identified in our theory.

Second, there exists the possibility that the empirical work below, which attempts to identify how self-employment status may affect the respondent's preferred number of children, suffers from some type of reverse causality. In other words, perhaps individuals who prefer larger family sizes decide to choose self-employment over working for some other business. While we cannot ignore this possibility, we adopt a few measures to help ensure that this problem does not cloud our findings. More importantly, we provide instrumental variables estimates which should help overcome any endogeneity bias from our estimated effect of self-employment status on the respondent's actual and expected number of children. As well, we provide additional evidence in sub-section 3.3 that

⁷This shortcoming is shared by the standard data sets that one might consider to use for testing our prediction – i.e. the the Panel Study of Income Dynamics (PSID), the National Longitudinal Survey of Youth (NLSY) and the General Social Survey (GSS), as either the question is limited to a specific age group, or the question changes, or is just asked in a few years.

⁸Interestingly, in 1996 the GSS did ask whether the respondent was a member of a family business. The question asked, “Do you work for pay in a farm or other business owned in whole or in substantial part by a member or members of your family (parents, grandparents, children, brothers, sisters, aunts, uncles, or first cousins?” 54 respondents answered yes, while 1381 responded no. Of the 54 who said yes, about half answered that they were self-employed. Note that whether an individual works for a family business is not quite the right question for what we are looking for. For instance, a respondent who works for a family firm and does not own it (or never anticipates owning it) would respond ‘Yes’ to the question of working for a family firm, would likely answer ‘No’ to being self-employed, and nevertheless would not, according to our line of theory, want to have more children. As this demonstrates, responses to the self-employment question are likely to be a better proxy for the incentives of an owner of a family business as compared to the family business question itself.

indicates that respondents who are self-employed work more hours and do not seem to have more “family friendly” jobs as compared to workers who are not self-employed. As such, this would suggest that self-employment status may not be a good haven for workers with lots of children as a way to better fit their work into their family lives.

3.1 The Data

In this section, we begin by describing the data employed in the paper. The GSS is an annual survey which asks respondents specific questions concerning family size and employment status. As well, it asks numerous other questions which we use to control for demographic variation and individual heterogeneity in their preferences for family size. The respondents are not re-interviewed across years so that the data are repeated cross sections rather than a panel. A description of the variables we examine is presented in the Data Appendix.⁹

The key variable of interest for this study is a measure of a respondent’s family size. We adopt a number of measures of this from the data available in the GSS: namely, the respondent’s actual number of children, KIDS, and their actual plus expected additional children, TOTKIDS. The latter may be an important measure of family size for younger respondents who may be planning to add children to their current family size in order to reach their preferred family size.

Of course, the self-employment status of the family is also a key variable of interest in this study. We denote SELFE to be a dummy variable equal to 1 if the respondents report that they are self-employed, and zero otherwise.¹⁰ As discussed above, self-employment is our proxy for a family business, as self-employment provides perhaps the greatest opportunity to pass down a business to a member of the next generation of one’s own family. Hence, if our theory is correct, the motivation for succession and insuring the business’s survival is likely to be key component in the fertility decisions of the self-employed. To control for additional family related factors influencing a respondent’s family size, we allow for a number of variables related to the respondent’s family.

⁹There are a total number of 15 thousand respondents who were not retired or in school (nor were their spouses if they had one) and responded to family size and self-employment questions throughout most years in 1972-2002. For the sample of currently married male respondents who are not in school nor retired (nor are their spouses), summary statistics for the key variables are reported in the Table 1.

¹⁰The exact wording is: ‘ (Are/Were) you self-employed or (do/did) you work for someone else? ’.

For instance, MARRIED is a dummy variable that indicates whether the respondent is currently married. As well, SPOUSEH is a dummy variable that takes the value 1 if the work status of the respondents spouse was ‘taking care of the house’, and 0 otherwise.

Additional explanatory variables which we use to control for observable factors that can systematically influence an agent’s preference for more children are the respondent’s age, AGE, sex, MALE, and a measure of their perceived financial status relative to others, FINRELA.¹¹ We also constructed measures of work and life experience: GENERATION is the year that the respondent was born, and LMEX is the respondent’s labor market experience, namely their age minus their number of years of schooling minus 6. We also use highest educational attainment for the respondent and the spouse as possible controls for the respondent’s preference for children: DIPHR, DIPJCR, DIPBAR, DIPGDR, DIPHSP, DIPJCSP, DIPBASP, and DIPGDSP are dummy variables for whether the highest diploma (DIP) was for high school (H), junior college (JC), four year college (BA) or for graduate school (GD), earned by the respondent (R) and spouse (SP).¹² Additional demographic information is contained in the dummy variables JEWISH and CATHOLIC, for respondent listed religion, WHITE and BLACK, for whether a respondent lists this as their race. City size effects are also controlled for: CTYSZ1 is a dummy variable if the respondent’s residence is less than 10,000, and 0 otherwise. CTYSZ2, CTYSZ3 and CTYSZ4 are similarly defined for population ranges 10,000 to 100,000, 100,000 to 1 million, and greater than 1 million, respectively.

Finally, we also control for the respondent’s reported INDUSTRY using a series of eleven dummy variables: Agriculture, Mining, Construction, Manufacturing, Transportation, Wholesale, Retail, Finance (FIRE), Entertainment, and Public Administration, and Professional Services. We also use the information on the respondent’s family background that they were raised in: SIBS is their number of siblings, while IMGRNT, IMGRNTPA and IMGRNTMA are dummy

¹¹FINRELA is measured on a self reported scale of 1 to 5, with 1 being that your financial status is well below the mean, 2 being that it is below the mean, 3 being that it is about the mean, 4 being that it is above the mean, and 5 being that it is far above the mean. While the GSS does report family income for some years, it does not do so for all years.

¹²Unfortunately, not all demographic questions are asked for both the respondent and his/her spouse. As noted below, this leads us to only use Male respondents for the majority of our empirical work. Note that the main result, that self-employed respondents have more children, holds for the full sample, and a large fraction of the sub-samples, including Males and Females separately.

variables that denote whether the respondent, the respondent’s father and mother were immigrants, respectively. Moreover, ETHNIC ORIGIN denotes the country the respondent’s ancestors came from: Africa, Austria, French Canada, Other Canada, China, Czech Republic, Denmark, England, Wales, Finland, France, Germany, Germany, Greece, Hungary, Ireland, Italy, Japan, Mexico, Dutch Holland, Norway, Philippians, Poland, Puerto Rico, Russia, Scotland, Spain, Sweden, Switzerland, West Indies, and Other.

3.2 Evidence

In the following sub-sections, we examine the main prediction from our theory: namely, that individuals who are self-employed are more likely to have larger families than those who are not self-employed. Sub-section 3.2.1 presents the stylized facts of the data we employ for this study. Sub-section 3.2.2 provides a more formal empirical test of our hypothesis.

3.2.1 Empirical Regularities

Table 1 provides an interesting set of empirical regularities in the data. For purposes described below, the primary data set we examine in this study includes only currently married males who are neither currently enrolled in school nor retired nor are their spouses.¹³ The first column of the data lists variable names, while columns two through four present the variable’s mean, standard deviation and median. The final three columns of the table report the correlation coefficient between the variable in the first column and the three key variables of interest in this study: namely, KIDS, KIDSTOT, and SELFE, respectively. As reported in the table, the respondent’s average number of actual and expected children is 2.2 and 2.4, respectively, while the average number of siblings is relatively larger at 3.8. The average respondent is just over 40 years old, has over 13 years of schooling and has approximately 23 years of labor market experience. The majority of the sample reports their race as WHITE, their religion as PROTESTANT and lives in a city size of under

¹³As we demonstrated below in Tables 2 and 4, the general finding that respondents who report that they are self-employed have significantly more children hold for both men and women. However, while the GSS reports the spouse’s self employment status, it does not provide much of the important demographic information for the respondent’s spouse. Hence we believe that given data availability the male only sample provides a more comprehensive picture of the relationship between self-employment status and family size.

100,000 (CTYSZ1 and CTSZ2 combined).¹⁴

The correlation coefficients reported in Table 1 also reveal a number of interesting findings between the variables of interest. First, as expected, the actual and expected number of children are highly correlated. Consequently, other variables have similar correlation patterns with these two variables. Second, age and labor market experience are positively and significantly correlated with the actual and expected number of children. Moreover, these correlations are large. Also of interest, the respondent's education is negatively correlated with the number of children, while his number of siblings and whether his wife works in the home is positively correlated with his number of children. Also, respondents that are BLACK, or from smaller cities, or that are CATHOLIC, tend to have more children.

Finally, approximately 18 percent of the respondents list themselves as 'self-employed'.¹⁵ Interestingly, self-employment status and the number of children is positively and significantly correlated. Also, self-employment status appears to be higher for older respondents with more labor market experience, as well as for those who report themselves as JEWISH. Finally, consistent with Fairlie and Meyer (1996,2000), self-employment is significantly higher for WHITE's and lower for BLACK's.¹⁶ Moreover, as indicated by their significant negative correlations with GENERATION, both self-employment and the number of children have decreased over time. As such, this will be an important variable to control for in our regression results below in sub-Section 3.2.2.^{17,18}

Table 2 presents the difference of means by self-employment status. The first column of the

¹⁴Friedlander and Silver (1967) argue that less dense populations are likely to have lower living costs, and thus would be more likely to have higher fertility rates.

¹⁵Again, this is for married men who are not retired, nor in school, nor are their spouses. The mean self-employment rate for women is 8.8 percent for married women who satisfy this same criteria.

¹⁶Fairlie and Meyer (2000) test several hypotheses in order to explain the difference in black and white self-employment rates. These hypotheses center on whether demographic changes have led to a decrease in the black/white gap and whether the black/white gap is due primarily to the historically low self-employment rates and past inexperience of blacks. They find that demographic changes that occurred in the twentieth century did not have a large effect on the black/white self-employment gap and using a simple intergenerational model of self-employment they find that it is not only initial conditions that explain low black self-employment but also continuing forces that depress black self-employment (i.e. discrimination or skills, capital, and intangibles that are passed intergenerationally).

¹⁷To note, the regression results not only include GENERATION as a right hand side variable, but they also include dummy variables for the year in which the interview took place.

¹⁸The four largest industries represented among the self employed are (in order) Finance, Construction, Professional Services, and Agriculture, respectively. The latter, which accounts for about 14 percent of the self employed in our sample but are only 4 percent of sample, is excluded from our baseline empirical work in columns (VII) and (VIII) of Table 4. The results are unaffected by their exclusion.

Table 1: Sample Statistics

VARIABLE	MEAN	STD	MEDIAN	CORRELATIONS		
				KIDS	KIDSTOT	SELFE
KIDACT	2.215	1.593	2			
KIDSTOT	2.411	1.546	2	.922***		
SELFE	.182	.386	0	.069***	.058***	
FINRELA	3.113	.827	3	-.064***	-.079***	.027*
LMEX	23.2	12.8	21	.421***	.294***	.132***
AGE	42.6	12.1	41	.392***	.261***	.150***
GENERATION	1944.9	13.5	1947	-.390***	-.302***	-.134***
SPWRKH	.341	.474	0	.164***	.162***	.002
SIBS	3.812	3.194	3	.170***	.186***	-.034**
DIPHR	.850	.357	1	-.214***	-.200***	-.003
DIPJCR	.054	.226	0	-.043***	-.023	-.055***
DIPBAR	.177	.382	0	-.103***	-.103***	.029*
DIPGDR	.103	.304	0	-.040***	-.047***	.052***
DIPHSP	.590	.492	1	.010	.008	.000
DIPJCSP	.061	.240	0	-.037***	-.036**	-.005
DIPBASP	.157	.364	0	-.101***	-.083***	.031**
DIPGDSP	.057	.231	0	-.076***	-.084***	.007
BLACK	.077	.267	0	.081***	.080***	-.074***
WHITE	.885	.319	1	-.076***	-.074***	.068***
JEWISH	.020	.140	0	-.024	-.032**	.109***
CATHOLIC	.275	.447	0	.056***	.078***	-.052***
PROTESTANT	.586	.493	1	.014	.003	.020
IMGRNTR	.080	.272	0	-.013	-.005	-.022
IMGRNTPA	.160	.367	0	.035**	.027*	.019
IMGRNTMA	.144	.351	0	.028*	.026*	.019
CTYSZ1	.353	.478	0	.032**	.031**	.068***
CTYSZ2	.434	.496	0	-.011	-.014	-.050***
CTYSZ3	.164	.371	0	-.021	-.020	-.025
CTYSZ4	.048	.214	0	-.010	-.003	.006

Note: The data sample size is 4220 and includes only currently married male respondents when neither spouse is retired or in school, MALES† & MAR. See Data Section 3.1 and the Data Appendix for data descriptions. Column 1 lists the variables. Columns 2 through 4 report the means, standard deviations and medians. Columns 5 – 7 report the correlations between the variables listed in the rows with those listed at the top of the columns. ***, **, and * indicate that the relationship is different from zero at or below the .01, .05 and .10 level of statistical significance, respectively.

table lists the samples of interest, while column two provides the dependent variables, namely KIDS and KIDSTOT. Columns three through five report the average number of actual and expected kids for the full sample, the proportion of the sample that is self-employed, and for the proportion of the sample that is not self-employed, respectively. Column six presents the p-value of the test that the means are equal across the two sub-samples (i.e. self-employed and not self employed), while column seven displays the number of observations for the sample of interest.¹⁹ The presentation of the means and the test of the equality of means across the sub-samples is performed for both measures of the respondent's number of children – KIDS and KIDSTOT.

A key and fundamental observation from the Table 2 bears directly on the model's main prediction: that is, for the full sample and for 81 out of the 94 sub-samples, the average number of children, both as measured by KIDS and KIDSTOT, is larger for the self-employed sub-sample versus the non-self-employed sub-sample. In other words, in 81 out of 94 cases the means in column (V) are larger than those for column (VI). Moreover, 56 out of the 94 tests of the equality of the means that are reported in Table 2 are statistically significant at or below the .1 level, and in all these cases the mean number of children of the self-employed is greater than that of the non-self-employed. And often in these cases where the null hypothesis of no difference between the two means fails to be rejected are typically for sub-samples with a small relatively number of observations. For example, the sample of males living in large cities (CTYSZ3 and CTYSZ4) have sample sizes of 693 and 203 and fail to show any significant difference between the average number of children. There are, however, a number of instances where the p-value is below .1 despite the small sample size. For example, the sample size for Jewish men is only 84 but for both measures of the dependent variable, the difference in the means is significant at the .1 level. Interestingly, there does seem to be some regional variation in differences of the means. These regional variables will be in the regression results below and are often statistically significant.

Table 2 provides some other interesting findings. First, regardless of their marital status, men who are self-employed have, on average, a higher number of actual and expected kids as compared to men that are not self-employed – the difference in means is significant at below the .01 level. Second, regardless of immigration status and the immigration status of the respondents

¹⁹The underlying t-test is obtained using a procedure that is robust to heteroskedasticity of unknown form.

Table 2: Tests for Differences in Means by Self-Employment Status

Sample (I)	Variable (II)	MEANS			p-value (VI)	NOBS (VII)
		FULL (III)	SELFE (IV)	NON-SELFE (V)		
ALL	<i>KIDS</i>	1.83	2.05	1.80	.000	15010
	<i>KIDSTOT</i>	2.09	2.24	2.07	.000	15010
ALL MALES	<i>KIDS</i>	1.70	2.00	1.64	.000	6859
	<i>KIDSTOT</i>	2.01	2.23	1.96	.000	6859
ALL FEMALES	<i>KIDS</i>	1.94	2.12	1.92	.003	8151
	<i>KIDSTOT</i>	2.16	2.25	2.15	.120	8151
MALES [†]						
& BMAR	<i>KIDS</i>	2.18	2.36	2.14	.000	5267
	<i>KIDSTOT</i>	2.36	2.50	2.33	.003	5267
& MAR1	<i>KIDS</i>	2.13	2.34	2.08	.000	3372
	<i>KIDSTOT</i>	2.34	2.49	2.31	.007	3372
& MAR	<i>KIDS</i>	2.22	2.45	2.16	.000	4220
	<i>KIDSTOT</i>	2.41	2.60	2.37	.000	4220
MALES [†] & MAR						
& IMGR	<i>KIDS</i>	2.14	2.56	2.07	.067	339
	<i>KIDSTOT</i>	2.38	2.79	2.31	.086	339
& IMGRPAR	<i>KIDS</i>	2.35	2.64	2.28	.020	782
	<i>KIDSTOT</i>	2.52	2.78	2.46	.036	782
& CATH	<i>KIDS</i>	2.36	2.58	2.32	.069	1161
	<i>KIDSTOT</i>	2.61	2.83	2.57	.053	1161
& JEWISH	<i>KIDS</i>	1.95	2.20	1.73	.082	84
	<i>KIDSTOT</i>	2.06	2.28	1.86	.104	84
& PROTESTANT	<i>KIDS</i>	2.23	2.46	2.18	.001	2471
	<i>KIDSTOT</i>	2.42	2.59	2.37	.009	2471
& BLACK	<i>KIDS</i>	2.66	3.07	2.62	.271	325
	<i>KIDSTOT</i>	2.84	3.41	2.79	.190	325
& WHITE	<i>KIDS</i>	2.17	2.40	2.12	.000	3735
	<i>KIDSTOT</i>	2.37	2.55	2.33	.001	3735
& RACEOTH	<i>KIDS</i>	2.33	3.15	2.17	.003	160
	<i>KIDSTOT</i>	2.51	3.15	2.38	.017	160
& SPWRKH	<i>KIDS</i>	2.58	2.82	2.52	.011	1440
	<i>KIDSTOT</i>	2.76	2.92	2.72	.090	1440
& NON-SPWKRH	<i>KIDS</i>	2.03	2.25	1.98	.000	2780
	<i>KIDSTOT</i>	2.23	2.43	2.19	.001	2780
& CTYSZ1	<i>KIDS</i>	2.29	2.43	2.25	.082	1491
	<i>KIDSTOT</i>	2.48	2.59	2.45	.179	1491
& CTYSZ2	<i>KIDS</i>	2.19	2.54	2.13	.000	1833
	<i>KIDSTOT</i>	2.39	2.67	2.33	.001	1833
& CTYSZ3	<i>KIDS</i>	2.14	2.29	2.11	.232	693
	<i>KIDSTOT</i>	2.34	2.41	2.33	.549	693
& CTYSZ4	<i>KIDS</i>	2.15	2.36	2.10	.366	203
	<i>KIDSTOT</i>	2.39	2.69	2.32	.177	203

Note: Continued.

Continued

Table 2: Tests for Differences in Means by Self-Employment Status

Sample (I)	Variable (II)	MEANS			p-value (VI)	NOBS (VII)
		FULL (III)	SELF (IV)	NON-SELF (V)		
MALES [†] & MAR						
& RELINC12	<i>KIDS</i>	2.42	2.74	2.34	.009	825
	<i>KIDSTOT</i>	2.64	2.85	2.60	.102	825
& RELINC3	<i>KIDS</i>	2.20	2.45	2.15	.003	2082
	<i>KIDSTOT</i>	2.41	2.64	2.36	.005	2082
& RELINC45	<i>KIDS</i>	2.12	2.28	2.07	.041	1313
	<i>KIDSTOT</i>	2.27	2.42	2.24	.068	1313
& NEW ENGLAND	<i>KIDS</i>	2.10	2.65	2.00	.021	252
	<i>KIDSTOT</i>	2.30	2.89	2.20	.004	252
& MID-ATLANTIC	<i>KIDS</i>	2.10	2.24	2.07	.295	667
	<i>KIDSTOT</i>	2.28	2.41	2.25	.308	667
& CENTRAL NE	<i>KIDS</i>	2.29	2.54	2.24	.052	912
	<i>KIDSTOT</i>	2.50	2.67	2.47	.185	912
& CENTRAL NW	<i>KIDS</i>	2.37	2.78	2.24	.011	388
	<i>KIDSTOT</i>	2.53	2.93	2.41	.010	388
& SOUTH ATL.	<i>KIDS</i>	2.05	2.48	1.96	.005	581
	<i>KIDSTOT</i>	2.26	2.55	2.20	.064	581
& CENTRAL SE	<i>KIDS</i>	2.22	2.18	2.22	.873	245
	<i>KIDSTOT</i>	2.32	2.27	2.34	.782	245
& CENTRAL SW	<i>KIDS</i>	2.34	2.47	2.31	.412	346
	<i>KIDSTOT</i>	2.56	2.56	2.56	.975	346
& PACIFIC	<i>KIDS</i>	2.12	2.08	2.13	.789	356
	<i>KIDSTOT</i>	2.31	2.29	2.32	.862	356
& NO DIPLOMA	<i>KIDS</i>	3.03	3.27	2.97	.153	634
	<i>KIDSTOT</i>	3.15	3.38	3.09	.182	634
& HIGH SCHOOL	<i>KIDS</i>	2.17	2.33	2.14	.026	2175
	<i>KIDSTOT</i>	2.37	2.46	2.35	.201	2175
& JUNIOR COLLEGE	<i>KIDS</i>	1.93	2.05	1.92	.694	227
	<i>KIDSTOT</i>	2.26	2.57	2.23	.223	227
& BA	<i>KIDS</i>	1.86	2.12	1.79	.018	749
	<i>KIDSTOT</i>	2.07	2.34	2.00	.010	749
& GRAD	<i>KIDS</i>	2.03	2.50	1.88	.000	435
	<i>KIDSTOT</i>	2.20	2.59	2.07	.001	435

Note: Continued.

Continued

Table 2: Tests for Differences in Means by Self-Employment Status

Sample (I)	Variable (II)	MEANS			p-value (VI)	NOBS (VII)
		FULL (III)	SELF (IV)	NON-SELF (V)		
MALES [†] & MAR						
& AGRICULTURAL	<i>KIDS</i>	2.77	2.85	2.62	0.455	166
	<i>KIDSTOT</i>	2.96	2.96	2.95	0.966	166
& MINING	<i>KIDS</i>	2.43	2.29	2.44	0.878	61
	<i>KIDSTOT</i>	2.61	2.29	2.65	0.725	61
& CONSTRUCTION	<i>KIDS</i>	2.39	2.54	2.32	0.217	452
	<i>KIDSTOT</i>	2.59	2.73	2.52	0.261	452
& TRANSPORTATION	<i>KIDS</i>	2.23	1.96	2.25	0.190	396
	<i>KIDSTOT</i>	2.41	2.07	2.43	0.135	396
& WHOLESALE	<i>KIDS</i>	2.21	2.21	2.21	0.999	205
	<i>KIDSTOT</i>	2.40	2.38	2.41	0.896	205
& RETAIL	<i>KIDS</i>	2.10	2.39	2.00	0.040	353
	<i>KIDSTOT</i>	2.33	2.44	2.29	0.419	353
& FINANCE	<i>KIDS</i>	2.18	2.37	2.08	0.054	552
	<i>KIDSTOT</i>	2.39	2.54	2.32	0.142	552
& PUBLIC ADMIN.	<i>KIDS</i>	2.11	2.67	2.10	0.319	375
	<i>KIDSTOT</i>	2.28	3.00	2.27	0.033	375
& MANUFACTURING	<i>KIDS</i>	2.26	2.38	2.25	0.493	1132
	<i>KIDSTOT</i>	2.45	2.68	2.43	0.183	1132
& ENTERTAINMENT	<i>KIDS</i>	1.78	1.57	1.84	0.645	32
	<i>KIDSTOT</i>	2.03	1.71	2.12	0.504	32
& PROFESSIONAL	<i>KIDS</i>	2.07	2.42	1.98	0.004	572
	<i>KIDSTOT</i>	2.26	2.53	2.20	0.024	572

Note: See Table 1 and Data Section 3.1. Column 1 reports the sample. Column 2 lists the measures of the number of Children. Columns 3 – 5 report the mean number of Children for the full sub-sample and then partitioned for the self-employed and non-self-employed portions of these sub-samples. Column 6 reports the p-value from the test (robust to heteroskedasticity of unknown form) of the null hypothesis that the means from the two partitioned sub-samples are identical. NOBS reports the number of observations. MALES[†] are MALE respondents who are neither retired nor in school, nor are their spouses.

father, those who are self-employed have a statistically significant higher average number of kids. Similarly, regardless of the respondents spouses work status (SPOUSEH), and the familys relative income (FINRELA), men who are self- employed on average tend to have a significantly larger number of actual and expected kids. Finally, there is also some heterogeneity of differences in the actual and expected number of children even within industry classifications. Interestingly, within each of the Retail, Financial and Professional Services industries, self-employed respondents have more actual and expected children. These findings are of interest, as many of these industries are likely to exemplify the type where a self-employed business could benefit from a child who sees the private value to the business (e.g. a retail shop, an insurance business, a dentist practice, etc...)²⁰ Taken together, we believe that the results in Table 2 provide extremely strong preliminary evidence in favor of the prediction: namely, *ceteris paribus*, self-employed respondents have more children.

3.2.2 Estimation

To test the hypothesis that the self-employed have more kids, we regress measures of family size (KIDS and KIDSTOT) on a dummy variable for whether the respondent is self-employed (SELFE) and a number of control variables to proxy for the respondents preferences for their family size.²¹ The control variables can be thought of as capturing household characteristics and demographics. Examples of the former are whether or not the spouse stays at home (SPOUSEH) and the respondents perceived financial relative position (FINRELA). Examples of the latter are features of residential area (CTYSZ), the respondents age (AGE); and trends that may affect the number of children a couple has over time: year of birth. The model is specified as follows:

²⁰ That Financial (Finance, Real Estate and Insurance) and Professional Services industries demonstrate this feature is particularly important for overcoming the following criticism: perhaps the cost of child rearing is cheaper for the self-employed as they have a source of cheap, unpaid help. While this criticism could apply to some retail stores (e.g. restaurants), it certainly would not apply to the more skilled professions inherent in Finance, Insurance and Real Estate or Professional Services.

²¹ Again, as in Table 1, the regression results will be only for married working males, though the results in Table 2 demonstrate that this fact is broader than for just this sub-sample of the data. The focus on men is primarily due to the data incompleteness for the respondent's spouse. Since married women are more likely to be part-time employed as compared to men, their self-employment status is likely to be less important in the fertility decision. Unfortunately, since we only have the parent's self- employment status for the respondent but not for the respondent's spouse, we are missing crucial information for the instrumental variables estimation when we look at female respondents.

$$N_i = \beta_0 + \beta_1 \cdot SELFE_i + \beta_2 \cdot X_i + \beta_3 \cdot T_i + \epsilon_i \quad (5)$$

where N_i represents the number of kids respondent i has or expects to have, X_i represent household characteristics and demographics of respondent i , and T_i represent time dummy variables for the year in which the interview was conducted. More specifically, the following variables are used as controls: SPOUSEH, FINRELA, JEWISH, CATHOLIC, AGE, AGE², GENERATION, LMEX, SIBS, BLACK, WHITE, 8 variables separately denoting the respondent and spouses highest degree obtained (DIPHR, DIPJCR, DIPBAR, DIPGDR, DIPHSP, DIPJCSP, DIPBASP, and DIPGDSP), three dummy variables for whether the respondent, father or mother were immigrants to the U.S. (IMGRNT, IMGRNTPA and IMGRNTMA) 26 ethnic origin dummy variables, dummy variables for city size (CITYSZ1-4), and the 8 Census regions.²² The data are described fully in the Data Appendix. Also, Table 1 and 2, as summarized previously, provide some raw data on the variables of interest and the control variables.

We estimate a number of empirical regressions of KIDS and KIDSTOT on the explanatory variables. To control for the possibility of endogeneity among some of the explanatory variables in particular, SELFE and SPOUSEH we also estimated the specifications using instrumental variables, IV.²³ While our theory predicts that the self-employed will tend to have more kids, we cannot ignore the possibility that families with more kids will see the benefits of a family business and choose to become self-employed. Also, since women who stay home instead of work in the market have more kids, a wife with more kids is more likely to stay at home – see the survey in Blau (1998). To control for the possibility that SELFE and SPOUSEH may be endogenous, we use a number of instruments, Z , that explain SELFE and SPOUSEH but are exogenous to KIDS and KIDSTOT. In particular, the instruments used are whether or not the respondents father was self-employed, the industry the respondent works in, stability of the respondents household at age 16, the religious affiliation of the respondent at age 16, and the financial relative status of the respondent at age

²²The specifications also contains a constant, so that where appropriate a dummy variable is excluded so as to avoid collinearity.

²³More specifically, the instrumental variables approach we adopt is the two-step efficient Generalized Method of Moments estimator.

16.^{24,25} As demonstrated in the p-values below, together these variables are statistically significant predictors of SELFE and SPOUSEH, and the instruments are not significantly correlated with the error term. As the number of instruments exceeds the number of estimated coefficients, below we will test the over-identifying restrictions that the residuals and the instruments are orthogonal – see the p-value of the J-test below.

Table 3 provides estimation results of the specified key variables for the OLS and IV regressions for KIDS and KIDSTOT, the explanatory variables of interest are presented in column one and the results for the IV regressions are reported in the odd numbered columns. The table provides the coefficient estimates, their levels of significance, and their standard errors. At the bottom of Table 3 are reported p-values for several tests to be explained later.

The key result from Table 3 is that, after controlling for a whole host of demographic variables such as religion, age, labor market experience, race, immigration status and other variables assumed to influence the number of kids an individual may choose to have, the estimated coefficient on SELFE is positive and statistically significant at below the .10 level in all regressions. The coefficient estimates from the IV regressions show that a male who is self-employed has approximately 0.4 more kids than his non-self-employed counterpart, and expects to have about 0.5 more kids than

²⁴For example, Hout and Rosen (2002) demonstrate that the offspring of self-employed fathers are more likely themselves to be self-employed, primarily for the reason we identify: namely, to hand down the business. Indeed, in our data, the probability of a respondent being self-employed given that his father was self-employed is equal to .279 (350/1255). In contrast, the probability falls to .141 (417/2965) if the respondent’s father is not self-employed. The difference is statistically significant at below the .001 level. However, unlike Rees and Shah (1986), we do not have earnings differentials to explain self-employment, but rather use industry dummy variables as a proxy. Moreover, Blanchflower and Oswald (1998) find evidence that a respondent’s receipt of a gift or inheritance increases the likelihood of self-employment. Their explanation, consistent with their theory, is that such a windfall reduces capital constraints on potential entrepreneurs. While such data is not available in the GSS, the likelihood of a gift or inheritance is likely to be correlated with one of our instruments: namely, the respondent’s financial status at age 16.

Interestingly, the ‘succession’ motive for why self-employed parents have more children may be connected to this ‘borrowing constraint’ literature for the following reason: The within-family succession of a self-employed business is less likely to fail because of borrowing constraints as compared to the sale of the business to an individual outside the family. The reason is that families can use informal and non-market arrangements to compensate the parents for the business. Moreover, the existence of borrowing constraints would make it even less likely for a business owner to sell the business at a fair market price to an outsider, in addition to the information and monitoring reasons we have already discussed. Hence, the fact that potential buyers outside the family will be more likely to be borrowing constrained as compared to inside family members will magnify the desire of the self-employed to have more children in order to raise the chances of obtaining a good internal match for the business’s succession.

²⁵In general, the instruments are good predictors of the potentially endogeneous regressors. In Tables 3 and 4, the p-value for SELFE and SPOUSEH reports the significance level of the F-test that, in a first stage regression, the coefficients in the instrumental variables Z are not all jointly equal to zero when regressed against these variables. In all cases, the p-value for SELFE is below .001 (with an associated F-statistic of over 30, while the p-value for SPOUSEH ranges from .01 to .1, with an associated F-statistic of about 3. As the former is the basis of inquiry for this study, we believe that our list of instruments is a good one.

Table 3: Baseline Regression of Number of Children on Self-Employment Status

Dep. Variable	KIDS		KIDSTOT	
	OLS	IV	OLS	IV
SELFE	.139*** (.059)	.407** (.183)	.161*** (.061)	.486*** (.185)
SPWRKH	.420*** (.049)	.553 (.627)	.353*** (.050)	.030 (.643)
FINRELA	-.066** (.029)	-.062** (.030)	-.074** (.029)	-.079** (.031)
JEWISH	-.084 (.175)	-.170 (.179)	-.097 (.177)	-.188 (.180)
CATHOLIC	.159*** (.058)	.170*** (.059)	.229*** (.059)	.252*** (.060)
AGE	.162*** (.023)	.162*** (.034)	.075*** (.024)	.057* (.035)
AGE ² /100	-.184*** (.013)	-.189*** (.022)	-.091*** (.014)	-.082*** (.022)
GENERATION	-.015* (.008)	-.014 (.012)	-.013 (.008)	-.018 (.012)
LMEX	.039** (.019)	.043** (.019)	.026 (.020)	.030 (.019)
SIBS	.040*** (.009)	.040*** (.009)	.049*** (.011)	.048*** (.011)
BLACK	.765*** (.298)	.748** (.298)	.708** (.294)	.717** (.290)
WHITE	-.042 (.169)	-.056 (.168)	-.012 (.163)	-.007 (.160)
IMGRNTR	.082 (.139)	.074 (.138)	.164 (.140)	.175 (.137)
IMGRNTPA	-.198** (.100)	-.218** (.106)	-.230** (.100)	-.224** (.107)
IMGRNTMA	-.073 (.109)	-.057 (.109)	-.085 (.111)	-.083 (.110)
p-values				
YEARS	.623	.594	.000	.000
City Size	.021	.026	.169	.191
Diplomas	.000	.039	.002	.022
Region	.000	.000	.000	.000
Ethnic Origin	.010	.007	.002	.013
Hausman		.262		.062
J-test		.628		.568
SELFE		.001		.001
SPOUSEH		.046		.046

Notes: See Tables 1 and 2. Estimated standard errors, robust to heteroskedasticity of unknown form, in parentheses. IV refers to two-step efficient GMM. The data sample is for MALES[†] & MAR, with 4220 observations in each regression. P-values for YEARS, City Size, Diplomas, Region and Ethnic Origin are from F-tests from excluding the linearly independent time dummies, CITYSZ1-3, the eight respondent and spouse diploma variables, seven Census regions and 34 Ethnic origin dummy variables, respectively. Hausman is the p-value for the specification test that the coefficients on SELFE and SPWRKH are the same across the OLS and IV estimates. J-test is the p-value for the Hansen test that the residuals are uncorrelated with the instruments, Z . SELFE and SPOUSEH are p-values for the test that, in a first stage regression, the coefficients on the instrumental variables Z are not all jointly equal to zero when regressed against these variables.

a male of equal standing but who is non-self-employed. For the OLS regressions, the coefficient estimates on SELFE are not as large in magnitude, though they are still statistically significant in all four columns. These results indicate that a self employed male has .139 more kids and expects to have .161 more kids than a male who is not self-employed.

An important result that is in accordance with previous research on the optimal number of kids is that the coefficient on FINRELA is negative and significant at the .05 level. The negative relationship between income and fertility decisions found in previous studies is assumed to explain a households decision to spread risk across children – see Appelbaum and Katz (1991) as well as Mulligan (1997). Another important result studied by many researchers is the role the mothers involvement in the labor market has to do with a households decision to have kids. Previous research has found that there is a negative relationship between a wifes labor market participation and the number of kids she gives birth to – see Willis (1973), Robinson and Tomes (1982), and Rosenzweig and Schultz (1985).²⁶ Whether or not the spouse stays at home is positive and highly significant in both OLS regressions but loses its significance when instrumented for in the IV regressions – see footnote 31.

Some other interesting results observed in Table 3 are that the CATHOLIC variable is positive and significant at the .01 level in each one of the regressions. Individuals who are Catholic tend to have approximately .16 more kids and tend to expect approximately .23 more kids. The JEWISH variable is negative in all regressions, though it is not statistically significant. Older men tend to have more actual and expected children, while the coefficient on the GENERATION variable, which is used to capture the change in demographics and changes in family structure over time, is negative for each of the regressions but is only significant for 1 of the 4 regressions. LMEX which is used to measure the husbands stability in the workforce which in turn provides a measure of the stability of the household is positive in all four of the regressions but only significant in both the OLS and the IV regressions for KIDS, a male with an additional year of labor market experience appears to have approximately .04 more kids.

Individual characteristics, for the most part, are as expected. The more siblings you have the

²⁶These authors argue that raising children is labor intensive which is why it accounts for the negative relationship with the mother's labor market participation, particularly in the short run.

more likely you are to have and expect more kids. Blacks tend to have more kids, and immigration status and mothers immigration status has no significant effect on the number of kids you have, while fathers immigration status has a negative and significant effect on the number of kids you have and expect to have.²⁷

The p-values for the F-test of whether variables in sub-categories are all equal to zero are reported at the bottom of Table 3. The diploma variables, the regional variables and the ethnic origin variables all reject the null hypothesis that the coefficients are all zero for each of the regressions, implying that education, location, and ethnicity influence the number of kids you have and expect to have.²⁸ On the other hand, the null hypothesis for the year variables is rejected for expected number of kids but not for actual number of kids. The null hypothesis that all the coefficients on the city size variables are zero is rejected for the actual number of kids but failed to be rejected for the expected number of kids.

Also reported at the bottom of Table 3 is the p-value for the Hausman test, which tests whether there is any significant difference between the OLS regression and the IV regression for the estimated coefficients on SELFE and SPOUSEH. Under the null hypothesis that there is no endogeneity bias to self-employment status and whether the respondents spouse stays at home, the estimated coefficients on these two variables should be the same across the OLS and IV estimation methods, with the exception that the IV estimate would be less efficient. If, however, the estimates using OLS and IV are different, then this would indicate evidence against the null hypothesis of no endogeneity bias. The results using the actual number of kids as the dependent variable suggests that there is no significant difference between the IV and OLS estimates of these key variables, which suggests that we do not have any endogeneity problems. However, when the expected number of kids is the dependent variable, there is a significant difference between the OLS and the IV regression at the .10 level, which suggests that the IV regression will provide a more accurate

²⁷Previous research such as Borjas (1986), and Fairlie and Meyer (1996), has found that several immigrant groups have a statistically higher self-employment rate than the native born. While we do not show a statistical positive correlation between immigration status and self-employment it could be due to the low number of observations or the fact that we did not distinguish directly between different immigrant groups, though we do control for a respondents ethnic origin.

²⁸In particular, the city size and Census Region variables – New England, Mid- Atlantic, Central NE, Central NW, South Atlantic, Central SE, Central SW, and Pacific – proxy for the fact that environmental factors may explain differences in birth rates – see Shultz (1969).

estimate of the effect of self-employment status on family size. The final three rows of Table 3 provide further evidence that the IV regressions are meaningful. The row denoted J-test is the orthogonality test between the estimated residuals and the instruments. The p-value indicates that there is no evidence to suspect that the instruments are correlated with the error term. Finally, the p-values denoted SELFE and SPOUSEH are the p-value from an F-test of the null hypothesis that in a linear regression of each of these variables on the instruments, that all the coefficients on the instruments would be jointly equal to zero. As such, this type of test indicates whether the instruments are correlated with the potentially endogenous regressors. As the p-values indicate, however, the instruments are statistically significant predictors of SELFE and SPOUSEH.

While the results in Table 3 provide consistently strong findings in support of the theory's prediction, the results in Table 4 provide additional results to show the robustness of our findings. For example, in columns (I) through (X) of Table 4, we present both OLS and IV estimates of the coefficients for sub-samples where respondents report that KIDS is positive, the respondent's spouse stays at home ($SPOUSEH = 1$), the respondent is currently in his first marriage ($MAR1=1$), for respondents that are not in the agricultural industry ($AG=0$), and for respondents that are not immigrants nor are their parents ($IMGRNT=0$).²⁹ To keep the number of results to a minimum, the dependent variable in each of these specifications is KIDSTOT.³⁰

Estimating the relationship between self-employment and family size over these sub-samples is helpful for a number of reasons. First, individuals may not have a particularly precise idea of their preferred family size (KIDSTOT) until they actually have some children ($KIDS > 0$). Note, however, that the results in columns (I) and (II) are very similar to those for the full sample. Second, since a prime determinant of whether a spouse stays at home is whether they have children (and or whether they have a larger number of children), the potential endogeneity of the variable *SPOUSEH* may be worrisome. Indeed, as shown in Table 3, while the coefficient on SPOUSEH is

²⁹Interestingly, Friedlander and Silver (1967) find that as a country's fraction of self-employed non-farm population rises, that their fertility rate rises. Though they do not provide an explanation for this finding other than to conjecture that there are lower child rearing costs for the self-employed, this fact is consistent both with our theory and empirical work on individual household behavior. See footnote 20 for our argument why this is not likely to be an explanation for our finding.

³⁰To note, the results are similar when KIDS is used as the dependent variable.

Table 4: Additional Regression on Total Number of Children (KIDSTOT)

SAMPLE ESTIMATOR	<i>KIDS</i> > 0		<i>SPOUSEH</i> = 1		<i>MAR1</i> = 1	
	<i>OLS</i> (<i>I</i>)	<i>IV</i> (<i>II</i>)	<i>OLS</i> (<i>III</i>)	<i>IV</i> (<i>IV</i>)	<i>OLS</i> (<i>V</i>)	<i>IV</i> (<i>VI</i>)
SELFE	.176*** (.059)	.605*** (.182)	.161 (.113)	1.034*** (.372)	.103* (.061)	.576*** (.197)
SPWRKH	.270*** (.048)	-.075 (.608)			.366*** (.053)	.773 (.545)
FINRELA	-.084*** (.028)	-.091*** (.029)	-.083 (.054)	-.098 (.062)	-.050 (.031)	-.050 (.032)
JEWISH	-.186 (.164)	-.309* (.167)	.013 (.285)	-.172 (.272)	-.027 (.174)	-.186 (.182)
CATHOLIC	.189*** (.057)	.215*** (.058)	.285*** (.107)	.327*** (.120)	.258*** (.061)	.250*** (.063)
AGE	.056*** (.023)	.047 (.035)	.096** (.045)	.085* (.045)	.054** (.026)	.064** (.033)
AGE ² /100	-.074*** (.013)	-.064*** (.023)	-.123*** (.021)	-.122*** (.021)	-.076*** (.014)	-.088*** (.022)
GENERATION	-.018** (.008)	-.022** (.011)	-.008 (.016)	-.013 (.016)	-.010** (.009)	-.005 (.011)
LMEX	.020 (.019)	.013 (.020)	.035 (.037)	.035 (.037)	.036* (.021)	.039* (.021)
SIBS	.048*** (.011)	.055*** (.012)	.065*** (.019)	.065*** (.019)	.052*** (.012)	.051*** (.012)
BLACK	.572** (.292)	2.629** (1.138)	.824* (.456)	.978** (.496)	.995** (.307)	1.007*** (.311)
WHITE	-.085 (.160)	-.173 (.173)	.133 (.295)	.213 (.328)	.097 (.153)	.124 (.154)
IMGRNTR	.052 (.131)	.068 (.128)	.110 (.260)	.112 (.250)	.088 (.153)	.094 (.158)
IMGRNTPA	-.154 (.097)	-.147 (.105)	-.214 (.159)	-.208 (.154)	-.234** (.104)	-.253** (.110)
IMGRNTMA	-.024 (.105)	-.005 (.106)	-.193 (.190)	-.207 (.183)	-.060 (.120)	-.065 (.123)
NOBS	3780	3780	1436	1436	3360	3360
p-values						
YEARS	.000	.000	.121	.126	.000	.000
City Size	.601	.178	.209	.425	.397	.575
Diplomas	.001	.009	.080	.047	.005	.179
Region	.000	.000	.008	.008	.000	.000
Ethnic Origin	.000	.129	.000	.030	.001	.001
Hausman		.015		.014		.037
J-test		.562		.270		.845
SELFE		.126		.000		.000
SPOUSEH		.100				.014

CONTINUED

Table 4: Additional Regression on Total Number of Children (KIDSTOT)

SAMPLE ESTIMATOR	<i>AG</i> = 0		<i>IMGRNT</i> = 0		<i>ALL</i>	<i>ALL</i>	<i>ALL</i>
	<i>OLS</i> (<i>VII</i>)	<i>IV</i> (<i>VIII</i>)	<i>OLS</i> (<i>IX</i>)	<i>IV</i> (<i>X</i>)	<i>TOBIT</i> (<i>XI</i>)	<i>IVTOBIT</i> (<i>XII</i>)	<i>OPRBT</i> (<i>XIII</i>)
SLFEMPR	.144*** (.065)	.494* (.263)	.116* (.066)	.645*** (.195)	.155** (.063)	.499*** (.190)	.120*** (.045)
SPWRKH	.343*** (.051)	.087 (.717)	.408*** (.057)	.280 (.713)	.400*** (.052)	.131 (.683)	.272*** (.037)
FINRELA	-.084*** (.030)	-.090*** (.031)	-.075** (.032)	-.076** (.033)	-.077*** (.031)	-.082** (.033)	-.048** (.022)
JEWISH	-.112 (.179)	-.211 (.190)	-.088 (.229)	-.207 (.236)	-.068 (.210)	-.162 (.217)	-.090 (.144)
CATHOLIC	.221*** (.060)	.243*** (.062)	.269*** (.065)	.285*** (.065)	.250*** (.063)	.263*** (.065)	.182*** (.044)
AGE	.082*** (.024)	.066* (.036)	.069** (.027)	.054 (.039)	.091*** (.024)	.078** (.037)	.066*** (.017)
AGESQ/100	-.095*** (.014)	-.088*** (.024)	-.094*** (.015)	-.091*** (.024)	-.106*** (.014)	-.100*** (.024)	-.072*** (.010)
GENERATION	-.014 (.008)	-.019 (.013)	-.012 (.009)	-.013 (.013)	-.013 (.010)	-.016 (.014)	-.010 (.006)
LMEX	.021 (.020)	.026 (.020)	.035 (.023)	.046** (.023)	.027 (.018)	.031 (.019)	.013 (.013)
SIBS	.048*** (.011)	.047*** (.011)	.053*** (.012)	.053*** (.013)	.051*** (.008)	.052*** (.008)	.033*** (.006)
RACEBL	.742** (.301)	.750** (.298)	.816** (.400)	1.290 (1.469)	.786*** (.273)	.831*** (.276)	.445** (.196)
RACEWH	.014 (.166)	.015 (.162)	.247 (.274)	.736 (.634)	.007 (.173)	.038 (.175)	.029 (.116)
IMGRNTR	.144 (.143)	.153 (.141)			.204 (.167)	.225 (.169)	.145 (.103)
IMGRNTPA	-.191* (.101)	-.187* (.110)			-.272*** (.104)	-.251*** (.112)	-.173*** (.074)
IMGRNTMA	-.111 (.110)	-.103 (.111)			-.095 (.112)	-.117 (.114)	-.073 (.080)
NOBS	4036	4036	3407	3407	4220	4220	4220
p-values							
YEARS	.000	.000	.000	.000	.000	.000	.000
City Size	.228	.236	.462	.439	.111	.117	.235
Diplomas	.002	.028	.017	.183	.004	.024	.008
Region	.000	.000	.000	.000	.000	.000	.000
Ethnic Origin	.001	.008	.001	.168	.007	.009	.000
Hausman		.141		.005		.078	
J-test		.236		.645		.820	
SELFE		.000		.000		.001	
SPOUSEH		.038		.098		.046	

Note: See following page.

CONTINUED

Table 4: Additional Regression on Total Number of Children (KIDSTOT)

Note: See Tables 1, 2 and 3. In columns denoted OLS, IV, TOBIT, IVTOBIT and OPRBT, the coefficients were estimated using ordinary least squares, instrumental variables, tobit, instrumental variables tobit, and ordered probit, respectively. As in Tables 1 and 3, the data set only includes only married males who are not retired, nor are their spouses (MALES† & MAR). In columns (I) and (II), the data set removes respondents who have no current children. In columns (III) and (IV), the data set includes only those respondents whose spouses stay at home. In columns (V) and (VI), the data set includes only those respondents who have been married once. In columns (VII) and (VIII), the data set removes respondents in the agricultural industry. In columns (IX) and (X), the data set removes respondents who are immigrants or if either of their parents were immigrants.

statistically significant in the OLS regressions, it is not so in the IV ones.³¹ Interestingly, as shown in columns (III) and (IV) of Table 4, the OLS estimate of the coefficient on SELFE is insignificantly different from zero, although the IV estimate is significantly different from zero at the .01 level. Moreover, the IV estimate on the coefficient on SELFE is quite large – i.e. it predicts that for respondents with stay at home spouses, self-employed have over 1.0 additional children than do the non-self-employed. Third, we report estimates for the sub-sample of male respondents who are currently married and have never been married before (MAR1).³² This sub-sample is important since the relationship between respondents and children from prior marriages may cloud the empirical relationships that we are attempting to investigate. However, as demonstrated in columns (V) and (VI), the estimated specification is very close to that for the full sample – see Table 3 columns (III) and (IV). Furthermore, as emphasized by Fairlie and Meyer (1996, 2000), we re-explore our empirical findings by removing respondents who are in the agricultural industry. As shown in columns (VII) and (VIII) of Table 4, however, the empirical results are virtually identical when these respondents are removed from the sample.³³ As a final attempt to demonstrate the

³¹ Of course, this is hampered by the fact that number of children is often an instrument for whether a spouse stays at home. For example, see the survey in Blau (1998).

³²The data is similar to that in Table 3 – that is, neither the respondent nor his spouse are retired or in school.

³³While the level of statistical significance is no longer at below the .05 level, it is significant at the .06 level.

robustness of our findings across sub-samples, columns (IX) and (X) report the model's estimates when respondents who are immigrants or whose parents are immigrants are removed from the sample. Such a control is of interest since respondents whose parents are immigrants are more likely to be self-employed: see Table 1. However, as shown in columns (IX) and (X), the results are unchanged from our baseline set of results when we remove respondents who are immigrants or whose parents were immigrants from the sample.³⁴

Table 4 also allows for a few changes to the econometric approach to estimating our basic specification. In particular, since the number of children is bounded below by zero, one may wonder whether our estimates are impacted by directly incorporating this into our estimation procedure. To this end, column (XI) and (XI) report Tobit and Instrumental Variables Tobit estimates of the key parameters. As the results suggest, however, the using this more sophisticated approach provides estimates and levels of significance almost identical to that using OLS and IV as presented in Table 3. Finally, in column (XIII) of Table 4, we re-estimate our specification using an Ordered Probit model. Such an estimator would directly incorporate the fact that the number of children discrete. Once again, however, the estimated pattern of signs and statistical significance are similar to our baseline estimates in Table 3.

3.3 Remarks

Taken together, the results in Tables 3 and 4 confirm the broad evidence presented in Table 2: self-employed households tend to have more children. We have demonstrated this result to be robust over key sub-samples of the data as well as to alternative methods for estimating the parameters. As such, we believe we have demonstrated that there is strong evidence in favor of our theory which is predicated on the idea that self-employed households increase their number of children in order to improve their chances of finding a match that improves the value of the business.

Nevertheless, doubts about the link between our theory and the evidence may persist. And

Moreover, the estimated coefficient value is very similar to that in our baseline specification in Table 3, although the larger estimated standard error is consistent with the smaller sample size.

³⁴Though not shown, the results are also unchanged when the sample is just whites who were not immigrants nor were there parents.

perhaps reasonably so. For example, one may conjecture that non-pecuniary benefits of having children may be higher for the self-employed for at least two reasons. First, the possibility exists that individuals may choose to be self-employed because the self-employed spend less time at work, which then allows them to spend more time with their children. Unfortunately, while we cannot fully test this hypothesis with the GSS data, there is some evidence in the GSS data that suggests just the opposite: namely, that the self-employed work more hours than fewer hours. For example, the GSS asked the following question: “How many hours did you work last week”. Interestingly, the simple correlations between KIDS and KIDSTOT with HOURS are $-.013$ and $-.021$, respectively, neither of which is statistically different from zero. However, the correlation between SELFE and HOURS is $.092$, which is statistically different from zero at below the $.0001$ level.³⁵ Put in another light, the average hours of worked for the self-employed is over 49 hours per week, while that for the non self-employed is around 45. Hence, more time spent at work suggests that self-employment does not free up extra hours with which to spend with one’s children, the latter which would generally be expected to be associated with respondents who prefer larger families. Finally, this mean gap of about 4 hours per week remains statistically significant even if we include all the control variables used in our control for all the explanatory variables in our estimating equation (5).³⁶

However, while self-employed respondents may work about 10 percent more hours than non self-employed respondents, perhaps there is a sense in which these jobs are more family friendly. Of course, while self-employment does provide some job flexibility (i.e. as your own boss you don’t have to ‘punch a clock’), the additional hours of work are likely to make other labor-leisure trade-offs more difficult to manage.³⁷ The GSS actually provides a number of questions, though only for a limited time period, about the extent to which work may impinge upon family obligations. For example, in 1996, the GSS asked respondents whether work had ever made them ‘miss a family occasion or holiday,’ ‘been unable to care for a sick child or relative’, or ‘been unable to do the work you usually do around the house’. Also, in 1998 they asked how important were ‘the person’s family

³⁵Again the sample is for MALE respondents, who are currently married, and neither they nor their spouses are retired or in school. Since fewer responded to the hours question, the number of observations is about 3800.

³⁶In other words, when we replace N_i with *HOURS* in regression (5), we find an effect equal to 3.8 hours with a robust t-statistic of 6.8. Moreover, this estimate and level of significance is unchanged even if we also include the instrument variables Z , which include industry dummy variables for the respondent.

³⁷Hamilton (2000) reports that the self-employed tend to earn less and work more than individuals with similar employment situations.

responsibilities?’ Importantly, the responses to these questions were all insignificantly correlated with the respondent’s self-employment status at or below the .1 level. Hence, self-employment (at least for the male sub-sample we have considered) is not a bargain with respect to increased flexibility towards family commitments. As such, this lessens the weight behind the criticism that households choose self employment in response to a larger actual or expected family size.³⁸

4 Conclusion

Self-employed entities face unique challenges that separate them from other publicly-owned enterprises. Aside from the profit-maximizing objective, self-employed individuals are typically concerned with ensuring that the business stays within the family. Reconciling the two objectives, namely, the success of the business and the control of the family over the business, implies that the parent must try to find a match among his children that would also guarantee the success of the business.

We provide a theory whereby self-employed households have an inducement to have more children in order to raise the expected return to their business. The important mechanism which generates this is that having more children can increase the likelihood that an inside family member will be a good match at running the business. Using data from the General Social Survey, we find empirical support for this finding. That is, overall, we demonstrate that there is a strong empirical relationship between a respondent’s family size and self-employment status. We find that, *ceteris paribus*, the self-employed have between .2 to .4 more actual and total (that is, actual plus expected) children than do the non-self-employed. This finding holds across a broad array of sub-samples in simple tests of the differences in means. It also holds in empirical regressions when control variables are included, and when the self-employment status of the respondent and whether the respondent’s spouse stays at home are all allowed to be endogenous.

³⁸In a recent important contribution, Frey and Benz (2003) provide evidence from Germany, the U.K. and Switzerland that self-employed are happier with their work not due to better financial outcomes, but rather because of non-pecuniary benefits such as their work’s independence and lower level of hierarchy. This work extends Blanchflower and Oswald’s (1998) and Blanchflower’s (2000) finding that the self-employed are happier than similarly employed workers.

References

- Altig, David, and Davis, Steven J., 1992. "The Timing of Intergenerational Transfers, Tax Policy, and Aggregate Savings." *The American Economic Review*, 82(5), 1199-1220.
- Appelbaum, Elie, and Eliakim, Katz, 1991. "The Demand for Children in the Absence of Capital and Risk Markets: A Portfolio Approach." *Oxford Economic Papers*, 43(2), 292-304.
- Arnott, Richard, and Stiglitz, Joseph E., 1991. "Moral Hazard and Nonmarket Institutions: Dysfunctional Crowding Out of Peer Monitoring?" *The American Economic Review*, 81(1), 179-190.
- Barro, Robert J., and Becker, Gary S., 1989. "Fertility Choice in a Model of Economic Growth." *Econometrica* 57(2), 481-501.
- Bernheim, B. Douglas, Shleifer, Andrei and Summers, Lawrence H., 1985. "The Strategic Bequest Motive." *The Journal of Political Economy* 93(6), 1045-1076.
- Bernheim, B. Douglas and Kotlikoff, Laurence J., 2001. "Household Financial Planning and Financial Literacy: The Need for New Tools," in Laurence J. Kotlikoff (eds.), *Essays on Saving, Bequests, Altruism, and Life-Cycle Planning*, MIT Press, 2001.
- Blanchflower, David G., 2000. "Self-Employment in OECD Countries." *Labor Economics* 7, 471-505.
- Blanchflower, David G. and Oswald, Andrew J., 1998. "What Makes an Entrepreneur?" *Journal of Labor Economics* 16(1), 26-60.
- Blau, Francine, 1998. "The Well-being of American Women, 1970-1995," *Journal of Economic Literature*, 106(1), 112-165.
- Borjas, George J., 1986. "The Self-Employment Experience of Immigrants." *The Journal of Human Resources*, 21(4): 485-506.
- Chami, Ralph, 2001. "What's Different About Family Business?" *IMF Working Paper*, # 01/70.
- Chami, Ralph, and Fischer, Jeffrey H., 1996. "Altruism, Matching, and Nonmarket Insurance." *Economic Inquiry*, 34(4), pp. 630-647.
- Chami, Ralph and Hess, Gregory D. 2002. "For Better or For Worse: State Level Marital Formation and Risk Sharing," *CEifo Working Paper* # 702.
- Clark, F., 1966. "The Commodore Left Two Sons," *American Heritage*, 17, 81-103.
- Cox, Donald 1987. "Motives for Private Income Transfers." *The Journal of Political Economy*, 95(3), 508-546.
- Dasgupta, Partha 1995. "The Population Problem: Theory and Evidence." *The Journal of Economic Literature*, 33(4), 1879-1902.

- Fairlie, Robert W., and Meyer, Bruce D., 1996. "Ethnic and Racial Self-Employment Differences and Possible Explanations." *The Journal of Human Resources*, 31(4): 757-793.
- Fairlie, Robert W., and Meyer, Bruce D., 2000. "Trends in Self-Employment among White and Black Men during the Twentieth Century." *The Journal of Human Resources*, 35(4), 643-669.
- Friedlander, Stanley, and Silver, Morris, 1967. "A Quantitative Study of the Determinants of Fertility Behavior." *Demography*, 4(1), 30-70.
- Frey, Bruno S. and Benz, Matthias, 2003. "Being Independent is a Great Thing: Subjective Evaluations of Self-Employment and Hierarchy." *CESifo Working Paper # 959*.
- Gersik, Kelin, Davis, John, Hampton, Marion McCollom and Lansberg, Ivan, (1997). "Generation to Generation: Life Cycles of the Family Business," Boston: *Harvard Business School Press*.
- Hamilton, Barton, 2000. "Does Entrepreneurship Pay? An Empirical Analysis of Returns to Self-Employment?" *The Journal of Political Economy*, 108(3), 604-632.
- Hess, Gregory D., 2004. "Marriage and Consumption Insurance: What's Love Got to do with It?" *The Journal of Political Economy*, forthcoming.
- Hout, Michael, and Rosen, Harvey, 2000. "Self-Employment, Family Background, and Race." *The Journal of Human Resources*, 35(4), 670-692.
- Leung, Siu Fai, 1991. "A Stochastic Dynamic Analysis of Parental Sex Preferences and Fertility." *The Quarterly Journal of Economics*, 106(4), 1063-1088.
- Mulligan, Casey B., 1997. *Parental Priorities and Economic Inequality*, Chicago: University of Chicago Press.
- Rees, Hedley, and Shah, Anup. 1986. "An Empirical Analysis of Self-Employment in the U.K." *Journal of Applied Econometrics*, 1(1), 95-108.
- Robinson, Chris, and Tomes, Nigel, 1982. "Family Labour Supply and Fertility: A Two-Regime Model." *The Canadian Journal of Economics* 15(4), 706-734.
- Rosenzweig, Mark R., and Schultz, T. Paul, 1985. "The Demand for and Supply of Births: Fertility and its Life Cycle Consequences." *The American Economic Review*, 75(5), 992-1015.
- Schultz, Theodore W., 1973. "The Value of Children: An Economic Perspective." *The Journal of Political Economy* Vol 81(2), S2-S13.
- Willis, Robert J., 1973. "A New Approach to the Economic Theory of Fertility Behavior." *The Journal of Political Economy*, 81(2), S14-S64.
- Zhang, Jie, and Zhang, Junsen, 1997. "Fertility and Wage Rates in an Overlapping-Generations Model." *The Canadian Journal of Economics*, 30(1), 224-234.

Data Appendix: Variables Used in the Analysis

Key Variables of Interest

KIDS The number of children respondent has ever had (includes those born from a previous marriage).

KIDSTOT The number of children respondent has plus the additional number they expect to have.

SELFE Dummy variable 1 if self-employed 0 if works for someone else.

Additional Control Variables

AGE Age of Respondent.

BLACK Dummy variable 1 if race is black, and 0 otherwise.

BMAR Dummy variable 1 if respondent has ever been married, and 0 otherwise.

CATHOLIC Dummy variable 1 if religion preference is Catholic, and 0 otherwise.

CTYSZ1 Dummy variable 1 if population at which interview took place is less than 10,000, and 0 otherwise.

CTYSZ2 Dummy variable 1 if population at which interview took place is greater than or equal to 10,000 but less than 100,000, and 0 otherwise.

CTYSZ3 Dummy variable 1 if population at which interview took place is greater than or equal to 100,000 but less than 1 million, and 0 otherwise.

CTYSZ4 Dummy variable 1 if population at which interview took place is greater than or equal to 1 million but less than 9 million 0 otherwise.

DIPHR(S) Dummy variable 1 if respondent's (spouse) highest degree received was from a high-school or GED certificate, and 0 if not.

DIPJCR(S) Dummy variable 1 if respondent's (spouse) highest degree received was from a junior college, and 0 if not.

DIPBAR(S) Dummy variable 1 if respondent's (spouse) highest degree received was from a 4 year college, and 0 if not.

DIPHGDR(S) Dummy variable 1 if respondent's (spouse) highest degree received was from a Graduate School, and 0 if not.

Ethnic Origin The Country the respondent's ancestors came from: Africa, Austria, French Canada, Other Canada, China, Czech Republic, Denmark, England, Wales, Finland, France, Germany, Germany, Greece, Hungary, Ireland, Italy, Japan, Mexico, Dutch Holland, Norway, Philippians, Poland, Puerto Rico, Russia, Scotland, Spain, Sweden, Switzerland, West Indies, Other.

FINRELA The respondent's household financial income relative to others. It is measured on a self reported scale of 1 to 5, with 1 being that your financial status is well below the mean, and 5 being that it is far above the mean.

IMGRNT Dummy variable 1 if respondent was not born in the U.S., and 0 otherwise.

IMGRNTMA Dummy variable 1 if respondents mother was not born in the U.S., and 0 otherwise.

IMGRNTPA Dummy variable 1 if respondents father was not born in the U.S., and 0 otherwise.

Industry Agriculture, Mining, Construction, Manufacturing, Transportation, Wholesale, Retail, Finance (Finance, Insurance and Real Estate), Entertainment, PROFESSIONAL (Professional Services), Administration.

JEWISH Dummy variable 1 if religion preference is Jewish , and 0 otherwise.

LMEX Respondent's labor market experience. Equal to Age minus years of school minus 6.

MALES Dummy variable 1 if male, and 0 if female.

MALES† Dummy variable 1 if MALES and if neither respondent nor spouse are in school or retired, and 0 otherwise.

MAR Dummy variable 1 if currently married, and 0 otherwise.

MAR1 Dummy variable 1 if respondent is currently married or widowed and never been previously married, and 0 otherwise.

PROTESTANT Dummy variable 1 if religion preference is Protestant, and 0 otherwise.

SIBS Number of siblings respondent has (includes those no longer alive, stepbrothers, stepsisters and children adopted by parents).

YRSCHR Number of years of schooling.

WHITE Dummy variable 1 if race is white, and 0 otherwise.

Regions New England, Mid-Atlantic, Central NE (North East), Central SE (South East), South Atlantic, Central SE, Central SW, and Pacific.

SPWRKH Dummy variable 1 if respondents spouse work status a week prior to the interview is either taking care of the house or student, and 0 if $spwrk = 1$.

CESifo Working Paper Series

(for full list see www.cesifo.de)

- 1037 Roberta Dessi and Sheilagh Ogilvie, Social Capital and Collusion: The Case of Merchant Guilds, September 2003
- 1038 Alessandra Casarico and Carlo Devillanova, Capital-skill Complementarity and the Redistributive Effects of Social Security Reform, September 2003
- 1039 Assaf Razin and Efraim Sadka, Privatizing Social Security Under Balanced-Budget Constraints: A Political-Economy Approach, September 2003
- 1040 Michele Moretto, Paolo M. Panteghini, and Carlo Scarpa, Investment Size and Firm's Value under Profit Sharing Regulation, September 2003
- 1041 A. Lans Bovenberg and Peter Birch Sørensen, Improving the Equity-Efficiency Trade-off: Mandatory Savings Accounts for Social Insurance, September 2003
- 1042 Bas van Aarle, Harry Garretsen, and Florence Huart, Transatlantic Monetary and Fiscal Policy Interaction, September 2003
- 1043 Jerome L. Stein, Stochastic Optimal Control Modeling of Debt Crises, September 2003
- 1044 Thomas Stratmann, Tainted Money? Contribution Limits and the Effectiveness of Campaign Spending, September 2003
- 1045 Marianna Grimaldi and Paul De Grauwe, Bubbling and Crashing Exchange Rates, September 2003
- 1046 Assar Lindbeck and Dennis J. Snower, The Firm as a Pool of Factor Complementarities, September 2003
- 1047 Volker Grossmann, Firm Size and Diversification: Asymmetric Multiproduct Firms under Cournot Competition, September 2003
- 1048 Dan Anderberg, Insiders, Outsiders, and the Underground Economy, October 2003
- 1049 Jose Apesteguia, Steffen Huck and Jörg Oechssler, Imitation – Theory and Experimental Evidence, October 2003
- 1050 G. Abio, G. Mahieu and C. Patxot, On the Optimality of PAYG Pension Systems in an Endogenous Fertility Setting, October 2003
- 1051 Carlos Fonseca Marinheiro, Output Smoothing in EMU and OECD: Can We Forego Government Contribution? A Risk Sharing Approach, October 2003

- 1052 Olivier Bargain and Nicolas Moreau, Is the Collective Model of Labor Supply Useful for Tax Policy Analysis? A Simulation Exercise, October 2003
- 1053 Michael Artis, Is there a European Business Cycle?, October 2003
- 1054 Martin R. West and Ludger Wößmann, Which School Systems Sort Weaker Students into Smaller Classes? International Evidence, October 2003
- 1055 Annette Alstadsaeter, Income Tax, Consumption Value of Education, and the Choice of Educational Type, October 2003
- 1056 Ansgar Belke and Ralph Setzer, Exchange Rate Volatility and Employment Growth: Empirical Evidence from the CEE Economies, October 2003
- 1057 Carsten Hefeker, Structural Reforms and the Enlargement of Monetary Union, October 2003
- 1058 Henning Bohn and Charles Stuart, Voting and Nonlinear Taxes in a Stylized Representative Democracy, October 2003
- 1059 Philippe Choné, David le Blanc and Isabelle Robert-Bobée, Female Labor Supply and Child Care in France, October 2003
- 1060 V. Anton Muscatelli, Patrizio Tirelli and Carmine Trecroci, Fiscal and Monetary Policy Interactions: Empirical Evidence and Optimal Policy Using a Structural New Keynesian Model, October 2003
- 1061 Helmuth Cremer and Pierre Pestieau, Wealth Transfer Taxation: A Survey, October 2003
- 1062 Henning Bohn, Will Social Security and Medicare Remain Viable as the U.S. Population is Aging? An Update, October 2003
- 1063 James M. Malcomson, Health Service Gatekeepers, October 2003
- 1064 Jakob von Weizsäcker, The Hayek Pension: An efficient minimum pension to complement the welfare state, October 2003
- 1065 Joerg Baten, Creating Firms for a New Century: Determinants of Firm Creation around 1900, October 2003
- 1066 Christian Keuschnigg, Public Policy and Venture Capital Backed Innovation, October 2003
- 1067 Thomas von Ungern-Sternberg, State Intervention on the Market for Natural Damage Insurance in Europe, October 2003
- 1068 Mark V. Pauly, Time, Risk, Precommitment, and Adverse Selection in Competitive Insurance Markets, October 2003

- 1069 Wolfgang Ochel, Decentralising Wage Bargaining in Germany – A Way to Increase Employment?, November 2003
- 1070 Jay Pil Choi, Patent Pools and Cross-Licensing in the Shadow of Patent Litigation, November 2003
- 1071 Martin Peitz and Patrick Waelbroeck, Piracy of Digital Products: A Critical Review of the Economics Literature, November 2003
- 1072 George Economides, Jim Malley, Apostolis Philippopoulos, and Ulrich Woitek, Electoral Uncertainty, Fiscal Policies & Growth: Theory and Evidence from Germany, the UK and the US, November 2003
- 1073 Robert S. Chirinko and Julie Ann Elston, Finance, Control, and Profitability: The Influence of German Banks, November 2003
- 1074 Wolfgang Eggert and Martin Kolmar, The Taxation of Financial Capital under Asymmetric Information and the Tax-Competition Paradox, November 2003
- 1075 Amihai Glazer, Vesa Kannianen, and Panu Poutvaara, Income Taxes, Property Values, and Migration, November 2003
- 1076 Jonas Agell, Why are Small Firms Different? Managers' Views, November 2003
- 1077 Rafael Lalive, Social Interactions in Unemployment, November 2003
- 1078 Jean Pisani-Ferry, The Surprising French Employment Performance: What Lessons?, November 2003
- 1079 Josef Falkinger, Attention, Economies, November 2003
- 1080 Andreas Haufler and Michael Pflüger, Market Structure and the Taxation of International Trade, November 2003
- 1081 Jonas Agell and Helge Bennmærker, Endogenous Wage Rigidity, November 2003
- 1082 Fwu-Ranq Chang, On the Elasticities of Harvesting Rules, November 2003
- 1083 Lars P. Feld and Gebhard Kirchgässner, The Role of Direct Democracy in the European Union, November 2003
- 1084 Helge Berger, Jakob de Haan and Robert Inklaar, Restructuring the ECB, November 2003
- 1085 Lorenzo Forni and Raffaella Giordano, Employment in the Public Sector, November 2003
- 1086 Ann-Sofie Kolm and Birthe Larsen, Wages, Unemployment, and the Underground Economy, November 2003

- 1087 Lars P. Feld, Gebhard Kirchgässner, and Christoph A. Schaltegger, Decentralized Taxation and the Size of Government: Evidence from Swiss State and Local Governments, November 2003
- 1088 Arno Riedl and Frans van Winden, Input Versus Output Taxation in an Experimental International Economy, November 2003
- 1089 Nikolas Müller-Plantenberg, Japan's Imbalance of Payments, November 2003
- 1090 Jan K. Brueckner, Transport Subsidies, System Choice, and Urban Sprawl, November 2003
- 1091 Herwig Immervoll and Cathal O'Donoghue, Employment Transitions in 13 European Countries. Levels, Distributions and Determining Factors of Net Replacement Rates, November 2003
- 1092 Nabil I. Al-Najjar, Luca Anderlini & Leonardo Felli, Undescribable Events, November 2003
- 1093 Jakob de Haan, Helge Berger and David-Jan Jansen, The End of the Stability and Growth Pact?, December 2003
- 1094 Christian Keuschnigg and Soren Bo Nielsen, Taxes and Venture Capital Support, December 2003
- 1095 Josse Delfgaauw and Robert Dur, From Public Monopsony to Competitive Market. More Efficiency but Higher Prices, December 2003
- 1096 Clemens Fuest and Thomas Hemmelgarn, Corporate Tax Policy, Foreign Firm Ownership and Thin Capitalization, December 2003
- 1097 Laszlo Goerke, Tax Progressivity and Tax Evasion, December 2003
- 1098 Luis H. B. Braido, Insurance and Incentives in Sharecropping, December 2003
- 1099 Josse Delfgaauw and Robert Dur, Signaling and Screening of Workers' Motivation, December 2003
- 1100 Ilko Naaborg,, Bert Scholtens, Jakob de Haan, Hanneke Bol and Ralph de Haas, How Important are Foreign Banks in the Financial Development of European Transition Countries?, December 2003
- 1101 Lawrence M. Kahn, Sports League Expansion and Economic Efficiency: Monopoly Can Enhance Consumer Welfare, December 2003
- 1102 Laszlo Goerke and Wolfgang Eggert, Fiscal Policy, Economic Integration and Unemployment, December 2003
- 1103 Nzinga Broussard, Ralph Chami and Gregory D. Hess, (Why) Do Self-Employed Parents Have More Children?, December 2003