DO STOCKHOLDERS SHARE RISK MORE EFFECTIVELY THAN NONSTOCKHOLDERS?

Fatih Guvenen*

Abstract—This paper analyzes the extent of risk-sharing among stockholders and nonstockholders. To evaluate the empirical importance of market incompleteness, it is essential to determine whether idiosyncratic shocks are important for the wealthy who have access to better insurance opportunities, but also face different risks, than the average household. We study a model where each period households decide whether to participate in the stock market by paying a fixed cost. Due to this endogenous entry decision, the testable implications of perfect risk-sharing take the form of a sample selection model, which we estimate using a semiparametric GMM estimator proposed by Kyriazidou (2001). Using data from PSID, we strongly reject perfect risk-sharing among stockholders, but perhaps surprisingly, do not find evidence against it among nonstockholders. This result appears to be robust to several extensions. This finding suggests further focus on risk factors that primarily affect the wealthy, such as entrepreneurial income risk.

I. Introduction

In the past several years, models with incomplete markets and uninsurable idiosyncratic shocks have achieved a central place in many fields of economics. These models are now used to study a wide range of economic questions, such as business cycle dynamics, fiscal policy, wealth inequality, and asset prices, among many others (cf. Aiyagari, 1994; Constantinides & Duffie, 1996; Krusell & Smith, 1998; Heaton & Lucas, 2000; Storesletten, Telmer, & Yaron, 2001).

A major motivation for these studies has been the decisive empirical rejection of perfect risk-sharing—the hypothesis that individuals are able to insure against all idiosyncratic shocks, and consequently, that their marginal utilities move in lockstep with each other. A number of empirical studies have found individuals' marginal utility growth (sometimes proxied by consumption growth) to be correlated with certain idiosyncratic shocks, violating the premise of perfect insurance (Cochrane, 1991; Nelson, 1994; Townsend, 1994; and Attanasio & Davis, 1996).

An important point to note is that these studies test whether perfect risk-sharing (PRS) holds among *all* households in the population. However, given that asset holdings and wealth are extremely concentrated—basically 90% of nonhousing wealth and 98% of stocks is held by the top 20% of the U.S. population—wealthy households play a crucial role in many economic interactions.[1] Thus, for a satisfactory analysis of the issues mentioned above, it is especially important to determine the extent of risk-sharing among these *wealthy households.*

On the one hand, there are good reasons to suspect that wealthy households stand a better chance of achieving perfect risk-sharing than the rest of the population. These households almost exclusively trade in stock markets, and therefore have access to the arguably most sophisticated market-based risk-sharing mechanism. Thus, it seems possible that the empirical rejection of PRS among all households could be driven by the lack of insurance among the poor and may not provide a justification that idiosyncratic risk is important for the wealthy. But, on the other hand, wealthy households are exposed to certain risks to a much larger extent than the rest of the population. For example, private capital—which is roughly as large as the capital in publicly traded companies and is potentially difficult to insure due to asymmetric information problems—is concentrated among the wealthy, exposing them to entrepreneurial income risk not faced by other households.[2] Furthermore, investors in financial markets face several trading frictions (such as transactions costs, margin requirements, and costs of information acquisition) that could prevent optimal diversification and expose stockholders to the idiosyncratic risk of the stocks in their portfolio. The first goal of this paper is then to formally investigate whether the (wealthy) stockholders—who face these various risks and have access to various risk-sharing mechanisms—are able to share risk effectively.

As emphasized in the literature, however, trading in financial assets is not the only channel for risk-sharing. There are a number of informal (or nonmarket) insurance mechanisms available to most households, such as redistributive government and social programs, gifts and loans between family members, financial assistance provided by charities, long-term contracts and labor hoarding by companies, and so on. Given also that the less wealthy households (nonstockholders) may not be exposed to certain types of risks as noted above, it is conceivable that they could also be sharing risk effectively among themselves. To investigate this possibility then, we also test for perfect risk-sharing among the less wealthy (nonstockholders). Applying the test

Received for publication September 5, 2003. Revision accepted for publication March 8, 2006.

* Department of Economics, University of Texas.

For helpful comments and suggestions, I would like to thank Daron Acemoglu (the editor), Chris Flinn, John Heaton, Boyan Jovanovic, Shakeeb Khan, Katerina Kyriazidou, Mico Mrkaic, Werner Ploberger, Tony Smith, Harold Zhang, two anonymous referees, and the seminar participants at NYU, University of Rochester, the 2002 SED Conference, and the North American summer meetings of the Econometric Society. Financial support from the William Larimer Mellon foundation is acknowledged. The previous version of this paper was circulated under the title "Does Stockholding Provide Perfect Risk-Sharing?" The usual disclaimer applies.

[1] See Guvenen (2006) for empirical evidence on the inequality of wealth and asset holdings in the United States. Because households in the top 20% own almost all stocks outstanding as well as 90% of the wealth, we also refer to stockholders as "the wealthy," and to nonstockholders as "the poor."

[2] Gentry and Hubbard (2000), and Heaton and Lucas (2000) document the extreme concentration of private capital and the business risks faced by entrepreneurs.

The Review of Economics and Statistics, May 2007, 89(2): 275–288

of PRS to this latter group also has the advantage of providing a benchmark to compare the results for stockholders.

We consider an economy with a complete set of financial assets traded in a stock market. Thus all stocks are *potentially* insurable. Each period, households decide whether to participate in the stock market by paying a per-period participation cost, or to stay out and trade a single risk-free asset. Households also make optimal portfolio and labor supply decisions. In addition, since perfect risk-sharing imposes restrictions on marginal utilities, a sufficiently flexible parameterization of the utility function is crucial for our purposes. To this end, we allow for nonseparabilities between consumption and the leisure times of head and spouse, and incorporate household-specific preference shifters. This specification is more general than most of those adopted in previous studies (with the exception of Altug & Miller, 1990 and Hayashi, Altonji, & Kotlikoff, 1996, which are similar to ours).

Due to the endogenous nature of stock market participation in this model, the testable implications of the risk-sharing hypothesis for stockholders take the form of a sample selection model, where the participation decision serves as the selection equation. To eliminate the selection bias, we implement a semiparametric GMM estimator recently proposed by Kyriazidou (1997, 2001) for panel data models, which does not require strong distributional assumptions about the error terms. To our knowledge, this is the first implementation of this estimator.

Our findings can be summarized as follows. Using data from the Panel Study of Income Dynamics (PSID), we strongly reject perfect risk-sharing among stockholders, but perhaps surprisingly, do not find similar evidence against it among nonstockholders. This result is robust to several extensions we considered, such as including future wages into the instrument set (as advocated by Hayashi et al., 1996), and using different moment conditions implied by PRS, among others. To interpret this finding, it would seem hard to argue that nonstockholders have access to better risk-sharing opportunities than stockholders, suggesting that the difference (in the tests of PRS) is likely to be due to the *additional uninsurable risks* faced by stockholders. Finally, we also strongly reject PRS for the whole population consistent with the existing literature, suggesting that those earlier rejections could also be driven by the lack of risk-sharing among wealthy households.

These results have important implications for modeling choices made in heterogeneous-agent models regarding the types of uninsurable risks. For example, the most common sources of uncertainty in these models are labor income risk, unemployment risk, health shocks, and mortality risk. However, it is also common to assume that different groups in the population—such as the wealthy and the poor—are exposed to these uninsurable risks to similar extents, which is hard to reconcile with our findings.[3] In contrast, some recent studies have begun incorporating entrepreneurial income risk—a risk that primarily affects the wealthy—into incomplete markets models (cf. Angeletos, 2005; Cagetti & De Nardi, 2003; Chari, Golosov, & Tsyvinski, 2005). Our results support this emphasis and suggest further focus on risks that fall disproportionately on the wealthy.

In terms of approach, this paper is most closely related to the literature that test for PRS among smaller groups in the population. The discouraging rejection of PRS in the whole population by the studies mentioned above led researchers to focus on smaller economic units that have strong ties, with the hope of uncovering full insurance within these groups. Examples include households living in the same geographical regions (Hess & Shin, 2000), inhabitants of small villages in various underdeveloped countries (Townsend, 1994; Ogaki & Zhang, 2001), and finally, family members (Hayashi et al., 1996). However, these studies treat participation in these groups as exogenous, and hence do not address selection bias.

The paper is organized as follows. In the next section we set up the model and derive the testable implications of PRS. In section III we develop the econometric techniques to analyze this problem in the presence of selection bias. Then in section IV we describe the data and discuss the estimation of the selection equation. Section V presents the results and discusses robustness. We then discuss some potential explanations for the findings, and conclude in section VI.

II. The Model

There are a finite number of households, each with a lifespan of T periods. Households derive utility from consumption as well as from head's and spouse's leisure times. Specifically, the lifetime utility of household i is given by

$$E_0[\sum_{t=1}^{T} \beta^t u(C_{it}, L_{1it}, L_{2it}\mathbf{X}_{it})], \tag{1}$$

where β is the subjective time discount rate; C_{it} is consumption in period t; L_{1it} and L_{2it} are the leisure times of head and spouse, respectively, in period t; and $\mathbf{X}_{it}$ is a vector of household-specific preference shifters that captures all the heterogeneity across households relevant for preferences.[4] The period utility function, u, is continuously differentiable and concave in the choice variables for each value of $\mathbf{X}_{it}$.

[3] See, for example, Hubbard, Skinner, and Zeldes (1995) and Carroll and Samwick (1997), who estimate very similar labor income processes for different education groups in the population. These estimates are extensively used to calibrate heterogeneous-agent models. Unemployment risk is sometimes assumed to vary across demographic groups, consistent with empirical evidence (highly skilled workers have a lower unemployment rate). But this assumption implies more risks for the poor in these models, in contrast to our results.

[4] Throughout the paper, bold letters are used to denote vectors in order to distinguish them from scalars.

A. Financial Markets and Participation

Let s_t denote a date-event pair (state), which constitutes a complete description of uncertainty for all the economy that is realized in *t,* and let $s^t = (s_1, s_2, \ldots, s_t)$ be the history of states realized up through period *t.* For example, s_t will contain the realization of wages of *all* households, the return on all assets in the economy in period *t,* etc. Each node s_t branches out into S possible states (successor nodes) in the next period. There is a complete set (S) of Arrow securities (one-period contingent claims) available in every state, each paying one unit of consumption good in exactly one state of the world tomorrow. From this description, it is clear that all shocks are potentially insurable.

In addition, there is also a risk-free bond available in the economy. While this bond can be traded freely by all households (that is, without incurring any fixed or proportional transaction costs), the same does not hold for the Arrow securities. Households must pay a fixed cost of ψ^P in every period they participate in financial markets (which we will also refer to as the "stock market") where these securities are traded.[5]

To better understand the choices facing a typical household, it is useful to express the decision problem recursively. Each period a household decides whether to participate in the stock market in the current period by paying ψ^P, or to stay outside and trade the risk-free bond only. Define $\mathbf{q}(s)$ to be the $(1 \times S)$ price vector of the Arrow securities when the current state is *s,* and q_0 to be the bond price. Similarly, let the $(S \times 1)$ vector $\mathbf{k}_i \equiv (k_{1i}, k_{2i}, \ldots, k_{Si})$ denote a current stockholder's portfolio choice vector of Arrow securities, and k_{0i} be the bond holdings of a nonstockholder. We drop the time subscript, and denote next period's variables by primes. Then a household's problem is

$$\upsilon(\omega_i, \mathbf{X}_i; s) = \max[\upsilon^h(\omega_i, \mathbf{X}_i; s), \upsilon^n(\omega_i, \mathbf{X}_i; s)],$$

where

$$\upsilon^h(\omega_i, \mathbf{X}_i; s) = \max_{C_i, L_{1i}, L_{2i}, k_i}\{u(C_i, L_{1i}, L_{2i}, \mathbf{X}_i) + \beta E(\upsilon(\omega_i', \mathbf{X}_i'; s')|s)\}$$

s.t

$$C_i + \mathbf{q}(s)\mathbf{k}_i \leq \omega_i + \sum_{j=1,2}(1 - L_{ji})W_{ji}(s) - \Psi^P$$

$$\omega_i' = k_{s'i}$$

and

$$\upsilon^n(\omega_i, \mathbf{X}_i; s) = \max_{C_i, L_{1i}, L_{2i}, k_{0i}}\{u(C_i, L_{1i}, L_{2i}, \mathbf{X}_i) + \beta E(\upsilon(\omega_i', \mathbf{X}_{i;\, s'}')|s)\}$$

s.t

$$C_i + q_0(s)k_{0i} \leq \omega_i + \sum_{j=1,2}(1 - L_{ji})W_{ji}(s)$$

$$\omega_i' = k_{0i},$$

where υ^h and υ^n are the value functions of current stockholders and nonstockholders respectively, W_{1i} (s) and W_{2i} (s) denote the wages of the head and the spouse respectively after history s, and ω_i denotes financial wealth. Finally, although the arguments of the choice variables are suppressed for clarity of notation, they are all functions of the state vector (ω_i, $\mathbf{X}_i$; s).

Note the difference between the budget constraints of current stockholders and nonstockholders. In particular, stockholders choose an unrestricted $(S \times 1)$ portfolio vector, implying that they can transfer any (budget-feasible) amount of wealth to a given state in the next period. Thus, markets are dynamically complete within the stock market community. In contrast, nonstockholders are restricted to choosing a constant wealth level, k_0, for the next period.

B. Perfect Risk-Sharing

In the model presented above, households optimally enter and exit the stock market in different periods, and thus face complete markets and incomplete markets at different points in time. This is in contrast to the canonical model used to test for PRS in the previous literature where (under the null hypothesis) households face complete markets every period. Despite this difference, the main testable implication of PRS derived in the canonical model continues to hold in our model, but only for households who participate in the stock market in two consecutive periods.[6] For clarity of exposition then, we present the derivation of the PRS condition in the canonical model, which can be obtained from our general framework by simply setting $\Psi^P \equiv 0$.

In this case, without the participation cost, a household faces complete markets in every period, and hence maximizes the preferences given in equation (1) subject to a single lifetime budget constraint:

[5] This per-period fee is intended to capture several explicit and implicit costs of trading in financial markets, such as the time and effort it takes to monitor and rebalance one's portfolio, the additional time it takes to file tax returns, and so on. Vissing-Jørgensen (2000) estimates these costs from microdata and concludes that even modest fixed costs (about $150 per year) are sufficient to keep a large fraction of households out of the stock market. However, it seems plausible that certain types of participation costs (especially those associated with learning how financial markets work, etc.) are incurred only once. This would suggest that there could be a one-time fixed cost as well as a per-period fixed cost. In the working-paper version (Guvenen, 2003) we allowed for both of these costs. This extension made the participation decision dynamic (and the model more complicated) but had no substantive effect on our results. Thus, we abstract from the one-time cost in the current version.

[6] As shown below, precisely because the PRS condition holds for households who self-select into the stock market, PRS must be tested jointly with the selection equation.

$$E_0[\sum_{t=0}^{T} \beta^t \lambda(s_t)(C_{it} - \sum_{j=1,2} (1 - L_{jit})W_{jit})] \leq 0, \tag{2}$$

where $\beta^t\lambda(s_t)$ is the time-zero price of one unit of consumption in state s_t. This budget constraint equates the discounted lifetime value of expenditures to the discounted lifetime value of wage income. Let ϕ_i be the multiplier associated with constraint (2). The first-order condition for consumption choice is given by

$$\frac{1}{\phi_i} u_1(C_{it}, L_{1it}, L_{2it}, \mathbf{X}_{it}) = \lambda_t, \tag{3}$$

where the subscript of u denotes the partial derivative with respect to the indicated argument. Similarly, the labor supply decision of the head (assuming an interior solution) is determined by[7]

$$\frac{1}{\phi_i} u_2(C_{it}, L_{1it}, L_{2it}, \mathbf{X}_{it}) = \lambda_t W_{1it}. \tag{4}$$

Taking the ratio of equation (3) for periods $t - 1$ and t to eliminate the unobservable component, ϕ_i, we get

$$\frac{u_1(C_{it}, L_{1it}, L_{2it}, \mathbf{X}_{it})}{u_1(C_{it-1}, L_{1it-1}, L_{2it-1}, \mathbf{X}_{it-1})} = \frac{\lambda_t}{\lambda_{t-1}}. \tag{5}$$

This last equation clearly illustrates the main implication of PRS: the marginal utility growth of a given household (on the left-hand side) should only be a function of aggregate variables (the right-hand side) and hence should not be correlated with any idiosyncratic variable. In contrast, with *incomplete markets* there is no reason to expect marginal utility growth to be equated across households. The main test of PRS is then to estimate the relationship in equation (5) and then see if any idiosyncratic variable is correlated with the resulting error term. Because of the stock market participation decision in our framework, the PRS condition holds for households who are in the stock market in periods $t - 1$ and t.[8] Consequently, this self-selection must be taken into account when testing condition (5).

C. *The Participation Decision*

In general a closed-form solution for the participation decision is not available, although it is easy to see that a numerical solution can be obtained by solving backward starting from the last period of a household's life. The focus of this paper is on risk-sharing, and our interest in the participation decision is mainly for having a good specification of the characteristics of households who self-select into the stock market. Thus, rather than explicitly solving for the participation decision, we seek variables that determine this choice, which can be obtained from the optimization problem above.

A household enters the market in period t if υ^h $(\omega_{it}, \mathbf{X}_{it}; s^t) > \upsilon^n$ $(\omega_{it}, \mathbf{X}_{it}; s^t)$ and stays outside if the reverse holds, implying that the decision rule will be a function of the state vector: $(\omega_{it}, \mathbf{X}_{it}; s^t)$. Although in general s^t is a high-dimensional vector containing the entire history of asset prices and wages in the economy, it can be simplified substantially by making two observations. First, the efficient-markets hypothesis holds that asset returns are not predictable. Although there is some evidence of predictability of stock returns over long horizons, at shorter horizons (one to two years) stock returns are difficult to predict (see Guvenen, 2005, table 4). Consistent with this evidence, we assume that asset returns are i.i.d. over time, implying that the history of asset prices can be excluded from s. Second, there is substantial empirical evidence indicating that the persistence in individual wage dynamics can be adequately represented by an AR(1) process (which includes a random walk as a special case. See Browning, Hansen, & Heckman, 1999, for a detailed review of this evidence). This implies that the current wage (W_{1it}, W_{2it}) is a sufficient state variable for predicting the future evolution of a household's wage process. Putting these two pieces together, we obtain $s = (W_{1it}, W_{2it})$.

To sum up, the participation decision rule for a typical agent can be written as

$$d_{it} = 1\{\pi(\omega_{it}, W_{1it}, W_{2it}, \mathbf{X}_{it}) \geq 0\},$$

where $1\{\cdot\}$ is an indicator function, and $\pi(\cdot)$ is determined by the solution to the problem.[9] Clearly, the set of variables included in $\pi(\cdot)$ represent all the *potential* determinants of stockholding, and it is likely that empirically only a subset of them are significant factors in participation choice. For example, variables that affect υ^h and υ^n symmetrically will leave $\upsilon^h - \upsilon^n$ unchanged and will have no impact on participation. Thus, identifying the significant determinants of stockholding is ultimately an empirical question, which we address in section IVB.

D. *Empirical Specification*

Since perfect risk-sharing imposes restrictions on marginal utilities, the specification of the utility function is especially important for the purposes of this paper. Following Altug and Miller (1990), we assume the following period utility function:

$$u(C, L_1, L_2, \mathbf{X}) \equiv \delta_0(\mathbf{X}) C^{\rho_0} L_2^{\rho_2} + \delta_1(\mathbf{X}) L_1^{\rho_1} L_2^{\rho_3}. \tag{6}$$

[7] A condition analogous to equation (4) holds for the female labor supply decision, although we do not use it in our analysis.

[8] See the working-paper version (Guvenen 2003, appendix A), which shows that equation (5) indeed holds for the present model with $\Psi^p > 0$.

[9] Notice that other parameters of the model (including Ψ^p, ρ_0, etc.) will also play a role in the participation choice. But these parameters are assumed to be identical across the population—except those already summarized in $\mathbf{X}_{it}$—so they are all soaked up into the functional form of π.

This specification is quite flexible, and indeed, it is more general than most of those considered in the previous literature. One limitation of the PSID consumption data that we use in our empirical work is that it consists of only food expenditures. We thus interpret the specification in (6) as the first two subutilities of a more general utility function in which nonfood consumption enters in a separable manner. (This has been the maintained assumption in all previous studies using PSID. See section IVA for further discussion.) Then the first subutility can be interpreted as a Cobb-Douglas home-production function where food expenditures serves as capital and female leisure hours as labor input.

The second subutility captures the possible nonseparability between the leisure times of head and spouse. Nonseparable specifications in both subutilities have empirical support (Browning & Meghir, 1991, Altug & Miller, 1990). Another possibility is to have male leisure also enter the first subutility. But, first, Hayashi et al. (1996) test for this possibility and do not find support for it. Second, if in fact male leisure and consumption are nonseparable, tests based on equation (5) are invalid due to observational equivalence (see Attanasio & Davis, 1996, p. 1235, for a discussion of this point).

Some components of the vector $\mathbf{X}_{it}$ may not be observable to the econometrician. Hence, it is convenient to write $\mathbf{X}_{it} = (\mathbf{x}_{it}, \varepsilon_{0i}, \varepsilon_{1i}, \varepsilon_{0it}, \varepsilon_{1it})$, where $\mathbf{x}_{it}$ is a vector representing the observable component, and the remaining elements denote the unobservables. Each subutility is weighted by indices which are log linear functions of $\mathbf{X}_{it}$:

$$\delta_m(\mathbf{X}_{it}) = \rho_m^{-1} \exp(b_m \mathbf{x}_{it} + \varepsilon_{mi} + \varepsilon_{mit}), \quad m = 0, 1. \tag{7}$$

Here b_m is a fixed vector of coefficients; ε_{mi} represents the fixed household-effect, and ε_{mit} is a zero-mean disturbance term that varies both over time and across households. Further assumptions on the error terms will be stated in the next section. Note that each subutility is scaled by ρ_0 and ρ_1.

For tractability, we specialize the selection function $\pi(\cdot)$ to a linear form, which allows us to write the decision rule as a standard binary-choice equation. Substituting the observable and unobservable parts of $\mathbf{X}_{it}$, we obtain

$$d_{it} = 1\{\boldsymbol{\theta}\mathbf{y}_{it} + \eta_i - \eta_{it} \geq 0\}, \tag{8}$$

where θ is a fixed vector of coefficients; $y_{it} \equiv (\omega_{it}, W_{1it}, W_{2it}, \mathbf{x}_{it})$; $\eta_i \equiv \varepsilon_{0i}+\varepsilon_{1i}$; and $\eta_{it} \equiv -(\varepsilon_{0it} + \varepsilon_{1it})$.

E. Moment Conditions

Using the parameterization for preferences, the risk-sharing condition for stockholders (5) yields our first moment condition. After taking logarithms, first-differencing, and rearranging, we obtain

$$(\rho_0 - 1)\Delta c_{it} = \Delta \ln(\lambda_t) - b_0 \Delta \mathbf{x}_{it} - \rho_2 \Delta l_{2it} + \Delta \varepsilon_{0it}, \tag{9}$$

where Δ_t denotes the difference operator between t and $t-1$, and lowercase letters denote the natural logarithms of their uppercase counterparts (except for $\mathbf{x}_{it}$).

Although this equation by itself is sufficient to test for risk-sharing, it cannot identify all the structural parameters of the model. For that purpose, we add another moment condition which is valid for both the stockholders and the nonstockholders. Using equations (3) and (4), we get

$$\frac{u_2(C_{it}, L_{1it}, L_{2it}, \mathbf{X}_{it})}{u_1(C_{it}, L_{1it}, L_{2it}, \mathbf{X}_{it})} = W_{1it}. \tag{10}$$

This is the familiar intratemporal efficiency condition equating the marginal rate of substitution between consumption and leisure to the wage rate.[10] Notice that this is a static condition, which does not depend on market completeness, so it holds for nonstockholders as well. It is convenient to take logarithms, then difference equation (10) to obtain

$$\begin{aligned}(\rho_0 - 1)\Delta c_{it} = & -\Delta w_{1it} + (b_1 - b_0)\Delta \mathbf{x}_{it} \\ & - (\rho_2 - \rho_3)\Delta l_{2it} + (\rho_1 - 1)\Delta l_{1it} + (\Delta\varepsilon_{1it} - \Delta\varepsilon_{0it}).\end{aligned} \tag{11}$$

III. Econometric Method

Since the disturbance terms, $\Delta\varepsilon_{0it}$ and $(\Delta\varepsilon_{1it} - \Delta\varepsilon_{0it})$, in the equations above have zero mean by construction, it might seem reasonable to look at equations (9) and (11) as defining orthogonality conditions, which could then be estimated using GMM. However, this strategy is not directly applicable in this framework due to sample selection bias. To clearly see this, first consider the PRS condition (9). Under the null hypothesis, only stockholders are able to share risk perfectly, so the appropriate moment condition is

$$E[\Delta\varepsilon_{0it} | d_{it} d_{i,t-1} = 1] = 0, \tag{12}$$

which requires

$$\begin{aligned}&E[\varepsilon_{0it} | d_{it} = 1] = E[\varepsilon_{0it} - 1 | d_{i,t-1} = 1] \Rightarrow \\ &E[\varepsilon_{0it} | \eta_{it} \leq \boldsymbol{\theta}\mathbf{y}_{it} + \eta_i] = E[\varepsilon_{0it-1} | \eta_{it-1} \leq \boldsymbol{\theta}\mathbf{y}_{it-1} + \eta_i]\end{aligned} \tag{13}$$

to hold. In general both sides of this equation will be nonzero, because ε_{0it} is correlated with η_{it}, which we earlier defined as $-(\varepsilon_{0it} + \varepsilon_{1it})$. In other words, the unobservable preference shock that affects the risk-sharing condition, ε_{0it},

[10] Notice that the MRS equation holds as an equality only when the head is working in a given period, which can potentially cause another selection problem. Altug and Miller (1990) estimate a tobit specification for selection into the labor market and find that the error term in the selection equation has a small and insignificant correlation with the error in the MRS equation. Moreover, we eliminate far fewer households compared to these authors, since we only require the head to work for two consecutive years to be included in the estimation (whereas they require this for fourteen consecutive years), so this problem is probably even less critical in this case.

also influences the stock market participation decision, creating a selection bias. Furthermore, while it is possible that these conditional expectations are nonzero, but still equal to each other, this is not likely to be the case either, given that these expectations are functions of $\boldsymbol{\theta}\mathbf{y}_{it}$, and will vary over time as this selection index changes.

This last observation, however, suggests a way of eliminating the selection bias, and forms the basis of the estimator proposed by Kyriazidou (1997, 2001). To explain the basic idea of this estimator, we first make two assumptions.

Let Z_{it} be a vector of instruments.

Assumption A1. $\{(\varepsilon_{0it}, \varepsilon_{1it})\}_{t=1}^{T}$ is *i.i.d.* over time for all i, conditional on $\zeta_i \equiv \{\eta_i, \mathbf{Z}_{i0}, (y_{i1}, \ldots, y_{iT})\}$.

Assumption A2. $(\varepsilon_{0it}, \varepsilon_{1it})$ is independent of $Z_{i\tau}$ for all $\tau < t$, and for all i conditional on ζ_i.

The first assumption is the same as condition $(A1')$ in Kyriazidou (2001). An important implication of this assumption is that all regressors in the selection equation must be exogenous. The second assumption is a slight weakening of her assumption $(A2')$ that allows us to have endogenous variables in the main equation and to instrument for them using lagged dependent variables.[11] Because η_{it} is the sum of ε_{0it} and ε_{1it}, it also satisfies both assumptions above.

The idea behind Kyriazidou's estimator can be explained as follows. From the discussion above, it is clear that the term $E[\varepsilon_{0it} \mid \eta_{it} \leq \boldsymbol{\theta}\mathbf{y}_{it} + \eta_i]$ will remain unchanged if the selection index, $\boldsymbol{\theta}\mathbf{y}_{it}$, (and consequently the conditioning set) is constant in two consecutive periods. Thus, under assumptions A1 and A2, the following modified version of condition (12) holds:

$$E[\Delta\varepsilon_{0it} | d_{it}d_{i,t-1} = 1, \Delta\boldsymbol{\theta}\mathbf{y}_{it} = 0] = 0. \tag{14}$$

One immediately observes, however, that if y_{it} contains any continuous variables, the set of households that satisfy $\{\boldsymbol{\theta}\mathbf{y}_{it} = \boldsymbol{\theta}\mathbf{y}_{it-1}\}$ may be very small, or even empty. One strategy is then to assign a weight to each observation which is inversely proportional to the change in the index of that household, $\Delta\boldsymbol{\theta}\mathbf{y}_{it}$, such that asymptotically only observations with constant indices are included in the estimation.[12]

The estimator can then be implemented as follows. In the first step, the selection equation (8) is estimated to obtain an estimate of θ (denoted by $\hat{\theta}$). Then, using this estimate we construct weights which we take to be "kernel density" functions of the following form:

$$\psi_{it}^{N} = \frac{1}{h_N} K\left(\frac{\Delta\hat{\boldsymbol{\theta}}\mathbf{y}_{it}}{h_N}\right), \tag{15}$$

where $K(\cdot)$ is a scalar density function that satisfies certain regularity conditions (described in appendix B), and h_N is a sequence of "bandwidths" that tends to 0 as the sample size $N \rightarrow \infty$. For a fixed value of $\Delta\hat{\boldsymbol{\theta}}\mathbf{y}_{it}$ the weight ψ_{it}^{N} shrinks as N increases, while for fixed N, a larger deviation in the selection index corresponds to a smaller weight.

The kernel-weighted GMM estimator is constructed as follows. Let $\mathbf{f}(\boldsymbol{\alpha})$ denote a column vector of orthogonality conditions, $\mathbf{f}(\boldsymbol{\alpha}, i)$ be its sample counterpart for the ith observation, and $\boldsymbol{\alpha}$ be the vector of identifiable parameters in that system. For example, $\mathbf{f}(\boldsymbol{\alpha})$ could be obtained by interacting $\Delta\varepsilon_{0it}$ in equation (14) with some appropriate instruments from the set $\mathbf{Z}_{it}$. The key difference from a standard GMM estimator is in the construction of the sample counterparts of the moment conditions, which are multiplied by the kernel weights in our case:

$$\mathbf{G}_N(\boldsymbol{\alpha}) \equiv \frac{1}{N}\sum_{i=1}^{N} \psi_{it}^{N}\mathbf{f}(\boldsymbol{\alpha}, i).$$

Once $\mathbf{G}_N(\boldsymbol{\alpha})$ is obtained, we proceed as in the case of a standard GMM estimation:

$$\hat{\boldsymbol{\alpha}}_N = \arg\min[\mathbf{G}_N(\boldsymbol{\alpha})^T \Phi_N^T \Phi_N \mathbf{G}_N(\boldsymbol{\alpha})],$$

where Φ_N is a stochastic matrix that converges in probability to a finite nonstochastic limit Φ_0, and the superscript "T" denotes the transpose of a matrix. This estimator is consistent and asymptotically normal with $\sqrt{Nh_N}$ convergence rate (Kyriazidou, 2001). Further details of the estimation method are in appendix B.

IV. Estimation

A. *The PSID Data*

This section briefly describes the data and the variables used in the empirical analysis. Further details are provided in appendix A. The data are drawn from the Panel Study of Income Dynamics (PSID) covering the period from 1982 to 1993. Although PSID data are available from 1968 onward, data on stock ownership were not collected until 1984, making the earlier period unsuitable for our purposes. (Data from 1982–1983 are used for constructing instruments.) The availability of income and consumption data,[13] along with

[11] This is the only difference between Kyriazidou's (2001) original estimator and the one used here. Kyriazidou considers the case where all regressors in the main equation are either strictly exogenous or lagged endogenous variables. Instead in our case, consumption and leisure are likely to correlated with contemporaneous preference shocks, and so we intrument for these variables. Assumption A2 ensures that such instruments exist. Furthermore, there are a number of additional regularity assumptions that are required for the estimator discussed in appendix B.

[12] It is clear that the same approach can be used to correct for the selection bias among nonstockholders, which we do when testing for PRS among this group. In addition, a similar selection problem also exists in the estimation of the MRS condition (11) because our sample selection procedure described in the next section eliminates households who change their stockholding status during the sample period: $d_{it} \neq d_{i,t-1}$. Even though this moment condition holds for the whole population, unlike the PRS condition, the error term $(\Delta\varepsilon_{1it} - \Delta\varepsilon_{0it})$ has zero mean over the entire population, whereas we need this expectation to be zero over the sample that we observe: $d_{it} = d_{i,t-1}$.

[13] Although consumption data are restricted to food and a few other expenditure items. We discuss this limitation below.

detailed demographic information, has made PSID attractive for studying perfect risk-sharing, and not surprisingly it has been used extensively for this purpose in the previous literature (among others, Altug & Miller, 1990; Cochrane, 1991; Hayashi et al., 1996; and Hess & Shin, 2000). To the extent possible, we follow these studies in our sample selection criteria.

We use households from the core sample of PSID, which is a representative sample of U.S. households. In addition, we select a family into the sample in year t if the head (a) was in the study for four consecutive years including 1984 or 1989; (b) was married to the same spouse in t and $t - 1$; and (c) had positive labor hours in t and $t - 1$. We further eliminate households who had missing or inconsistent data on some key variables, as described in appendix A. Filtering out these observations leaves a total of 8,941 household-years (observations) that can be used in estimation.

For each household $i \in \{1, \ldots, N\}$ we have data on (a) annual leisure hours of head and spouse; (b) real average hourly earnings of head and spouse; (c) age of head and spouse, denoted age_{1it} and age_{2it} respectively; (d) real household food consumption expenditures (which is the sum of "food at home," "food away from home," and "the cash value of food stamps"); (e) number of household members, hsz_{it}; (f) completed education of head, E_{it}; and (g) a dummy indicating whether the household is a stockholder, d_{it}. These variables are available in every year from 1982 to 1993, except for food data which are missing in 1988 and 1989, and the stockholding variable which is only available in 1984 and 1989.

Table 1 provides the summary statistics of the data for both groups. First, note that there are roughly twice as many nonstockholders as there are stockholders in our sample (65% versus 35%). This point is important to keep in mind when interpreting the relative power of the tests of PRS in the two subsamples. Second, the average annual work hours, both for males and females, are similar across the two groups. Third, the average hourly earnings of stockholders is higher—by 70% for males, and by 50% for females—than that of nonstockholders. And fourth, the average food consumption is 18% higher for stockholders. However, despite these differences in levels, there is no significant difference between the dispersion of any of these variables across the two groups: the coefficient of variations of hours, earnings, and consumption are similar for stockholders and nonstockholders.

Finally, there are a few points concerning the use of food expenditures as the measure of consumption that should be addressed. First, separability between food and nonfood has been the maintained assumption in all studies on risk-sharing using PSID data, which makes our results comparable (Altug & Miller, 1990; Cochrane, 1991; Hayashi et al., 1996; Hess & Shin, 2000; etc.). Second, Atkeson and Ogaki (1996) provide evidence that food and nonfood consumption are separable.[14] Moreover, Ogaki and Zhang (2001) find virtually the same results regarding PRS when they replicate their tests using nondurable consumption instead of food expenditures. Third, a possible concern could be that food consumption may not be sufficiently variable, causing risk-sharing tests to have low power. But, if anything, the volatility of food consumption (from PSID) is higher than the volatility of nondurables consumption calculated from the Consumption Expenditure Survey. This is true for both stockholders and nonstockholders. Finally, in section VB we derive and test another implication of the perfect risk-sharing hypothesis that does not rely on consumption data. As we further discuss there, the results of that test confirm our findings using food expenditures.

B. *First Step: The Selection Equation*

The first step in the procedure is to estimate the parameter vector θ in the participation equation. This question has received a lot of attention in the recent literature, and as a result, there now exists a wealth of information on the empirical determinants of stock market participation (cf. Haliassos & Bertaut, 1995; Hurst, Luoh, & Stafford, 1998; Vissing-Jørgensen, 2000; Guiso, Haliassos & Jappelli, 2001; Curcuru et al., 2005). Many of these papers use data sets containing detailed wealth and portfolio information, such as the Survey of Consumer Finances (SCF), that are more suitable for this estimation than the PSID. Given this existing body of work, we do not reestimate the selection equation in this paper. Rather, we construct the kernel weights based on the findings of this literature, which we discuss here.

In an early paper, Haliassos and Bertaut (1995) estimate a logit model for stockholding choice using data from the SCF with a large number of explanatory variables that include demographics, preferences toward risk, income, wealth, and occupation. In particular, their list of variables contains all the variables included in our specification of the selection equation (8), except that they use labor income instead of wages. They find (a) race, (b) education, (c) risk aversion measures, (d) labor income, (e) financial net worth, and (f) "whether the individual has a managerial occupation" to be significant determinants of stockholding. Subsequent studies mentioned above have confirmed these findings using various alternative specifications and different data sets.

An important point to observe about these findings is that except for labor income and financial wealth, these explanatory variables represent individual (or household) characteristics that show little or no change over time. Since kernel weights are constructed based on the *time change* in the selection index, all but two of these regressors become redundant. Consequently, the coefficient estimates on these

[14] See also, however, Attanasio and Browning (1995), who argue against separability between food and nonfood consumption.

TABLE 1.—A LIST OF KEY VARIABLES AND THEIR SIMPLE STATISTICS

	Stockholders	Nonstockholders	All
Number of observations	3,178	5,763	8,941
Percentage of sample	34.8	65.2	100
Hours and earnings			
Average annual hours of head	2,213	2,177	2,189
	(646.1)	(686.5)	(672.1)
Average annual hours of spouse	1,451	1,501	1,483
	(741.5)	(706.8)	(718.2)
Average hourly earnings of head	\$17.83	\$10.41	\$12.99
	(13.79)	(7.45)	(9.76)
Average hourly earnings of spouse	\$10.10	\$6.82	\$7.96
	(9.08)	(5.71)	(6.88)
Average annual food consumption	\$5,249	\$4,419	\$4,708
	(2,806)	(2,253)	(2,445)
Demographic variables			
Average age of head	43.8	39.9	41.2
	(11.3)	(11.4)	(11.3)
Average education of head	6.07	4.9	5.31
	(1.54)	(1.62)	(1.59)
Average household size	3.3	3.6	3.5
	(1.13)	(1.21)	(1.18)

Notes: Standard deviations are in parentheses. The sample selection criteria are detailed in appendix A.

fixed characteristics do not affect the second step estimation, as they are always differenced out.

The conclusion we draw from this analysis is that, in order to correct for the selectivity bias, we need to mainly consider movements over time in labor income and financial wealth. Nevertheless, labor income and financial wealth are both endogenous variables and are thus likely to be correlated with the preference shifters included in the main equation (for example, the MRS and PRS conditions), which is not allowed by assumption A1 above. In our work then, we proxy for both variables with the sum of the head's wage and spouse's wage, $W_{1it} + W_{2it}$, and construct a univariate kernel with this variable only (see Guvenen (2003) for further justifications for using this proxy).

C. *Second Step: The Main Equation*

There are a number of different ways the risk-sharing hypothesis can be tested. The first and most obvious one is to estimate equation (9) alone for stockholders and use Hansen's J-test as a model specification test. If stockholders are not able to share risk perfectly, then their marginal utility growth cannot be explained by aggregate shocks alone, and the residuals will be correlated with idiosyncratic variables. By including household-level variables in the instrument set, this correlation will be caught by the J-test as a model specification error. This idea forms the basis of the previous tests implemented in the literature.

A second method, whose advantage will become clear in a moment, is the following: First estimate the MRS condition (11), which holds for the entire population. Then append equation (9) multiplied by d_{it} as an additional moment condition and estimate the two jointly, and test for PRS as an overidentifying restriction of the model. Specifically, if the additional orthogonality condition imposes p_1 extra restrictions and identifies p_2 additional parameters (and $p_1 > p_2$), then Nh_N (effective sample size) times the increment in the GMM criterion function has a χ^2 distribution with $(p_1 - p_2)$ degrees of freedom. The second approach has the advantage of exploiting more information, thereby increasing the efficiency of the estimator and the power of our hypothesis test. We will report the results using both approaches.

The Choice of Optimal Bandwidth: A standard Gaussian density is used for the kernel function $K(\cdot)$. As noted by Kyriazidou (2001), the choice of bandwidth is potentially more important for the performance of the estimator than the choice of kernel. The optimal bandwidth is determined with a cross-validation procedure described in appendix B. This procedure yields an optimal bandwidth of $h_N^* = 0.24$. However, the cross-validation criterion function is quite flat between 0.2 and 0.5, although it increases steeply outside this region. Due to the exponential nature of weights, small differences in the value of h_N in this range results in large variations in kernel weights. For example, a household whose selection index changes by 50% between two periods is weighted by 0.61, 0.13, and 0.004 for $h_N = 0.5$, 0.24, and 0.15, respectively. To make sure that our conclusions are robust to values of h_N in this range, in the next section we will also report the results for $h_N = 0.5$ as well.

V. Results

In this section we report our empirical findings. First, we present the results obtained from tests of risk-sharing for stockholders, nonstockholders, and the whole population. Then we consider several extensions and alternative tests to examine for the robustness of these results. We conclude by discussing parameter estimates.

In PSID, consumption (food) data is not available in 1988 and 1989, which leaves us with six time differences that can be used in estimation: 1984–85, 1985–86, 1986–87, 1990–91, 1991–92, 1992–93.

Instruments. Our main set includes the following seven variables for the estimation using the time difference between $t - 1$ and t: a constant; age of head at time t, age_{1it}; age of spouse at time t, age_{2it}; the contemporaneous change in household size, Δhsz_{it}; twice-lagged change in log consumption, $\Delta c_{i,t-2}$; twice-lagged change in log wage of head, $\Delta w_{1i,t-2}$; and the twice-lagged change in log spouse's wage, $\Delta w_{2i,t-2}$.

This instrument set is used for the PRS condition. Notice that we have not included the first lags of variables that are susceptible to measurement error, such as consumption and wages, because the resulting correlation with variables in the PRS equation would make them invalid. For the MRS condition, we use the same instruments as in the PRS condition, with two additions (to increase the precision of the estimates of preference parameters): we include the dummy for stockholding status, d_{it}; and we use the levels of household size in $t - 1$ and t, hsz_{it-1} and hsz_{it}, instead of the change.

For the empirical specification of $\mathbf{x}_{it}$, we choose a square and a cubic polynomial of head's age, ${age_{1it}}^2$, ${age_{1it}}^3$, and the household size, hsz_{it}.

A. Tests of Risk-Sharing

Table 2 presents the main result of the paper. We first estimate the PRS condition for stockholders only, using the optimal bandwidth ($h_N^* = 0.24$). If perfect risk-sharing holds, this equation should adequately describe the marginal utility growth for stockholders, and the model specification *J*-test should not reject the estimated moment condition. However, this is not the case: in column 1, the PRS condition has a *p*-value of 0.004, strongly rejecting perfect risk-sharing among stockholders.

Next we turn to nonstockholders. As discussed in the introduction, nonstockholders also have access to informal or nonmarket insurance mechanisms, and may face different (or even fewer types of) risks than stockholders. Therefore, it is conceivable that they might be able to share risk effectively among themselves. To investigate this possibility then, we now test for PRS only among nonstockholders, which is reported in the second column. The *p*-value of the test is 70%, showing no evidence against perfect risk-sharing among nonstockholders.

To examine whether this result is sensitive to the assumed bandwidth for the kernel function, we repeat the same test with a wider bandwidth: $h_N = 0.5$ (columns 4 and 5). The *p*-value for stockholders slightly increases to 0.009, still indicating a rejection at the 1% level. The *p*-value for nonstockholders falls to 0.421, but is still far away from rejection. Similarly, using a tighter bandwidth of $h_N = 0.15$ (columns 6 and 7) has only a small effect on these results: the *p*-value for stockholders is 0.025, and for nonstockholders it rises to 0.66. As noted earlier, because of the exponential nature of the kernel weights, the effective sample size shrinks quickly with a smaller bandwidth, which might be partly responsible for the slightly higher *p*-values for the stockholders in this latter case.

B. PRS Tests: Robustness

Before looking for interpretations for these findings, it is important to address several issues regarding the robustness of these results. The first question is whether the model specification test could be rejecting the PRS moment condition because of invalid instruments. One way to check this possibility is to estimate the MRS equation with instruments that were also included in the instrument set for the PRS equation. This will be informative because under the null hypothesis of perfect risk-sharing, the error term in the PRS equation is $\Delta\varepsilon_{0it}$, which also appears in the MRS equation (compare equations 9 and 11). So, if some instruments are invalid, they are also likely to result in the rejection of the MRS equation. In column 3 of table 2, we estimate the MRS equation using data on all households in the sample. The *J*-test has a *p*-value of 0.166, showing no evidence against the MRS equation, and consequently, against the validity of the instruments. Moreover, if instruments were indeed invalid, it is not clear why this would affect the moment conditions of the stockholders significantly while not being revealed in the nonstockholders' moment conditions.

Second, could these results be explained by the poor finite sample properties of the kernel-weighted GMM estimator? For example, if the model specification test tends to over-reject in small samples, it will be more likely to reject the

TABLE 2.—TESTS OF RISK-SHARING USING THE PRS EQUATION ONLY

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		$h_N^* = 0.24$		$h_N = 0.50$		$h_N = 0.15$	
Group	*H*	*N*	All	*H*	*N*	*H*	*N*
Moment conditions	PRS	PRS	MRS	PRS	PRS	PRS	PRS
Test statistics							
χ^2	52.1	24.3	54.1	49.8	29.9	45.7	25.3
df	29	29	45	29	29	29	29
p-value	0.004	0.709	0.166	0.009	0.421	0.025	0.660

Notes: *H* and *N* denote stock*H*olders and *N*onstockholders respectively. *P*-value (PRS) refers to the significance level associated with the PRS moment condition; *df* is the degrees of freedom for the moment conditions in a given column. The instrument set for the PRS equation includes a constant, age of head, age of spouse, change in household size, consumption growth lagged twice, and head's and spouse's wage growth lagged twice. The instrument set for MRS adds a dummy indicating stock ownership to the previous list, and uses the levels of household size instead of its change.

TABLE 3.—TESTS OF RISK-SHARING USING LEAD INSTRUMENTS

	$h_N^* = 0.24$		$h_N = 0.50$	
Group	*H*	*N*	*H*	*N*
Moment conditions	PRS	PRS	PRS	PRS
Test statistics				
χ^2	48.7	33.8	52.2	36.4
df	29	29	29	32
p-value (PRS)	0.012	0.247	0.005	0.162

Notes: The instrument set includes a constant, age of head, age of spouse, change in household size, and change in head's log wage from $t - 3$ to $t + 1$, and from $t - 2$ to $t + 1$.

null for stockholders than for nonstockholders (since the sample size is smaller for the former group). However, note that the effective sample size for the estimator is Nh_N. So, for example, the stockholders' effective sample size when $h_N = 0.5$ ($3{,}178 \times 0.5 = 1{,}589$) is close to that of nonstockholders when $h_N = 0.24$ ($5{,}763 \times 0.24 = 1{,}383$). Similarly, the effective sample size of nonstockholders when $h_N = 0.15$ is close to that of stockholders when $h_N = 0.24$. Since the results of PRS tests are the same for all these bandwidth sizes, this particular concern does not appear to be critical. Furthermore, consistent with this finding, the Monte Carlo evidence in Kyriazidou (1997, 2001) suggests that the small sample properties are quite well-behaved for sample sizes around those considered in this paper.[15]

Third, is it possible that we fail to reject PRS in nonstockholders' sample because there is more unobserved variability (and hence less information) in that sample, whereas the reverse is true in stockholders' sample, leading to a rejection? This does not seem likely to be the case, because the nonstockholders' sample is twice as large as that of stockholders (table 1), so everything else equal specification tests should have more power to reject in the former sample. As another reflection of this fact, in the next section we find that the estimates obtained from the nonstockholders' sample are more precise than those from the stockholders' sample. Thus, it is unlikely that less information (or larger variances) in the nonstockholders' sample accounts for differences in the PRS test results.

Fourth, as noted by Hayashi et al. (1996), tests of perfect risk-sharing may not have high power against the alternative of self-insurance if the instrument set only includes lagged values of variables such as wages and consumption.[16] To investigate this possibility, we replace the lagged wage changes ($\Delta w_{1i,t-2}$ and $\Delta w_{2i,t-2}$) with the bracketing wage changes of the head from $t - 2$ to $t + 1$ and from $t - 3$ to $t + 1$ in the main instrument set used before, as suggested by Hayashi et al. As table 3 displays, PRS is rejected for stockholders as before, with *p*-values between 0.005 and 0.012. For the nonstockholders, although the *p*-value is lower than before (0.162 and 0.247), the PRS cannot be rejected at conventional significance levels.

A final concern discussed earlier is about, the use of food expenditures as a measure of consumption. It is possible to test for PRS without relying on consumption data. To develop such a test, note that PRS also imposes structure on the cross section of marginal utility of leisure growth through equation (4). After taking logs, differencing, and rearranging this equation (for household heads that are working in years t and $t - 1$), we get

$$\rho_2 \Delta l_{2nt} = \Delta w_{1nt} + \Delta \ln(\lambda_t) - \mathbf{b}_0 \Delta x_{nt} - (\rho_1 - 1)\Delta l_{1nt} - \Delta \varepsilon_{0nt}. \tag{16}$$

The results reported in table 4 are similar to those found above: using the optimal bandwidth, risk-sharing is rejected for stockholders with a *p*-value of 0.009, but not for nonstockholders (*p*-value = 0.471). Increasing h_N to 0.50 has little effect as in previous cases.

Before closing this section, we compare our findings to existing work reviewed in the introduction, which strongly rejected perfect insurance in the whole population. We repeat our main test of PRS for the whole population, and append the MRS equation to increase the power of the test. In column 1 of table 5, the overidentifying restriction for PRS has a *p*-value of 0.013, rejecting the null of perfect risk-sharing for the whole population.

To further increase the power of the test, we add a simple wage equation (C2)—which expresses the wage of the head as a function of individual characteristics—as an additional moment condition (see appendix C for the description of this equation). This equation acts as a "seemingly unrelated regression" and is also used to obtain more precise estimates in the next section. In the next column, we report the results with this wage equation added: the *p*-value of PRS now goes further down to 4×10^{-4}. Considering a higher bandwidth value of $h_N = 0.5$ makes the rejection even stronger with *p*-values never higher than 10^{-4}. Thus, our results are consistent with earlier results, indicating strong rejection of perfect risk-sharing for the whole population. In

TABLE 4.—TESTS OF RISK-SHARING USING THE MARGINAL UTILITY OF LEISURE EQUATION

	$h_N^* = 0.24$		$h_N = 0.50$	
Group	*H*	*N*	*H*	*N*
Moment conditions	PRS	PRS	PRS	PRS
Test statistics				
χ^2	49.7	28.9	51.7	30.1
df	29	29	29	29
p-value (PRS)	0.009	0.471	0.006	0.409

Notes: The instrument set is the same one as the one used in table 2.

[15] Another well-known problem is that small sample properties could deteriorate when the number of instruments is large. To address this possibility, the working-paper version (Guvenen, 2003, table 4) reports additional tests of PRS, where we eliminate several instruments to reduce the degrees of freedom. The results reported there show that this has no appreciable effect on the test results.

[16] This is because even with incomplete markets, the permanent income hypothesis implies that lagged endogenous variables will be uncorrelated with current forecast errors. If, further, the forecast error can be written as the sum of an aggregate and an idiosyncratic component, then lagged variables will have zero correlation with the idiosyncratic component even when markets are incomplete. See Hayashi et al. (1996) for further discussion.

TABLE 5.—TESTS OF RISK-SHARING IN THE WHOLE POPULATION

	(1)	(2)	(3)	(4)
	$h_N^* = 0.24$		$h_N = 0.50$	
Group	$H + N$	$H + N$	$H + N$	$H + N$
Moment conditions	MRS & PRS	MRS & PRS & WAGE	MRS & PRS	MRS & PRS & WAGE
Test statistics				
χ^2	102.6	137.67	149.2	196.0
df	74	114	74	114
p-value (model)	0.016	0.003	0.000	0.000
p-value (PRS)	0.013	0.000	0.000	0.000

Notes: The instrument set is the same one as the one used in table 2. For the wage equation, (C2), we exclude spouse's twice-lagged wage change and twice-lagged consumption growth, but add the education of head, E_{1nt}, to the instrument set above.

light of these findings, it seems that the rejection of PRS in the whole population obtained in the literature is likely to be due to the rejection of PRS not among the nonstockholders, but instead among the stockholders.

C. Parameter Estimates

All the structural parameters of the model can be identified by jointly estimating (a) the MRS equation and (b) the PRS condition for either stockholders or nonstockholders. The wage equation (C2) is added as a third moment condition to obtain more precise estimates. Because the PRS equation is strongly rejected for stockholders, but not for nonstockholders, in table 6 we report the estimates of structural parameters obtained by estimating the PRS equation for nonstockholders only, in addition to the MRS and wage equation estimated for the whole population.

TABLE 6.—ESTIMATES OF THE STRUCTURAL PARAMETERS

	Utility Function:		
	$\delta_0(\mathbf{X})C^{\rho_0}L_2^{\rho_2} + \delta_1(\mathbf{X})L_1^{\rho_1}L_2^{\rho_3}$		
Bandwidth Model rejected at 5%?	$h_N^* = 0.24$ No	$h_N = 0.50$ No	$h_N = \infty$ Yes
Curvature Parameters			
ρ_0	0.27	0.49	0.99
	(0.43)	(0.33)	(3.63)
ρ_1	−6.16	−4.83	−6.26
	(1.96)	(1.10)	(41.54)
ρ_2	−24.1	−16.32	−0.87
	(31.2)	(28.1)	(265.8)
ρ_3	−42.0	−29.90	−34.2
	(27.4)	(23.7)	(336.1)
Demographic Effects			
b_0^1 (age-squared)	−0.88	−0.42	−0.01
	(2.55)	(1.47)	(2.47)
b_0^2 (age-cubed)	−2.50	−1.18	−0.03
	(1.24)	(0.77)	(3.64)
b_0^3 (family size)	0.67	0.32	−0.01
	(1.05)	(0.68)	(3.88)
b_1^1 (age-squared)	−1.16	−0.54	−0.43
	(2.65)	(1.48)	(4.12)
b_1^2 (age-cubed)	−2.30	−1.12	0.17
	(2.63)	(1.39)	(3.65)
b_1^3 (family size)	1.24	0.92	1.54
	(2.54)	(1.44)	(13.64)
Test Statistics			
χ^2 (model)	124.3	120.5	149.2
df (model)	114	114	114
p-value (model)	**0.239**	**0.320**	**0.014**
p-value (PRS)	0.566	0.762	0.022

Notes: The moment conditions used are the MRS and wage equations (11 and C2) for the whole population, and the PRS equation (9) for nonstockholders. The structural parameters are exactly identified, and the standard errors for parameter estimates are in parentheses.

The first column presents the parameters obtained when the optimal bandwidth value is used. First, the curvature of consumption, ρ_0, is estimated to be 0.27 and is not significantly different from 0 suggesting logarithmic consumption preferences. In the next column, which reports the results when the bandwidth is $h_N = 0.50$, the estimate of ρ_0 is only slightly higher (0.49) and still not significantly different from 0. This result is consistent with earlier studies using PSID data (cf. Altug & Miller, 1990).

Second, the curvature of male leisure is estimated to be −6.16 and is statistically significant at the 1% level. The implied elasticity of male labor supply with respect to wages (holding the marginal utility of wealth constant) is $L_{1it}\,((1 - L_{1it})\,(1 - \rho_1))^{-1}$, which can easily be derived from the first-order condition for labor choice, equation (4). Given that the average time spent at work is approximately 2,200 hours per year in our sample (see table 1), assuming 16 hours of discretionary time per day, we get $L_{1it} = 0.37$. Then the implied elasticity is 0.08. Similarly, if we take the estimate of $\rho_1 = -4.83$ from column 2, the implied elasticity would be 0.10. These values are within the range found in the previous literature, reported for example in Browning et al. (1999, table 3.3). As for the curvature parameters of female leisure, ρ_2 and ρ_3, while the point estimates are negative, neither one is statistically different from 0. The negative signs are consistent with the estimates of Altug and Miller (1990) and Hayashi et al. (1996).

The estimated coefficients of household characteristics seem reasonable. The coefficients on the age polynomial are all negative, indicating that, everything else held constant, the utility derived from both subutilities decreases with age. The structural coefficient of hsz_{it} is positive, which means that an increase in household size (which is mostly due to a new child, since our sample contains only married couples) increases both subutilities. The implied Frisch elasticity of male labor supply with respect to family size is $(b_1^3 hsz_{it})\, L_{1it}\, ((1 - L_{1it})\,(\rho_1 - 1))^{-1}$. Evaluating this formula using the estimate of ${b_1}^3 = 1.24$ from the first column, and the average family size of 3.5 from table 1, yields an elasticity of −0.36. The negative value suggests that a new child reduces the work hours of the head, most probably in response to the increased home production demand associated with child rearing.

Finally, the last column reports the parameter estimates when no kernel is used: $h_N = \infty$. First, notice in the last row that the model is rejected at any significance level higher than 1.4%. Second, while the parameter estimates are not dramatically different, the standard errors are substantially higher in most cases. Similarly, we found that when stockholders' PRS condition is used (instead of nonstockholders') the parameter estimates are erratic and often have the wrong sign with large standard errors (available upon request). As noted above, in this case the joint moment conditions (MRS, PRS, and wage equations) are always rejected regardless of kernel bandwidth.

VI. Discussion and Conclusion

In this paper we found strong evidence against perfect risk-sharing among stockholders, but we were unable to reject it among nonstockholders. Furthermore, risk-sharing in the whole population is also strongly rejected. Overall, these findings suggest that the failure of PRS in the whole population, also found by many previous studies, is likely to be due to the failure of the wealthy to insure the additional risks they face.

One potential source of risk faced primarily by the wealthy (stockholders) is entrepreneurial income risk.[17] The literature offers several reasons—based on asymmetric information (agency) problems—for why entrepreneurial income is difficult to insure (Bitler, Moskowitz, & Vissing-Jørgensen, 2005). Thus, while stockholders may have access to additional insurance opportunities, their income is harder to insure as well.[18] In contrast, the main sources of income for nonstockholders are wages and salaries, which already include several implicit or explicit sources of insurance (unemployment and disability insurance, long-term contracts, labor hoarding behavior by firms, and so on). In addition, a number of informal risk-sharing mechanisms (for example, inter vivos transfers, charitable donations, and borrowing and lending) further eliminate the risks faced by most households. Note that many of these insurance opportunities may not be effective in insuring the potentially large losses experienced by business owners.

A second potential source of additional risk for stockholders could be the stock market itself. If stockholders face certain trading frictions, such as information acquisition costs (for each stock they trade), transactions costs, and short-selling constraints, this could easily prevent them from forming optimal portfolios. As a result, if each stockholder holds only a few stocks instead of a well-diversified optimal portfolio, there is no reason to expect that they will be equating their marginal utility growth, since each investor is exposed to substantial idiosyncratic uncertainty of the stocks in their portfolio. Indeed, empirical studies have shown that, especially before the 1990s, more than one-third of stockholders were seriously underdiversified, holding a portfolio containing five stocks or fewer (see Curcuru et al., 2005, for a detailed review of these empirical facts). However, during the 1990s, investors increasingly switched from holding individual stocks to mutual funds, improving diversification (although still far from perfect). Although it would be interesting to explore whether this trend has improved risk-sharing among stockholders, unfortunately PSID does not contain detailed portfolio information to conduct such an analysis.

Accounting for differences in idiosyncratic risks between the wealthy and the poor could potentially help explain several differences in behavior between these groups, such as why the wealthy have a higher savings rate, and why they demand such a high return for holding risky assets. Overall these results underscore the need to focus on risks faced by wealthy households as important sources of market incompleteness. Further work is needed to pinpoint the exact types of risks that are uninsurable for wealthy households.

[17] According to the 1995 Survey of Consumer Finances, the fraction of business owners (with a business value above $10,000) among nonstockholders is only 3%.

[18] Although it is beyond the scope of this paper, in principle it is possible to identify the role of entrepreneurial risk for stockholders. In particular, looking at stockholders only, one can test the PRS separately for those who own private businesses and those who do not. If risk-sharing is rejected for the former group but not for the latter, then this would suggest entrepreneurial risk as an important source of market incompleteness. However, entrepreneurship introduces another selection equation that must be dealt with as well. So, we leave this question for future research.

REFERENCES

Aiyagari, S. Rao, "Uninsured Idiosyncratic Shock and Aggregate Saving," *Quarterly Journal of Economics* 109:3 (1994), 659–684.

Altug, Sumru, and Robert A. Miller, "Household Choices in Equilibrium," *Econometrica* 58:3 (1990), 543–570.

Angeletos, George-Marios, "Uninsured Idiosyncratic Investment Risk and Aggregate Saving," National Bureau of Economic Research working paper no. 11180 (2005).

Atkeson, Andrew, and Masao Ogaki, "Wealth-Varying Intertemporal Elasticities of Substitution: Evidence from Panel and Aggregate Data," *Journal of Monetary Economics* 38:3 (1996), 507–534.

Attanasio, Orazio P., and Martin Browning, "Consumption over the Life Cycle and over the Business Cycle," *American Economic Review* 85:5 (1995), 1118–1137.

Attanasio, Orazio P., and Steven J. Davis, "Relative Wage Movements and the Distribution of Consumption," *Journal of Political Economy* 104:6 (1996), 1227–1262.

Bierens, Herman J., "Kernel Estimators of Regression Functions," in Truman F. Bewley (Ed.) *Advances in Econometrics: Fifth World Congress* (Cambridge, MA: Cambridge University Press, 1987).

Bitler, Marianne P., Tobias J. Moskowitz, and Annette Vissing-Jørgensen, "Testing Agency Theory with Entrepreneur Effort and Wealth," *Journal of Finance* 60:2 (2005), 539–576.

Browning, Martin, Lars P. Hansen, and James J. Heckman, "Micro Data and General Equilibrium Models," in John B. Taylor and Michael Woodford (Eds.), *Handbook of Macroeconomics,* vol. 1B (Amsterdam: Elsevier Science B.V., 1999).

Browning, Martin, and Costas Meghir, "The Effects of Male and Female Labor Supply on Commodity Demands," *Econometrica* 59:4 (1991), 925–951.

Cagetti, Marco, and Mariacristina De Nardi, "Entrepreneurship, Frictions, and the Distribution of Wealth," Federal Reserve Bank of Minneapolis staff report no. 322 (2003).
Carroll, Christopher, and Andrew A. Samwick, "The Nature of Precautionary Wealth," *Journal of Monetary Economics* 40:1 (1997), 41–71.
Chari, V. V., Mikhail Golosov, and Aleh Tsyvinski, "Business Start-ups, the Lock-in Effect, and Capital Gains Taxation," University of Minnesota working paper (2005).
Cochrane, John H., "A Simple Test of Consumption Insurance," *Journal of Political Economy* 99:5 (1991), 957–976.
Constantinides, George M., and Darrell Duffie, "Asset Pricing with Heterogeneous Consumers," *Journal of Political Economy* 104:2 (1996), 219–240.
Curcuru, Stephanie, John Heaton, Deborah Lucas, and Damien Moore, "Heterogeneity and Portfolio Choice: Theory and Evidence," University of Chicago working paper (2005).
Davidson, Russell, and James G. MacKinnon, *Estimation and Inference in Econometrics* (Oxford: Oxford University Press, 1993).
Gentry, William M., and R. Glenn Hubbard, "Entrepreneurship and Household Saving," Columbia University working paper (2000).
Gourinchas, Pierre-Olivier, and Jonathan A. Parker, "Consumption Over the Life Cycle," *Econometrica* 70:1 (2002), 47–89.
Guiso, Luigi, Michael Haliassos, and Tullio Jappelli, "Household Stockholding in Europe: Where Do We Stand and Where Do We Go?" *Economic Policy* (2001).
Haliassos, Michael, and Carol C. Bertaut, "Why Do So Few Hold Stocks?" *The Economic Journal* 105 (1995), 1110–1129.
Hayashi, Fumio, Joseph Altonji, and Laurence Kotlikoff, "Risk-Sharing Between and Within Families," *Econometrica* 64:2 (1996), 261–294.
Heaton, John, and Deborah Lucas, "Asset Pricing and Portfolio Choice: The Importance of Entrepreneurial Risk," *Journal of Finance* 55:3 (2000), 1163–1198.
Guvenen, Fatih, "Does Stockholding Provide Perfect Risk Sharing?" University of Rochester working paper (2003).
——— "A Parsimonious Macroeconomic Model for Asset Pricing," University of Texas at Austin working paper (2005).
——— "Reconciling Conflicting Evidence on the Elasticity of Intertemporal Substitution: A Macroeconomic Perspective," *Journal of Monetary Economics* 53:7 (October 2006), 1451–1472.
Hess, Gregory D., and Kwanho Shin, "Risk-Sharing by Households within and across Regions and Industries," *Journal of Monetary Economics* 45:3 (2000), 533–560.
Hubbard, R. Glenn, Jonathan Skinner, and Stephen P. Zeldes, "Precautionary Saving and Social Insurance," *Journal of Political Economy* 103:2 (1995), 360–399.
Hurst, Erik, Ming Ching Luoh, and Frank P. Stafford, "The Wealth Dynamics of American Families, 1984–94," *Brookings Papers on Economic Activity* 1998:1 (1998), 267–337.
Investment Company Institute, "Equity Ownership in America," research report (2002), available online at http://www.ici.org/statements/res.
Krusell, Per, and Anthony A. Smith, "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy* 106:5 (1998), 867–896.
Kyriazidou, Ekaterini, "Estimation of a Panel Data Sample Selection Model," *Econometrica* 65:6 (1997), 1335–1364.
——— "Estimation of Dynamic Panel Data Sample Selection Models," *Review of Economic Studies* 68:3 (2001), 543–572.
Nelson, Julie, "On Testing for Full Insurance Using Consumer Expenditure Survey Data," *Journal of Political Economy* 102:2 (1994), 384–394.
Ogaki, Masao, and Qiang Zhang, "Decreasing Relative Risk Aversion and Tests of Risk-Sharing," *Econometrica* 69:2 (2001), 515–526.
Storesletten, Kjetil, Chris Telmer, and Amir Yaron, "Asset Pricing with Idiosyncratic Risk and Overlapping Generations," Carnegie-Mellon University working paper (2001).
Townsend, Robert M., "Risk and Insurance in Village India," *Econometrica* 62:3 (1994), 539–591.
Vissing-Jørgensen, Annette, "Towards an Explanation of Household Portfolio Choice Heterogeneity: Nonfinancial Income and Participation Cost Structures," University of Chicago working paper (2000).

APPENDIX A

Data

Starting with the PSID family files covering 1982–1993 waves, we use the following sample selection criteria to select our main sample. Specifically, we include household-years in t and $(t - 1)$ in estimation if the head of the family

(a) is in the study for at least four consecutive years $(t - 3, \cdots, t)$ including 1984 or 1989,
(b) is married to the same spouse at least in the last two years $(t - 1, t)$ of the same period,
(c) has a positive labor income at least in the last two years $(t - 1, t)$ of the same period.

These criteria produced a sample of 2,350 households who were in the study between 1984 and 1993, not necessarily for all years. We further eliminate a household-year if

(d) annual family food consumption expenditure is less than $150,
(e) the head's education variable is missing for the last two years $(t - 1, t)$ of this period,
(f) if the head's or spouse's reported annual labor hours exceeded 4,860 hours,
(g) if the head or spouse had positive annual labor hours but zero annual labor income, or vice versa.

Conditions (d)–(g) are similar to those used to eliminate irregular observations in the literature (cf. Altug & Miller, 1990; Hayashi et al., 1996).

(h) Finally, if a household changed its stockholding status from 1984 to 1989, we eliminate that observation from estimation between these two dates.

An important concern is coding errors. To eliminate potential outliers, we first isolated observations on total consumption, head's wage, and spouse's wage, if the following bound was violated: $E(x_t) - 2std(x_t) \leq x_t \leq 2E(x_t) + 2std(x_t)$, where x_t denotes the variable. There were a total of 46 observations violating this bound for at least one of the three variables. Upon closer inspection of the time series of these variables, we eliminated 41 observations which had small standard deviations and an outlier that was significantly away from the sample average. These criteria produced a total of 8,941 of observations. The breakdown for each moment condition is as follows: 1,292 observations for the 1983–1984 moments, 1,289 for 1984–1985, 1,302 for 1985–1986, 1,761 for 1989–1990, 1,709 for 1990–1991, and 1,588 for 1991–1992.

Wages: The average hourly labor earnings (wages) of head and spouse reported in PSID and adopted in this paper are calculated from the sum of the following types of income and total annual hours: V19127, Labor Part of Farm Income; V19128, Labor Part of Business Income; V19129, Salary Income; V19131, Bonuses, Overtime, Commissions; V19132, Income from Professional Practice or Trade; V19133, Labor Part of Market Gardening Income; V19134, Labor Part of Roomers and Boarders Income.

Stockholding: The definition of stockholding adopted in this paper includes ownership of shares of stock in publicly held corporations, mutual funds, and investment trusts, including stocks in IRAs. This definition corresponds to PSID variables V10912 for 1984 and V17325 for 1989. All households that indicate they do not own any of these assets are considered nonstockholders that year.

This definition of stockholding does not include indirect ownership of stocks through pension funds. Notice that, first, indirect holding was more modest during the 1980s and has become much more popular in the 1990s (see Investment Company Institute, 2002, for direct and indirect stock ownership rates). And second, pension funds impose several restrictions on the use of funds, which is not consistent with our null hypothesis that stockholders can optimally use their assets to insure against shocks to their budget sets.

PSID collects stock ownership data every five years (from 1984 on), whereas for our empirical work we need this information for every year. We identify a household as a stockholder (alternatively, nonstockholder) in every year between 1984 and 1989, if the household is present in the sample in both years as a stockholder (nonstockholder). Second, if a household switches between these two groups from 1984 to 1989, we eliminate those observations from the sample between these two dates since we are not able to determine when the switch exactly happens. Clearly, this step creates another selection bias, which the econometric method is able to handle as explained in the text. Finally, for years after 1989 we take the status of a household as it is given in 1989. This identification scheme is clearly not perfect, but notice that the estimation method asymptotically assigns zero weight to an observation if the probability of being a stockholder changes ($\Delta\theta y_{it} \neq 0$). Thus, a household that moves into or out of the stock market between 1984 and 1989 will receive a small (and asymptotically zero) weight in estimation, because this move is likely to be accompanied by a change in the selection index.

APPENDIX B

Estimation

We describe the procedure for the estimation of the PRS condition (9) for the stockholders; the case for the MRS equation and the wage equation (equation [C2]) described in the next appendix) are analogous. First, in order to construct $f(\alpha,i)$, for each year t we pick $(r \times 1)$ dimensional vectors of instruments Z_{it} satisfying

$$E[\mathbf{Z}_{it}\Delta\varepsilon_{0it}|\Delta\theta\mathbf{y}_{it} = 0, d_{it}d_{i,t-1} = 1] = 0. \tag{B1}$$

Denote these moment conditions for year t by $f_t(\alpha,i)$. From the moment conditions (9) and (11), it is clear that for testing PRS and identifying the structural parameters of the model, one only needs data in two consecutive periods. Thus, we reduce the panel data estimation into cross section by forming the following T^*r dimensional vector where T^* is our panel length:

$$\mathbf{f}(\boldsymbol{\alpha},i) = (\mathbf{f}_1(\boldsymbol{\alpha},i), \ldots, \mathbf{f}_{T^*}(\boldsymbol{\alpha},i))'. \tag{B2}$$

$T^* = 6$ and $r = 39$ for the PRS equation ($r = 51$ for the MRS equation, and 47 for the wage equation). We then construct the sample counterparts by weighting each observation by its corresponding kernel weight: $\mathbf{G}_N(\boldsymbol{\alpha}) \equiv \frac{1}{N}\sum_{i=1}^{N} \psi_{it}^N \mathbf{f}(\boldsymbol{\alpha},i)$, which is then used to construct the GMM objective function:

$$\hat{\alpha}_N = \arg\min[\mathbf{G}_N(\boldsymbol{\alpha})^T\Phi_N^T\Phi_N\mathbf{G}_N(\boldsymbol{\alpha})].$$

Under certain regularity conditions outlined in Kyriazidou (2001), this estimator is consistent and asymptotically normal with $\sqrt{Nh_N}$ convergence rate. In the absence of a formula for the optimal weighting matrix, we will choose Φ_0^* such that $\Phi_0^{*}T\Phi_0^* = E((\psi_{nt}^N)^2\mathbf{f}(\boldsymbol{\alpha},n)^T\mathbf{f}(\boldsymbol{\alpha},n))^{-1}$, which is optimal in the standard GMM case.

Unlike in the standard GMM case, the J-test for the kernel-weighted estimator is noncentral *chi*-squared (because of the asymptotic bias of the estimator), with the noncentrality parameter (NCP) equal to the squared mean of $(1/(\sqrt{Nh_n}))\sum_{n=1}^{N}\hat{\psi}_N f_j(\alpha_j, n)$. Even though this quantity can be estimated in principle, this is very difficult in practice (see Bierens, 1987, for a detailed discussion). The Monte Carlo experiments in Kyriazidou (1997, 2001) suggest that this bias is small in general, suggesting that the NCP is also likely to be small. We use the central *chi*-squared distribution to perform the hypothesis tests. In the worst case, this will bias the results toward rejection if the NCP is large (see Davidson & MacKinnon, 1993, pp. 412–414).

Kernel Function: We assume that the K satisfies the standard regularity conditions. In particular, $\int |K(v)|dv < \infty$, $K(v)\,dv = 1$, and we consider symmetric kernels: $\int vK(v)\,dv = 0$. Moreover, the smoothness of the kernel affects the asymptotic convergence rate, which imposes restrictions on the empirical choice of the function $K(\cdot)$. We work with a Gaussian kernel which satisfies these conditions. Because asymptotically optimal kernel functions perform only slightly better even in the limit, normal density is a reasonable choice in practice.

The Choice of Optimal Bandwidth: The first step is to choose h_N. As is usually the case with semiparametric methods, asymptotically optimal methods for selecting the bandwidth provide little guidance for practical implementation with a fixed sample size. However, observing that the estimated weighting function, ψ_{it}^N, has a structure similar to a kernel density estimator, a sensible approach is to select h_N as the cross-validated value for the estimation of the density of the selection index, θy_{it}. Hence, the bandwidth is chosen by minimizing the mean integrated squared error of the kernel density estimator.

APPENDIX C

The Wage Equation

Provided that the labor market is competitive, the wage rate of a given household head can be written as follows:

$$W_{1it} = \delta_w(\mathbf{X}_{wit})\overline{W_t}, \tag{C1}$$

where $\delta_w(\mathbf{X}_{wit})$ is an efficiency index function; $\mathbf{X}_{wit}$ is a vector of household characteristics possibly containing some elements not included in $\mathbf{X}_{it}$; and $\overline{W_t}$ is the market wage rate. This wage equation is used in some specifications (described in the text) to increase the asymptotic efficiency of the estimator through the correlation of the error terms. We assume that $\mathbf{X}_{wit} = (\mathbf{x}_{wit}, \varepsilon_{wit})$, where $\mathbf{x}_{wit}$ and ε_{wit} denote the observable and unobservable components. Moreover, similarly to the specification in (7) we assume that $\delta_w(\mathbf{X}_{wit}) = \exp(b_w\mathbf{x}_{wit} + \varepsilon_{wit})$. After taking logs, first-differencing, and rearranging, we have

$$\Delta w_{1it} = \Delta\ln(\overline{W_t}) + \mathbf{b}_w\Delta\mathbf{x}_{wit} + \Delta\varepsilon_{wit}. \tag{C2}$$

The variable $\mathbf{x}_{wit}$ includes a constructed experience variable: $(A_{1nt} + E_{nt}^2)$, where E_{nt} is the completed education of the individual.