

Asset Pricing with Idiosyncratic Risk and Overlapping Generations*

Kjetil Storesletten,[†] Chris I. Telmer,[‡] and Amir Yaron[§]

May 2001
(first draft: June 1996)

Abstract

Constantinides and Duffie (1996) show that for idiosyncratic risk to matter for asset pricing the shocks must (i) be highly persistent and (ii) become more volatile during economic contractions. We show that data from the Panel Study on Income Dynamics (PSID) are consistent with these requirements. Our results are based on econometric methods which incorporate macroeconomic information going beyond the time horizon of the PSID, dating back to 1910. We go on to argue that life-cycle effects are fundamental for how idiosyncratic risk affects asset pricing. We use a stationary overlapping-generations model to show that life-cycle effects can either mitigate or accentuate the equity premium, the critical ingredient being whether agents accumulate or decumulate risky assets as they age. Our model predicts the latter and is able to account for both the average equity premium and the Sharpe ratio observed on the U.S. stock market.

*A predecessor of this paper circulated as “Persistent Idiosyncratic Shocks and Incomplete Markets.” In addition to numerous participants at seminars and conferences, we thank Dave Backus, Geert Bekaert, David Bowman, Martin Browning, John Cochrane, Wouter den Haan, Mick Devereux, Ron Gallant, René Garcia, Rick Green, John Heaton, Burton Hollifield, Mark Huggett, Narayana Kocherlakota, Per Krusell, Deborah Lucas, Pierre Mella-Barral, Bob Miller, Michael Parkin, Victor Ríos-Rull, Amlan Roy, Tony Smith, Paul Willen, Steve Zeldes, Stan Zin and three anonymous referees for helpful comments and suggestions. We have benefitted from the support of NSF grant SES-9987602 and the Rodney White Center at Wharton. Valuable research assistance was provided by Xiaorong Dong. The current version of this paper is available at <http://bertha.gsia.cmu.edu>.

[†]Institute for International Economic Studies, Stockholm University and CEPR; storeslk@iies.su.se

[‡]Graduate School of Industrial Administration, Carnegie Mellon University; chris.telmer@cmu.edu

[§]Wharton School, University of Pennsylvania and NBER; yarona@savage.wharton.upenn.edu

1 Introduction

The essence of Mehra and Prescott's (1985) equity premium puzzle is that investing in equity looks like a free lunch; an individual who consumes aggregate consumption places far greater value on the stock market than does the stock market itself. A large literature has asked if aggregation lies at the heart of the puzzle. The idea is that individuals face idiosyncratic risks, are unable to insure against them, and that this affects the way they value financial assets. The plausibility of this story seems apparent; non-financial wealth, human capital in particular, is subject to substantial risks and is larger in magnitude than financial wealth. Upon closer inspection, however, the story runs into difficulties. Idiosyncratic risks are, by definition, uncorrelated with aggregate risks. In contrast, asset pricing relies on dependence between sources of risk and asset returns in order to explain why some assets pay a higher average return than others. The challenge for a theory of asset pricing driven by idiosyncratic risk, therefore, is to generate such dependence while still having the idiosyncratic shocks wash-out at the aggregate level.

Constantinides and Duffie (1996), following work by Mankiw (1986), have overcome this difficulty by assuming that the volatility of idiosyncratic shocks covaries negatively with aggregate shocks. Our paper begins by asking if non-financial earnings data from the Panel Study on Income Dynamics (PSID) are consistent with this assumption. We find that they are, documenting new evidence of both countercyclical volatility as well as high persistence in idiosyncratic shocks. Our results contrast those of previous studies, in particular Heaton and Lucas (1996). The main reason is that our statistical approach explicitly incorporates household age, which allows us to capture an important interaction between age, aggregate shocks and idiosyncratic shocks.

The remainder of our paper argues that, in addition to countercyclical volatility, life-cycle effects are fundamental. The reasons are as follows. In our model, and most related work, idiosyncratic risk takes the form of shocks to nontradeable labor market income. The essence of the question, then, is how shocks to human capital affect the pricing of financial capital. The magnitude of human capital relative to financial capital is inherently a life-cycle phenomenon. Idiosyncratic risk, therefore, has an inescapable life-cycle component: the young face more than the old. The implications for asset pricing are important. The most obvious example is retirement. Retired people face little such risk, yet they comprise roughly 20 percent of the adult population and their participation rate in equity markets is even higher. If working individuals dislike stocks because of labor market shocks — and therefore demand a premium to hold them — why don't the retirees hold all the stocks, thus driving down the equity premium and resurrecting the puzzle? The same question applies, to a lesser degree, to older workers, for whom idiosyncratic shocks are relatively unimportant because human capital is small relative to financial capital.

We answer these questions using a class of life-cycle economies. We start with an OLG generalization of the no-trade, Constantinides-Duffie model and show that introducing retirement dramatically reduces the equity premium. We then motivate trade by introducing a realistic distribution of human wealth relative to total wealth; old agents have little of the former and young agents, who must save for retirement, have little of the latter. Given both trade and retirement, the equity premium depends primarily on the age-old question of whether to accumulate or decumulate risky assets as one ages. If agents accumulate, life-cycle effects dampen the equity premium and deepen the puzzle. If agents decumulate then workers are instrumental for pricing and, because the volatility of the shocks they face is countercyclical, they demand a high premium which tends to resolve the puzzle. Our calibration generates the latter.

The driving force in our model is a life-cycle interaction between idiosyncratic and aggregate risk which goes beyond the countercyclical-volatility interaction emphasized by Constantinides and Duffie (1996). Aggregate shocks in our model have a larger impact on financial wealth than on human wealth (*i.e.*, stock returns are more volatile than the wage rate). The total wealth of young agents is heavily weighted toward human capital, whereas the opposite holds for old agents. These two features, in isolation, suggest an intergenerational transfer of aggregate risk from the old to the young which will reduce the aggregate risk premium (*e.g.*, Ríos-Rull (1996)). What this ignores, however, is that human capital is subject to idiosyncratic risk. Young agents in our model are dissuaded from taking-on aggregate risk, primarily because of the Constantinides-Duffie, countercyclical-volatility effect. The net effect is that aggregate risk falls disproportionately on the backs of middle-aged agents, who demand a relatively high premium to hold it. Our main point, then, is that idiosyncratic risk affects the equity premium because it acts as a brake on the intergenerational sharing of aggregate risk. An implication is that stock-holding is concentrated on mature workers, the youngest workers hold zero stock, and retirees hold diversified portfolios of stocks and bonds. This is broadly consistent with portfolio composition in the United States (*e.g.* Amerkis and Zeldes (2000), Heaton and Lucas (2000)).

Quantitatively, our model is able to account for both the U.S. equity premium and the Sharpe ratio with a risk aversion coefficient of 6.7. In contrast, our calibration of the Constantinides and Duffie (1996) model requires a coefficient of 9.4 and is unable to simultaneously match both moments. An important caveat is the mechanism with which we generate variability in the return on equity. For computational reasons, we use a production economy with a linear technology and a stochastic depreciation rate. While the consumption side of the model is realistic, the production side is not (investment and output growth are excessively variable).

Our econometric approach is motivated by our emphasis on the life cycle. We interpret PSID data as arising from finite-lived processes. This leads to a method-of-moments estimator which conditions on age and, therefore, is able to capture a striking empirical regularity: cross-sectional dispersion in earnings increases with age.

This is primarily what leads to high estimates for the autocorrelation of idiosyncratic shocks, exceeding 0.90 in all cases. Just as importantly, age-dependence in cross-sectional moments, along with the fact that households born at different times have lived through different series' of macroeconomic shocks, allows us to incorporate much more information on aggregate shocks than have previous studies. Our estimates are based on aggregate variation over the period 1910-1992 whereas previous work has been confined to the time horizon of the panel, which in our case would be 1968-1992. We find the implications to be important, concluding that the conditional standard deviation of idiosyncratic shocks increases by roughly 120 percent as the economy moves from a boom to a recession.¹

An advantage of our model relates to risk-sharing behavior. U.S. data on income and consumption indicate that, while complete markets may not characterize the world, neither does a distinguishing feature of the Constantinides and Duffie (1996) framework: autarky. The cross-sectional standard deviation of U.S. consumption, for instance, is roughly 35 percent smaller than that of non-financial earnings.² The Constantinides-Duffie model, in contrast, features consumption and income volatility which are equal. Given the question being asked — How does the market price of risk depend on idiosyncratic risk? — this inconsistency seems of first-order importance. Surely the *amount* of idiosyncratic risk inherent in consumption allocations must matter? An advantage of our framework is that, even with unit root shocks, allocations feature partial risk sharing behavior. This is a characteristic of how the life-cycle and buffer-stock savings motives interact. Our model generates too little risk sharing (*i.e.*, consumption variability exceeds that of the data). Nevertheless, it suggests that life-cycle effects may help reconcile observed consumption behavior with the high degree of persistence in idiosyncratic shocks which observed asset prices seem to require.

A number of studies have examined the quantitative implications of the Constantinides and Duffie (1996) model.³ To understand our contribution, it is important

¹Previous papers which have used panel data to examine the time-series properties of labor earnings as they relate to risk sharing include Abowd and Card (1989), Altonji, Hayashi, and Kotlikoff (1991), Altug and Miller (1990), Attanasio and Davis (1996), Deaton (1991), Deaton and Paxson (1994), Hubbard, Skinner, and Zeldes (1994), MaCurdy (1982) and Mrkaic (1997). The finding of high persistence is a common theme among many of these papers. We are unaware of existing work which has used an age-dependent GMM methodology similar to ours.

²A large number of papers, including Altonji, Hayashi, and Kotlikoff (1991), Attanasio and Davis (1996), Attanasio and Weber (1992), Cochrane (1991) Deaton and Paxson (1994) Mace (1991) and Storesletten, Telmer, and Yaron (2000a) provide evidence which is suggestive of imperfect risk sharing. Altug and Miller (1990) find opposing evidence. The numerical value cited here is based upon evidence in both Deaton and Paxson (1994) and Storesletten, Telmer, and Yaron (2000a).

³Our paper is also related to a large body of work on asset pricing with heterogeneous agent models. Most closely related is Krusell and Smith (1997), which serves as an important benchmark, both in terms of computational methods and results. Our paper is primarily distinguished by its emphasis on the life cycle. In addition, they require extreme borrowing constraints (essentially zero) in order to generate risk premiums, whereas borrowing constraints play no role in our study. Other related papers include Aiyagari (1994), Aiyagari and Gertler (1991), Alvarez and Jermann (1999), den Haan (1994),

to understand their main result. They show that any given collection of asset price processes are consistent with a heterogeneous agent economy in which agents have ‘standard’ preferences and face idiosyncratic shocks with a particular volatility process. Their model’s testable restrictions can be thought of in two ways. First, because the economy admits the construction of a representative agent, it restricts the joint behavior of aggregate consumption, asset returns and the cross-sectional variation in consumption. That is, conditional on knowledge of the cross-sectional variance, the model’s first-order conditions can be tested without individual-level data. Papers by Balduzzi and Yao (2000), Brav, Constantinides, and Geczy (2000), Cogley (2000) Ramchand (1999) and Sarkissian (1999) investigate these restrictions and find mixed evidence. Second, if one asks what gives rise to the first-order conditions, the model restricts the joint behavior of individual labor income, asset returns, and individual consumption. Most critical is the requirement that labor income be a unit-root process with innovations which become more volatile during aggregate downturns. Our paper focuses on these restrictions.⁴ The advantages to doing so are both related to data — income is certainly easier to measure than consumption — and the ability to understand how idiosyncratic risk interacts with asset pricing at a structural level.

The remainder of our paper is organized as follows. Section 2 uses PSID data to evaluate the restrictions placed on labor earnings data by the Constantinides and Duffie (1996) model. Section 3 incorporates a notion of life-cycle into the Constantinides-Duffie no-trade economy and shows that the equity premium is dramatically diminished. Section 4 overturns this result in an economy with trade, motivated by life-cycle savings. Section 5 concludes.

2 Measuring Idiosyncratic Risk

The critical ingredients of both our theory and Constantinides and Duffie’s (1996) are the persistence of idiosyncratic shocks and the extent to which their volatility varies over the business cycle. The following statistical process captures these features. Denoting Y_t as the logarithm of per-capita labor earnings at time t and y_{it}^h as the logarithm of labor earnings for household i of age h at time t , we decompose y_{it}^h into two components:

$$y_{it}^h = g(x_{it}^h, Y_t) + u_{it}^h . \tag{1}$$

Heaton and Lucas (1996), Huggett (1993), Lucas (1994), Mankiw (1986), Marcet and Singleton (1999), Ríos-Rull (1994), Telmer (1993), Weil (1992), and Zhang (1997). The stationary OLG framework we develop owes much to previous work by Ríos-Rull (1994), Huggett (1996) and Storesletten (1999).

⁴Krusell and Smith (1997), emphasize the Constantinides-Duffie model’s unrealistic implications for financial wealth distribution. We focus, instead, on its implications for the distribution of financial wealth relative to human wealth. Given the inherent nature of the question — How do nontradeable shocks to human wealth affect the pricing of financial wealth? — this seems of primary importance.

The component $g(x_{it}^h, Y_t)$ is comprised of aggregate shocks as well as x_{it}^h , deterministic components of household-specific earnings attributable to age, education and so on. The component u_{it}^h is the random component of a household's earnings which is idiosyncratic to them. The condition which identifies u_{it}^h is,

$$\tilde{E}_t(u_{it}^h) = 0 \quad , \quad \forall t$$

where \tilde{E}_t denotes the cross-sectional mean, conditional on aggregate information at date t .

The specification of $g(x_{it}^h, Y_t)$, while critical for identifying u_{it}^h , is not central for our question. We therefore relegate its discussion to Appendix A, focusing here on the idiosyncratic process, u_{it}^h . We use an ARMA(1,1) specification with a regime-switching process for the conditional variance:

$$\begin{aligned} u_{it}^h &= z_{it}^h + \varepsilon_{it} \\ z_{it}^h &= \rho z_{i,t-1}^{h-1} + \eta_{it} \quad , \end{aligned} \tag{2}$$

where household age, h , is made explicit only when the conditional distribution of a variable depends upon it. We assume that $\varepsilon_{it} \sim \text{Niid}(0, \sigma_\varepsilon^2)$, $\eta_{it} \sim \text{Niid}(0, \sigma_t^2)$, $z_{it}^0 = 0$ and,

$$\begin{aligned} \sigma_t^2 &= \sigma_E^2 \text{ if aggregate expansion at date } t \\ &= \sigma_C^2 \text{ if aggregate contraction at date } t \quad . \end{aligned}$$

The idea underlying Constantinides and Duffie (1996) is that $\sigma_C > \sigma_E$, but we do not impose this *a priori*.

The initial conditions $z_{it}^0 = 0$, and the fact that age h is finite, allow us to interpret (2) as a finite process. This has a number of advantages. It allows us to condition on age, which, as will become apparent, is fundamental to our entire approach. It also implies that the distribution of u_{it}^h is well defined for any value of ρ , thereby avoiding the well-known difficulties associated with non-stationary time series analysis. This all comes at the cost of the strong assumptions on initial conditions. In Appendix A we discuss alternatives and argue that our results are robust.

The age dependence in equations (2) is what distinguishes both our approach and our results. Most important is how it affects the cross-sectional variances:

$$\text{Var}(u_{it}^h) = \sigma_\varepsilon^2 + \sum_{j=0}^{h-1} \rho^{2j} (I_{t-j} \sigma_E^2 + [1 - I_{t-j}] \sigma_C^2) \quad , \tag{3}$$

where $I_t = 1$ if the economy is in an expansion and $I_t = 0$ otherwise. If $|\rho| < 1$, the summation term converges to the familiar $\sigma^2 / (1 - \rho^2)$, where σ^2 is a probability weighted average of σ_E^2 and σ_C^2 . For finite h , however, $\text{Var}(u_{it}^h)$ is increasing in h , at a

rate determined by the value of ρ . This allows us to incorporate a striking feature of U.S. data on income and consumption: inequality increases with age.⁵ The relatively large estimates of ρ obtained below are primarily a consequence of incorporating this property.

Age dependence also plays an important role in measuring countercyclical volatility. It admits restrictions between age, cross-sectional variance and the business cycle which identify σ_E and σ_C . Consider, for instance, the cohort of 60 year old workers in the first year of our panel, 1968. According to equation (3), the cross-sectional variance for this cohort involves indicator variables I_{t-j} dating as far back as 1930 (we assume $h = 1$ corresponds to a 23 year old worker). Given that we can classify each year between 1930 and 1968 as either an expansion or a contraction, we can form moments based on (3) which incorporate aggregate shocks dating back to 1930, far beyond the 1968-1992 confine of our panel. Doing so across all 25 years of our panel is what identifies σ_E and σ_C . To see this, compare the 60 year old workers in 1968 to the 60 year old workers in 1992. Each have worked for 38 years, however the 1968 cohort have worked through more contractions, including the Great Depression. The process (2) predicts that, if $\rho > 0$ and $\sigma_C > \sigma_E$, then the cross-sectional variance of a particular cohort is increasing in the number of contractionary years they worked through. Therefore, inequality among the 1968 cohort should exceed that of the 1992 cohort. This is exactly what we document below, it is what distinguishes the GMM estimator in Section 2.3, and it is the main reason we reach substantially different conclusions than have previous studies.

2.1 PSID Data

Our microeconomic data is defined at the household level and is taken from the Panel Study on Income Dynamics (PSID), 1968-1991. Earnings are defined to equal the wage receipts of all adult household members, plus any transfers received such as unemployment insurance, workers compensation, transfers from non-household family members, and so on. Transfers are included because our model abstracts from the implicit insurance mechanisms which these payments often represent. That is, we wish to measure earnings variation *net* of risks which are insured via programs such as unemployment insurance. Along similar lines, we study the household as a single unit so as to abstract from shocks which are insured via intra-household variation in labor force participation.

We depart from the common practice of using a longitudinal panel: an equal number of time series observations on a fixed cross section of households. A longitudinal

⁵Deaton and Paxson (1994) demonstrate that this is a robust feature of U.S. data from the Consumer Expenditure Survey, as well as data from Great Britain and Taiwan. Storesletten, Telmer, and Yaron (2000a) show that the same holds for earnings data from the PSID.

panel necessarily features an average age which increases with time. This is problematic for our approach which emphasizes restrictions between age and aggregate variation. For instance, were we to use a longitudinal panel for the years 1968-1991, a large fraction of the household heads will have been retired, or at least been in their late earning years, during the last 2 of only 5 business cycles witnessed during the period 1968-1991.⁶ A longitudinal panel will also feature a relatively small sample size — only 610 households in our case — and be subject to survivorship bias in that only the relatively stable households are likely to remain in the panel for all 23 years.

We overcome these issues by constructing a sequence of overlapping three-year sub-panels. For each of the years 1968-1991, we construct a three year sub-panel consisting of households which reported strictly positive total household earnings (inclusive of transfers) for the given year and the next 2 consecutive years in the survey. For example, our 1970 sub-panel is essentially a longitudinal panel on 1,663 households over the years 1970, 1971 and 1972. Doing this for all years results in sequence of 22 overlapping sub-panels. The overall data structure contains more than enough time-series information to identify the parameters in equation (2), while at the same time survivorship bias is mitigated and the cross-sectional distribution of age is quite stable. The mean and standard deviation of the average age in each sub-panel are 44.2 and 1.1, respectively. The mean and the standard deviation (across panels) of the number of households is 2045 and 228, respectively. Additional details, including a number of filters related to measurement error and demographic stability, are outlined in Appendix A.

2.2 Graphical Analysis

The key to understanding the econometric results in the next section is understanding how cross-sectional variance in the PSID varies with age and with time. Variation across age is what drives our estimate of autocorrelation, ρ . Variation across time drives our estimates of σ_E and σ_C .

Panel A of Figure 1 graphs the cross-sectional variance of idiosyncratic earnings, $Var(u_{it}^h)$, against age. The graph shows that between ages 23 and 60, the variance increases by a factor of 3 and that, roughly speaking, the increase is linear. Inspection of equation (3) indicates that this implies a non-zero value of ρ , that the rate of increase depends on the values of σ_E and σ_C , and that the linear shape requires a value of ρ near unity. In other words, the slope of the graph dictates the magnitude of the conditional variance while the curvature dictates persistence. The relatively large estimates of ρ we obtain in Section 2.3 are essentially a formalization of this which incorporates additional information not represented in the graph, most notably autocovariances.

⁶Specifically, average age increases from 39 to 62 over the years 1968-1991

Panels B and C of Figure 1 consider variation over time instead of age. Panel B graphs the overall cross-sectional standard deviation for the years 1968-1991. It also graphs the detrended cross-sectional mean. Even at this informal level, where we pool together all age cohorts in a given year, we see striking evidence of countercyclical volatility. The correlation between the mean and the standard deviation is -0.85 . The magnitude of the changes, however, must be interpreted with caution. In Section 2.4 we show that, because it is a close cousin of the unconditional variance, the pooled cross-sectional variance in Figure 1-B, will always understate variation in what we are ultimately interested in: the conditional variance σ_t^2 .

Finally, Panel C of Figure 1 provides a visual characterization of how we break the time series confines of our panel and incorporate aggregate shocks dating back to 1910. The graph contains one data point for each cohort in our panel, a cohort being a group of households with heads born in the same year. Our approach is based upon one simple implication of the cross-sectional variances in equation (3); if cohort x has worked through more contractionary years than cohort y , then the cross-sectional variance among members of cohort x should exceed that of cohort y . Figure 1-C shows that this is a characteristic of our panel. The horizontal axis represents the fraction of years during a given cohort's working life which were contractions, the latter being defined as the NIPA measure of GDP growth being below average. The vertical axis is an estimate of the cross-sectional variance for this cohort. Although there are a number of cohorts who worked through relatively few contractions and nevertheless exhibit substantial differences in cross-sectional dispersion (*i.e.*, the cluster of points in the south-west corner of the graph), the positive relationship is apparent. The OLS slope coefficient is 0.93 with a standard error of 0.11. This relationship between age, cross-sectional variance, and macroeconomic history is the main reason that our conclusions regarding variation in σ_t^2 differ from those of previous studies.

2.3 Estimation

The moments in Figure 1 suggest a value of ρ near unity and values $\sigma_E < \sigma_C$. This, however, is based only on cross-sectional variances and ignores the most natural moments for estimating a time-series model, autocovariances. We now formulate a GMM estimator which incorporates autocovariances while at the same time formalizing the visual evidence above.

Table 1 reports GMM estimates of the parameters from the process (2), based on the following moment conditions:

$$\tilde{E}_t \left[(u_{it}^h)^2 - \sigma_\varepsilon^2 - \sum_{j=0}^{h-1} \rho^{2j} (I_{t-j} \sigma_E^2 + [1 - I_{t-j}] \sigma_C^2) \right] = 0 \quad (4)$$

$$\tilde{E}_t \left[u_{it}^h u_{it-1}^{h-1} - \rho \sum_{j=1}^{h-1} \rho^{2(j-1)} (I_{t-j} \sigma_E^2 + [1 - I_{t-j}] \sigma_C^2) \right] = 0 \quad (5)$$

$$\tilde{E}_t \left[u_{it}^h u_{it-2}^{h-2} - \rho^2 \sum_{j=2}^{h-1} \rho^{2(j-2)} (I_{t-j} \sigma_E^2 + [1 - I_{t-j}] \sigma_C^2) \right] = 0 \quad (6)$$

Equations 4 (one for each age, h) are cross-sectional variances, identical to (3). Equations (5) and (6) are first and second-order autocovariances. Complete details of our estimation procedure are provided in Appendix A.

The distinctive thing about the moments (4)-(6) is that they condition on a particular realization of aggregate shocks. That is, we require knowledge of the variables I_{t-j} — the indicator variables which denote expansion or contraction — in order to compute the cross-sectional mean. Table 1 uses three different measures, based on either GDP growth, the NBER business cycle series, or aggregate unemployment.

The estimated autocorrelation, ρ is between 0.91 and 0.93 in all cases. Relative to the implications of Figure 1-A, the incorporation of autocovariances therefore drives down the estimate of ρ . This is qualitatively consistent with previous work (*e.g.*, Heaton and Lucas (1996)) as well as our related paper, Storesletten, Telmer, and Yaron (2000a). The estimated conditional variances indicate that the persistent innovations are substantially more volatile than the transitory ones, with the standard deviation being roughly twice as large in all cases (based on the average of σ_E and σ_C). Finally, our estimate of countercyclical volatility is striking. In each case σ_C is more than twice the size of σ_E . The estimated conditional standard deviation increases by roughly 120 percent between expansion and contraction.⁷

Our findings relate to previous work as follows. Heaton and Lucas (1996) find much weaker evidence for countercyclical volatility, concluding that the analogous increase in the conditional standard deviation is roughly 27%. We attribute the differences to our emphasis on age-dependent moments, our panel’s stable age distribution (their panel is longitudinal), and our incorporation of many more aggregate shocks. In Appendix A we provide further comparisons, including a replication of their findings using a longitudinal panel.

Our estimates of autocorrelation, ρ , are substantially larger than those of Heaton and Lucas (1996) (who obtain 0.53), but are quite similar to those of a number of other papers, including Abowd and Card (1989), Hubbard, Skinner, and Zeldes (1994) and MaCurdy (1982). Again, as Section 2.2 emphasizes, the main reason is our emphasis on how cross-sectional variance increases with age, a fact which is confirmed by Deaton and Paxson (1994). In Storesletten, Telmer, and Yaron (2000a) we use an alternative estimator, which uses a less restrictive specification for aggregate effects than our model dictates here, and obtain estimates of ρ which are indistinguishable from unity.

Finally, our estimates of conditional variance are somewhat larger than several previous papers. Where we find a conditional standard deviation (assuming homoskedas-

⁷If we redefine ‘expansion’ and ‘contraction’ to be ‘bull market’ and ‘bear market,’ the analogous estimate is 80 percent.

tic innovations) for the persistent shocks of 0.25, Heaton and Lucas (1996) and Hubbard, Skinner, and Zeldes (1994) report 0.24 and 0.18, respectively. This is not surprising given that our overlapping panel admits a wider variety of households and, to a certain extent, mitigates survivorship bias. We also find the fraction of the innovation variance attributable to transitory shocks to be smaller than previous studies. Hubbard, Skinner, and Zeldes (1994), for example, estimate that the transitory and persistent variances are roughly of the same magnitude.

2.4 Temporal Variation in the Cross-Sectional Variance

Figure 1-B indicates that, over the period 1968-1991, the largest change in the cross-sectional standard deviation of earnings was roughly 13 percent. At first blush, this may seem inconsistent with our estimates, which indicate a change of greater than 100 percent over the course of the business cycle. The missing link is that the standard deviation in Figure 1-B is closely associated with the unconditional cross-sectional distribution of earnings, while the estimates in Table 1 are associated with the conditional distribution. In this section we ask whether the two can be quantitatively reconciled, thereby providing one last corroborative check on our estimates.

The overall cross-sectional variance in Figure 1-B is a weighted average of the conditional variances σ_E^2 and σ_C^2 . This is true in two senses. First, equation (3) indicates that, in the usual AR(1) sense, the cross-sectional variance among agents of a given age h is a moving average of the variances of past innovations. The only wrinkle is that the terms in the moving average change over time, depending on the macroeconomic history which the cohort has lived through. Second, the overall variance is a weighted average of the cohort-specific variances, where the weights are the population shares. Algebraically,

$$cv_t = \sum_{h=1}^H \varphi_h \left(\sigma_\varepsilon^2 + \sum_{j=0}^{h-1} \rho^{2j} (I_{t-j} \sigma_E^2 + [1 - I_{t-j}] \sigma_C^2) \right) , \quad (7)$$

where φ_h are population shares and cv_t denotes the pooled cross-sectional variance.

Inspection of equation (7) indicates that cv_t is a moving average which varies between an upper and a lower bound. Given that $\sigma_C > \sigma_E$, the upper bound coincides with a long sequence of aggregate contractions (*i.e.*, $I_{t-h} = 0$ for all $0 < h \leq H$), and is equal to $\sigma_\varepsilon^2 + \sigma_C^2 / (1 - \rho^2)$ for large H and $\rho < 1$. Similarly, the lower bound coincides with a long sequence of expansions and is equal to $\sigma_\varepsilon^2 + \sigma_E^2 / (1 - \rho^2)$. The way in which cv_t varies between these extremes depends on persistence in both aggregate and idiosyncratic shocks, and the economy's demographic structure (the φ_h 's). Figure 2 graphs this for one particular realization of our economy. It plainly illustrates the main point of this section, that large changes in the variance of the conditional distribution are necessarily associated with smaller changes in the overall cross-section.

We use this fact as another check on the plausibility of our estimates. We conduct a Monte Carlo experiment based on a large number of replications of our economy with a time period of 24 years per replication (which corresponds to the time dimension of our PSID panel). We define two test statistics. The first is based on the ratio of the maximum to the minimum values of cv_t which are observed in any given replication:

$$\frac{\max(cv_t)}{\min(cv_t)} . \quad (8)$$

The second is based on the maximal absolute change between any two adjacent years:

$$\max(|cv_t - cv_{t-1}|) \quad (9)$$

For each test statistic we compute the p-value, the probability of observing the sample statistics reported in Figure 1-B, given that our model generates the data. The p-values associated with the ratios (8) and (9) are 0.64 and 0.60, respectively.

To summarize, the estimated large changes in the conditional variance reported in Table 1 are quantitatively consistent with the relatively small changes we see in the overall cross-section, Figure 1-B. This serves to put the validity of our estimates on firmer ground, especially to the extent that variation in the overall cross-section is more easily and accurately measured.

3 An OLG Version of the Constantinides-Duffie Model

We now examine the asset pricing implications of the estimates from the previous section. We use a version of the Constantinides and Duffie (1996) model, modified to incorporate a life cycle component. Our formulation nests theirs as a special case. This allows us to examine their model directly as well as highlight the main point of our paper.

Due to the autarkic nature of the Constantinides-Duffie model, we are free to choose any specification for financial markets, without loss of generality, as long as trade in idiosyncratic outcomes is ruled out. We choose the simplest setup, a one-period riskless bond and an equity claim to a dividend process, D_t . The bond and equity prices are denoted q_t and p_t , respectively.

The economy is populated by H overlapping generations of agents, indexed by $h = 1, 2, \dots, H$, with a continuum of agents in each generation. Agents are born with one unit of equity and zero units of bonds. Preferences are

$$U(c) = E_t \sum_{h=1}^H \beta^h (c_{it+h}^h)^{1-\gamma} / (1-\gamma) , \quad (10)$$

where c_{it}^h is the consumption of the i^{th} agent of age h at time t and β and γ denote the discount factor and risk aversion coefficients, respectively.

Each agent receives nontradeable endowment income of y_{it}^h ,

$$y_{it}^h = G_t \exp(z_{it}^h) - D_t, \quad h = 1, 2, \dots, (H-1) \quad (11)$$

$$y_{it}^H = G_t \exp(z_{it}^H) - (p_t + D_t), \quad (12)$$

where G_t is an aggregate shock (defined more explicitly below) and the idiosyncratic shocks, z_{it}^h , follow a unit root process with heteroskedastic innovations,

$$z_{it}^h = z_{i,t-1}^h + \eta_{it} \quad (13)$$

$$\eta_{it} \sim N(-\sigma_t^2/2, \sigma_t^2) \quad (14)$$

$$\sigma_t^2 = a + b \log(G_t/G_{t-1}) \quad (15)$$

$$z_{i,t}^0 = 0. \quad (16)$$

This structure is essentially identical to the Constantinides-Duffie model, the only exception being that in the last period of life the amount $p_t + D_t$ is subtracted from income, instead of just D_t . Equilibrium is autarkic with individual consumption $c_{it}^h = G_t \exp(z_{it}^h)$. Bond and equity prices satisfy

$$q_t = \beta^* E_t \lambda_{t+1}^{-\gamma^*} \quad (17)$$

$$p_t = \beta^* E_t \lambda_{t+1}^{-\gamma^*} (p_{t+1} + D_{t+1}), \quad (18)$$

where $\lambda_{t+1} = G_{t+1}/G_t$, $\beta^* = \beta \exp(\gamma(1+\gamma)a/2)$ and $\gamma^* = \gamma - b\gamma(1+\gamma)/2$ (see Constantinides and Duffie (1996) for derivations). A cross-sectional law of large numbers implies that the variable G_t , and therefore the growth rate λ_t , coincides with per-capita consumption, which we denote C_t (the reason for making a potential distinction will become apparent in the next section),

$$C_t = \frac{1}{H} \tilde{E}_t \sum_{h=1}^H G_t \exp(z_{it}^h) = G_t,$$

where \tilde{E}_t is a cross-sectional expectations operator which conditions on time t aggregate information. Since $C_t = G_t$, the pricing equations (17) and (18) represent a representative agent equilibrium where the agent's preference parameters (β^*, γ^*) are amalgamations of actual preference parameters (β, γ) and technological parameters (a, b) . The main idea behind the Constantinides-Duffie model is that (i) because $\beta^* > \beta$, the model may resolve the 'risk-free rate puzzle,' and (ii) if $b < 0$ (*i.e.*, the volatility of idiosyncratic shocks is countercyclical) then 'effective' risk aversion exceeds actual risk aversion ($\gamma^* > \gamma$), and the model may resolve the equity premium puzzle.

We examine this quantitatively by conducting a calibration exercise similar to Mehra and Prescott (1985), but where the representative agent's preference parameters

are re-interpreted according to the formulae for β^* and γ^* . To do so, we map our PSID estimates from Section 2 into numerical values for the idiosyncratic risk parameters, a and b . Specific details are given in Appendix B. The results are $a = 0.038$ and $b = -0.516$. Next, we choose the ‘effective’ discount factor, β^* , to match the average U.S. riskfree interest rate, and the effective risk aversion coefficient, γ^* , to match either the U.S. Sharpe ratio or the unlevered U.S. equity premium. Table 2 reports the implications for the ‘actual’ risk aversion coefficient, γ . To match the Sharpe ratio, a value of $\gamma^* = 34.5$ is required. This corresponds to an actual risk aversion coefficient of $\gamma = 9.4$. To match the equity premium $\gamma^* = 13.9$ is required, which corresponds to $\gamma = 5.3$. Time preference is characterized by $\beta^* = 1.01$ ($\beta = 0.16$) and $\beta^* = 1.12$ ($\beta = 0.61$), respectively.

The Constantinides-Duffie model, then, is successful at what it sets out to do; given a realistic parameterization for idiosyncratic risk, it accounts for the equity premium without resorting to extreme values for risk aversion and/or negative time preference. Along other dimensions, of course, the model is counterfactual. It generates excessive volatility in both risky and riskless asset returns and cannot account for the ubiquitous rejections of Euler equation tests based on (17) and (18) (*i.e.*, such tests typically reject for all values of β^* and γ^*). Constantinides and Duffie (1996) prove that this can be rectified with an alternative process for the conditional variance σ_t^2 from equation (15). The remainder of our paper, however, focuses on a more fundamental set of the model’s restrictions, those which involve age and risk sharing.

3.1 The Implications of Retirement

The Constantinides-Duffie model has the feature that all agents face idiosyncratic income risk. This seems an important restriction for assessing the asset-pricing effects of labor market shocks. Retired people comprise roughly 20 percent of the adult population, they are disproportionately important in terms of equity market participation, yet they face little if any such shocks. We now add retirees to the model and ask to what extent its success is mitigated.

There are two senses in which the process (11)–(16) does not capture retirement. First, it confronts agents with idiosyncratic income shocks in all periods of life. Second, it implies that agents receive income each period until death, thus obviating the need to save for retirement. We begin by incorporating the first feature, which can be analyzed in the no-trade environment. The second requires trade and is incorporated in Section 4.

We define a retired agent as one who does not receive an idiosyncratic shock beyond some retirement age so that, for retirees, $a = b = 0$. Given this, equations (17) and (18) no longer describe autarkic equilibrium prices. Marginal rates of substitution (at

autarky) are

$$\text{workers: } \quad \beta E_t \left(\frac{G_{t+1}}{G_t} \right)^{-\gamma} e^{\gamma(1+\gamma)a/2} \left(\frac{G_{t+1}}{G_t} \right)^{\gamma b(1+\gamma)/2} , \quad (19)$$

$$\text{retirees: } \quad \beta E_t \left(\frac{G_{t+1}}{G_t} \right)^{-\gamma} . \quad (20)$$

Retirees differ from workers in two ways. First, with $a > 0$ the exponential term in equation (19) is positive, implying that retirees discount future consumption more than workers. Intuitively, the absence of idiosyncratic risk reduces their demand for precautionary savings and they assign a lower price to a riskfree bond. Second, if $b < 0$, retirees appear less risk averse than workers, assigning a relatively high value to risky assets or, equivalently, demanding a relatively small risk premium. By removing the countercyclical volatility from the retiree's endowments we have effectively given them a greater capacity to bear aggregate risk.

We can now do one of two things to characterize an equilibrium. We can allow trade and solve for market clearing prices to replace equations (17) and (18). This would involve a substitution of consumption from retirement toward the working years, and an increased exposure to aggregate risk for retired individuals. Alternatively, we can follow Constantinides and Duffie (1996) and characterize endowments which give rise to a no-trade equilibrium, but subject to the constraint that retirees do not receive idiosyncratic shocks. The difference between these endowments and those in equations (11) and (12) will be suggestive of what will characterize an equilibrium with trade.

A three-generation example, $H = 3$, will make the point. Generations 1 and 2 receive endowments according to equations (11)–(16). Generation 3 — the old agents — receive

$$y_{it}^3 = f_t G_t \exp(z_{it}^3) - (p_t + D_t) , \quad (21)$$

but with $z_{it}^3 = z_{it}^2$ (*i.e.*, the innovation in equation (13) equals zero), and

$$f_t = e^{-a(1+\gamma)/2} \left(\frac{G_t}{G_{t-1}} \right)^{-b(1+\gamma)/2} .$$

Given the endowment (21), the prices (17) and (18) once again support an autarkic equilibrium. Relative to the original endowment, the old now receive less goods (on average) with more aggregate risk, just as the above intuition suggests. What has changed, however, is aggregate consumption. Assigning a population weight of 20 percent to the old generation (corresponding to the U.S. population), aggregate consumption is

$$\begin{aligned} C_t &= \tilde{E}_t \left(0.8[G_t \exp(z_{it}^1) + G_t \exp(z_{it}^2)] + 0.2f_t G_t \exp(z_{it}^3) \right) \\ &= G_t \left(0.8 + 0.2 e^{-a(1+\gamma)/2} \left(\frac{G_t}{G_{t-1}} \right)^{-b(1+\gamma)/2} \right) , \end{aligned} \quad (22)$$

which, because we've added aggregate risk to the endowment of the old, can be substantially more variable than G_t .

The prices (17) and (18) are now valid, but only in an economy with more variability in aggregate consumption growth than the original. The above calibration (which underlies Table 2) is therefore invalid. Aggregate consumption growth, as implied by equation (22), now has a standard deviation of 9.6 percent, roughly three times larger than U.S. data. In this sense, adding retirees implies that, without changing preferences, the model can only account for asset prices with an unrealistically high amount of aggregate variability.

An alternative is to re-calibrate the process G_t/G_{t-1} so that aggregate consumption growth, C_t/C_{t-1} from equation (22), has mean, standard deviation and autocorrelation which match the U.S. data. Results are given in the 5th and 6th rows of Table 2. Holding risk aversion fixed, we find that the required reduction in the variability of aggregate consumption growth causes the model's Sharpe ratio to fall from 41.2 percent to 19.7 percent. The equity premium falls from 13.1 percent to 2.9 percent. For the alternative calibration (row 6), the Sharpe ratio and equity premium fall from 23.7 percent to 13.5 percent and 4.1 percent to 1.3 percent, respectively. Finally, if we instead re-calculate the risk aversion coefficient required to match, respectively, the Sharpe ratio and equity premium, we arrive at values of 19.8 and 11.2. Without retirees we required values of 9.4 and 5.3.

To summarize, retirement, defined here as old agents receiving fixed incomes, has the effect one might expect. Because retirees do not face countercyclically heteroskedastic shocks — the driving force in the Constantinides-Duffie model — they are less averse to bearing aggregate risk. An autarkic allocation must therefore skew the aggregate risk toward the old, who are content to hold it in return for a relatively low expected return. In this sense, the incorporation of retirement resurrects the equity premium puzzle.

4 Models With Trade

The previous section emphasized the importance of how idiosyncratic shocks are distributed over the life cycle. Equally important is the distribution of what is being shocked: the human wealth represented by the flow of income, y_{it}^h . Human wealth typically accounts for a large fraction of total wealth for young people and a small fraction for older people. Given the nature of our question — How do shocks to human wealth affect the valuation of financial wealth? — incorporating this seems of first-order importance. It may also overturn the implication of the previous section, which was driven by older agents bearing the lion's share of the aggregate risk. If a realistic human/financial wealth distribution reverses this, making the younger agents

who face the idiosyncratic risk instrumental in pricing the aggregate risk, the incorporation of retirement may actually help the model to account for the equity premium.

The major cost of incorporating a life-cycle wealth distribution is that, necessarily, we must allow for trade (*e.g.*, if nontradeable income is zero after retirement, the young must save and the old must dissave). With several exceptions, Gertler (1999) for example, this means using computational methods to analyze the model. The benefits, however, are (i) we can make the model more realistic along certain dimensions which are important for calibration (*e.g.*, the demographic structure) and (ii) the model will display partial risk-sharing behavior — an undeniable aspect of U.S. data on income and consumption — even with unit root idiosyncratic shocks. With this in mind, we make the following changes to the framework of Section 3.

Financial markets.

With trade, the menu of assets is no longer innocuous. We now *limit* asset trade to a riskless and a risky asset. The latter takes the form of ownership of an aggregate production technology. The main reason for adding production is computational tractability: the resulting price of the risky asset will always be equal to unity. Agents rent capital and labor to a single firm which then splits its output between the two. Labor is supplied inelastically and, in aggregate, is fixed at N . Denoting aggregate consumption, output and capital as Y_t , C_t and K_t respectively, the production technology is

$$Y_t = r_t K_t + w_t N \quad (23)$$

$$K_{t+1} = Y_t - C_t + (1 - \delta_t) K_t \quad (24)$$

$$r_t = \theta Z_t K_t^{1-\theta} N^{1-\theta} - \delta_t \quad (25)$$

$$w_t = Z_t w, \quad (26)$$

where r_t is the return on capital (the risky asset), w_t is the wage rate, θ is capital's share of output, Z_t is an aggregate shock, w controls the average wage rate, and δ_t is the depreciation rate on capital. The depreciation rate is stochastic:

$$\delta_t = \delta + (1 - Z_t) \frac{s}{Std(Z_t)}, \quad (27)$$

where δ controls the average and s is, approximately, the standard deviation of r_t .⁸

This production process delivers four key ingredients: (i) the model is tractable (solving the analogous endowment economy is substantially more difficult), (ii) the volatility of the return on equity can be calibrated realistically, (iii) the volatility of

⁸Greenwood, Hercowitz, and Huffman (1988) and Greenwood, Hurcowitz, and Krusell (1997) have used a similar production technology in a business cycle context. Boldrin, Christiano, and Fisher (2001) have done so in an asset pricing context. Our technology is essentially a reduced-form representation of, for instance, Greenwood, Hurcowitz, and Krusell (1997), equation (B3).

aggregate consumption growth can be calibrated realistically, (iv) the return on human capital — essentially the wage rate — can be substantially less volatile than the return on equity. Each ingredient is critical for our question. The first two are obvious. The third ensures that the aggregate part of the asset-pricing Euler equations is realistic (*i.e.*, see equations (17) and (18)), which is essential if we are to isolate the incremental impact of idiosyncratic risk. The fourth is instrumental in determining which age cohorts hold equity in equilibrium and, consequently, whether or not idiosyncratic risk is priced. Unfortunately, these ingredients come at a cost. They imply, for example, excessively volatile investment and output, a feature shared by most existing production-based models should they be calibrated to have realistic variability in asset returns. We do not resolve such issues here. We view our model in the same way we view an endowment economy; as an economy with a potentially unrealistic production side which, nevertheless, yields informative restrictions on consumption and asset returns.

Endowments.

The endowments (11)–(16) are of a special form required to support an autarkic outcome. Since this is no longer required, and because of the incorporation of production, we reformulate them as follows. First, to capture the fact that young people have relatively little financial wealth relative to human wealth, we endow all newborn agents with zero units of equity and zero units of bonds. Next, the nontradeable endowment now takes the form of labor efficiency units, not units of the single good.⁹ At time t the i th working agent of age h is endowed with n_{it}^h units of labor which they supply inelastically. Retirees, agents for whom h exceeds a retirement age \hat{H} , receive $n_{it}^h = 0$. For workers,

$$\log n_{it}^h = \kappa_h + z_{i,t}^h \quad , \quad (28)$$

where κ_h is used to characterize the cross-sectional distribution of mean income across ages, and

$$z_{it}^h = \rho z_{i,t-1}^{h-1} + \eta_{it} \quad , \quad \eta_{it} \sim N(0, \sigma_t^2) \quad ,$$

with $z_{it}^0 = 0$. Our specification for the conditional variance σ_t^2 reverts back to our original two-state specification (this is also necessary for computational reasons):

$$\begin{aligned} \sigma_t^2 &= \sigma_E^2 \text{ if } Z \geq E(Z) \\ &= \sigma_C^2 \text{ if } Z < E(Z) \quad . \end{aligned}$$

⁹Strictly speaking, this is inconsistent with our empirical approach which measured idiosyncratic risk using labor *income*, not hours worked. To reconcile the two, we have generated simulated data on labor income from our model and used the estimators described in Section 2 to estimate the parameters of the process (28). Owing in large part to relatively low variability in the wage rate, w_t , the results were very similar to those reported in Table 1. In this sense, the population moments for labor income in our model have been calibrated to sample moments on non-financial income from the PSID.

Individual labor income now becomes the product of labor supplied and the wage rate: $y_{it}^h = w_t n_{it}^h$.¹⁰

With $\rho = 1$ this process is analogous to the Constantinides-Duffie process, (11)–(16). The exceptions are that (i) income is now a share of the aggregate wage bill instead of aggregate consumption, (ii) financial income is no longer ‘taxed’ at 100 percent as in (11)–(16), and (iii) the variance of the innovations to z_{it}^h is now a discrete function of the technological shock Z , not a continuous function of aggregate consumption growth.

4.1 Equilibrium

The state of the economy is a pair, (Z, μ) , where μ is a measure defined over an appropriate family of subsets of $S = (\mathcal{H} \times \mathcal{Z} \times \mathcal{A})$, \mathcal{H} is the set of ages, $\mathcal{H} = \{1, 2, \dots, H\}$, \mathcal{Z} is the product space of all possible idiosyncratic shocks, and \mathcal{A} is the set of all possible beginning-of-period wealth realizations. In words, μ is simply a distribution of agents across ages, idiosyncratic shocks and wealth. The existence of aggregate shocks implies that, necessarily, μ must evolve stochastically over time (*i.e.*, μ belongs to some family of distributions over which there is defined yet another probability measure). We use G to denote the law of motion of μ ,

$$\mu' = G(\mu, Z, Z') .$$

The bond price and the return on equity can now be written as time-invariant functions $q(\mu, Z)$ and $r(\mu, Z)$. The wage rate is $w(\mu, Z)$. Omitting the (now redundant) time t and individual i notation, the budget constraint for an agent of age h is,

$$\begin{aligned} c_h + k'_{h+1} + b'_{h+1}q(\mu, Z) &\leq a_h + n_h w(\mu, Z) & (29) \\ a_h &= k_h r(\mu, Z) + b_h \\ k'_{h+1} &\geq 0 \\ b'_{H+1} &\geq 0 \end{aligned}$$

where a_h denotes beginning-of-period wealth, k_h and b_h are beginning-of-period capital and bond holdings, and k'_{h+1} and b'_{h+1} are end-of-period holdings. The third equation rules out shortselling (which turns out to be innocuous) and the fourth restricts terminal wealth to be non-zero. Note that, beyond terminal wealth, we do not impose borrowing constraints. Section 4.5 examines the effects of more restrictive assumptions.

Denoting the value function of an agent of age h as V_h , the choice problem can be represented as,

$$V_h(\mu, Z, z_h, a_h) = \max_{k'_{h+1}, b'_{h+1}} \left\{ u(c_h) + \right.$$

¹⁰Our model assumes that bequests are zero. This provides focus on our main point: the effect of intergenerational dispersion in the ratio of human to total wealth.

$$\beta E [V'_{h+1}(G(\mu, Z, Z'), Z', z'_{h+1}, k'_{h+1}r(G(\mu, Z, Z'), Z') + b'_{h+1})] \} \quad , \quad (30)$$

subject to equations (29). An equilibrium is defined as stationary price functions, $q(\mu, Z)$, $r(\mu, Z)$ and $w(\mu, Z)$, a set of cohort-specific value functions and decision rules, $\{V_h, k'_{h+1}, b'_{h+1}\}_{h=1}^H$, and a law of motion for μ , $\mu' = G(\mu, Z, Z')$, such that r and w satisfy equations (25) and (26), the bond market clears,

$$\int_S b'(\mu, Z, z_h, a_h) d\mu = 0 \quad ,$$

aggregate quantities result from individual decisions,

$$\begin{aligned} K(\mu, Z) &= \int_S k_h(\mu, Z, z_h, a_h) d\mu \\ N &= \int_S n_h d\mu \quad , \end{aligned}$$

agents' optimization problems are satisfied given the law of motion for (μ, Z) (so that $\{V_h, k'_{h+1}, b'_{h+1}\}_{h=1}^H$ satisfy problem (30)), and the law of motion, G , is consistent with individual behavior. We characterize this equilibrium and solve the model using the computational methods developed by Krusell and Smith (1997) and described further in Appendix C.

4.2 Quantitative Properties

Our model now has three main motives for trade: the life-cycle distribution of idiosyncratic shocks, the life cycle distribution of the ratio of human to total wealth, and the possibility that $\rho < 1$. In order to focus on life-cycle issues and maintain a tangible link with the Constantinides-Duffie benchmark, we put the latter aside until Section 4.4.

We calibrate the above economy according to the criteria outlined in Appendix B. The most important features are as follows.

1. The standard deviation of aggregate consumption growth matches the annual U.S. sample value of 3.7 percent. As equations (17) and (18) emphasize, realistic properties for aggregate consumption are essential here, just as they are in representative agent models. The cost, in our case, is excessively volatile output and investment, something which is commonplace in models with production. Full details are provided in Appendix B.
2. The standard deviation of the return on capital matches the annual sample moment of 10 percent from the unlevered CRSP value-weighted index. This is achieved in a reduced-form manner, via the stochastic depreciation process (27). We therefore have little to say about why the return on the equity market is as

variable as it is. What we can say, however, is that the main consequence — the return on financial capital being substantially more volatile than the return on human capital — has stark implications for life-cycle portfolio choice and, therefore, for how idiosyncratic shocks interact with asset pricing.¹¹

3. Idiosyncratic risk, captured by equation (28), follows a unit-root process with a regime-switching conditional variance function chosen to match our estimates Table 1. We scale down the variances in our model so that, with $\rho = 1$, the unconditional variance over the life-cycle matches that implied by the $\rho < 1$ estimates in the Table. This results in $\sigma_E = 0.0977$ and $\sigma_C = 0.2161$.
4. Young agents are born with zero assets and retired agents receive zero labor income. This serves as the primary motive for trade. It also results in a realistic life-cycle distribution of human to financial capital — younger agents hold most of the former whereas older agents hold most of the latter — which, as we'll see, plays an important role in portfolio choice.
5. Retired agents comprise roughly 20 percent of the population.

Table 2 reports the first two moments of the asset returns in this economy. In order to match the observed U.S. Sharpe ratio and mean equity premium, our model requires a risk aversion coefficient of 6.7.¹² In contrast, the Constantinides-Duffie model with retirement, where retirement simply means not facing idiosyncratic risk (Section 3.1), requires a coefficient of 19.8. The incorporation of life-cycle savings, therefore, more than offsets the negative effect of retirees not receiving shocks. More generally, our model also requires lower risk aversion than the Constantinides-Duffie model without retirement: a coefficient of 6.7 versus one of 9.4. The overall effect of the life cycle, then, is one of magnifying the effect of idiosyncratic risk on asset pricing.

To understand this, consider the portfolio behavior depicted in Figure 3. The figure reports the fraction of financial wealth invested in bonds for the mean individual of each age. We see that the youngest agents choose to hold only bonds. Agents between age 29 and 43 and those older than 63 tend to hold diversified portfolios of bonds and equity. Finally, mature workers — those between age 43 and 63 — issue bonds, thus maintaining a levered position in equity. There are two main forces at work:

- (i) The youngest workers face the most idiosyncratic risk and, therefore, are relatively ill-equipped to bear aggregate risk. More specifically, the young hold

¹¹A potentially important counterfactual aspect of our economy is that the return on capital is far more correlated with the wage rate — and therefore with labor earnings — than we see in the data. This can be overcome, at a non-trivial computational cost, by making the depreciation process (27) a stochastic function of the technology shock Z .

¹²Interestingly, this value for risk aversion is similar to what Cogley (2000) finds using CEX data in a model with time-varying cross-sectional moments. He estimates risk aversion of 8 with 6.7 lying well within a 90% confidence interval.

almost all of their wealth as human wealth, which is directly subject to idiosyncratic shocks. These shocks are large and they persist over much of a young agent's life. Moreover, their volatility covaries negatively with the risky return, something which discourages risky asset holding for any working agent. Because this is particularly true for the youngest workers, in equilibrium they choose to hold no risky assets whatsoever (recall that shortselling is prohibited).¹³

- (ii) Retired agents are also relatively ill-equipped to bear aggregate risk, but for a different reason. Retirees must finance consumption entirely out of financial wealth. Aggregate variation in the return on equity is large relative to that of human wealth.¹⁴ A retiree holding all of their wealth as equity, therefore, would face substantially more aggregate risk than would a worker holding all of their *financial* wealth as equity, because the worker also owns some human wealth. The result, as we see in Figure 3, is that retirees choose to diversify and allocate some of their wealth toward bonds. These bonds are issued by mature workers.¹⁵

Young workers, then, dislike the risky asset because of the countercyclical nature of the idiosyncratic risk they face. Retirees dislike it because it has a highly variable return and they no longer receive labor income. What Figure 3 shows is that middle-aged workers represent a bridge between the two; they dislike aggregate risk for the same reasons, but in each case to a lesser degree. They hold part of their wealth as financial wealth and therefore care less than the young do about idiosyncratic shocks. They face the same variability in equity returns as the old, but their labor income mitigates their overall exposure to aggregate shocks. The end-result is that the middle-aged hold levered equity by issuing bonds to the young and the old. The resulting hump-shaped pattern in equity ownership is broadly consistent with U.S. data and has been the focus of a number of recent studies (Amerkis and Zeldes (2000), Heaton and Lucas (2000))¹⁶.

¹³This interpretation is confirmed by experiments in which the volatility of idiosyncratic shocks is constant. In this case, the bond portfolio share profile is monotone increasing, as in Jagannathan and Kocherlakota (1996).

¹⁴To see this, note that our production technology implies that the correlation of the wage rate and the return on capital is essentially unity, with the latter being substantially more variable than the former. Given this, and assuming, for simplicity, that hours worked are an *i.i.d.* process, it is easily shown that the return on human capital is a convex combination of the return on the bond and the return on capital, where the weight on the capital return is $H^{-1} \times \text{Stdev}(W)/\text{Stdev}(R)$, with H denoting the value of human capital, W denoting the wage rate and R denoting the return on capital. Consequently, the variability of the return on human capital will be small insofar as (i) the variability of W is small relative to that of R and (ii) the size of H is large. The former is a feature of our economy, whereas the latter is increasingly the case, the younger is an agent.

¹⁵This interpretation is confirmed by experiments in which the return on equity is calibrated to be *less* than that of the wage rate. In these cases, we observe retired agents holding levered positions in the risky asset and an average bond share which is monotone increasing with age.

¹⁶In terms of individual portfolio choice, our results are driven by bond-holding decision rules being U-shaped as a function of wealth (Amerkis and Zeldes (2000) discuss a similar phenomenon). The

This outcome has a natural interpretation in terms of the intergenerational redistribution of aggregate risk. Because the return on capital is more variable than the aggregate component of the return on labor, an efficient allocation will feature aggregate risk being transferred from retirees — those who receive none of labor’s share of output — to workers. One mechanism which achieves this is retirees holding bonds which are issued by workers. (*e.g.*, Ríos-Rull (1996)). Young workers, however, face an additional source of aggregate risk, manifested in the conditional variance of their idiosyncratic risk process.¹⁷ An efficient allocation, therefore, will stop short of transferring aggregate risk to the youngest, resulting instead in the U-shaped pattern we see in Figure 3.

Finally, understanding asset pricing in our economy is straightforward. The essence of Section 3.1 was that if retired agents hold most of the aggregate risk, then, because they face no idiosyncratic risk, the implications of countercyclical volatility may be strongly mitigated and the equity premium puzzle may be resurrected. What we find here is that the life-cycle savings effect — most importantly, the fact that older agents derive most of their income from financial assets — counteracts this, essentially because of the life-cycle distribution of *aggregate* risk. Aggregate risk motivates retirees to hold a diversified portfolio and, because the young choose a corner solution, the lion’s share of the aggregate risk falls on the backs of the middle-aged. These agents are workers and the countercyclical variance in their idiosyncratic shocks matters for risk premia, just as it does in Constantinides and Duffie (1996). How much it matters is a quantitative question, depending mainly on middle-aged agents’ ratio of human to total wealth as well as their relative size in the population distribution. The ability to answer this question, as we do in Table 2, is the main benefit of undertaking our computational exercise.

4.3 Risk Sharing

An counterfactual implication of the Constantinides and Duffie (1996) model is that the equilibrium features no risk sharing while the bulk of the existing evidence suggests that partial risk sharing better characterizes the world. This seems important for the question, which essentially asks how idiosyncratic consumption risk affects the market price of risk. Surely the *magnitude* of the consumption risk which agents face — a synonym for the degree of partial risk sharing — is relevant for this question?

An advantage of the life-cycle model is that, even with unit root shocks, allocations exhibit partial risk sharing. The reason involves the way in which the life-cycle savings

average profile in Figure 3 inherits this shape because financial capital’s share of total wealth is, on average, an increasing function of age.

¹⁷In Storesletten, Telmer, and Yaron (2000b) we examine the welfare consequences of this form of aggregate risk more explicitly. We find that the welfare costs of business cycles can be quite large, should the elimination of business cycles also imply the elimination of heteroskedasticity in idiosyncratic shocks.

interacts with ‘buffer-stock savings:’ the savings reaction to an unexpected shock. In our model, provided that financial wealth is positive, the marginal propensity to save out of current income is increasing in the level of current income but decreasing in the level of wealth. The implication is that, in spite of being characterized by unit-root shocks, our economy displays a type of contingent, self-insurance behavior.

The risk-sharing behavior of our model is depicted in Figure 4. Panel A shows that, with the exception of the youngest, the cross-sectional variance in consumption is less than that of income. Averaged over age, consumption is roughly 10 percent less variable (in terms of the standard deviation). In U.S. data, this value is roughly 35 percent (see Deaton and Paxson (1994) or Storesletten, Telmer, and Yaron (2000a)), so our model exhibits too little risk sharing. Panel B reports the cross-sectional variance in the *growth rate* of consumption, which is more directly related to the essence of risk sharing: the equalization of marginal rates of substitution. In this case, we see a larger difference between our model and the autarkic outcome. Autarky implies that, for workers, the graph is flat at 0.17, as shown. Our model features a monotonically decreasing graph, starting at roughly autarky and falling to near zero. The main reason is what we’ve emphasized above: a decreasing ratio of human to total capital and the resulting mitigation of the impact of idiosyncratic shocks. Risk sharing behavior is yet another dimension of our model for which this ratio is the main economic force at work.

4.4 Mean-Reverting Idiosyncratic Shocks

We assumed $\rho = 1$, mainly to provide a coherent link to the Constantinides-Duffie model and isolate the effects of retirement/life-cycle savings. Our estimate in Table 1, however, is $\rho = 0.92$. We are able to solve economies with mean-reverting shocks, although the computational burden increases substantially. We find that the U-shaped pattern in life-cycle portfolio holdings (see Figure 3) is robust to $\rho = 0.92$. The main economic message of our paper, therefore, does not rely on unit-root shocks. Quantitatively, Table 2 shows that, holding risk aversion fixed, the equity premium and Sharpe ratio increase slightly relative to the unit-root economy. The riskfree rate, however, is -0.50 percent and the variability of aggregate consumption growth is 6.7 percent, roughly double that of our main model. Were we to re-calibrate the model to reconcile these values with data, we’d expect a reduced risk premia. In this sense, reductions in persistence mitigate the asset-pricing effects of idiosyncratic risk, just as previous work suggests. In the previous version of this paper, for instance, we show that values of ρ in the neighborhood of 0.5 reduce Sharpe ratios by an order of magnitude. High persistence, therefore, is an important part of the story regardless of the cyclical pattern in volatility which the literature has emphasized.

4.5 Borrowing Constraints

Unlike many related studies (*e.g.*, Heaton and Lucas (1996), Krusell and Smith (1997), Telmer and Zin (1995)), our model has placed no constraints on borrowing other than the terminal wealth condition. Were we to impose such constraints, not surprisingly, risk premiums in our model would increase. For instance, we find that a borrowing limit equal to 5% of expected, next-period income increases the Sharpe ratio and equity premium to 51.5 percent and 5.0 percent, respectively. Tightening the constraint to 1% increases these values further, to 56.3 percent and 5.5 percent. We do impose a short-sales constraint, but find the effects of removing it to be quantitatively very small (a change in the Sharpe ratio of less than 1 percent). Portfolio constraints, therefore, play no role in our model.

5 Conclusions

Our main question is whether idiosyncratic labor-market risk matters for the pricing of aggregate risk. An inescapable aspect of idiosyncratic risk is that it necessarily has a life-cycle component: the young face more than the old. This arises both directly — workers face shocks but retirees don't — and indirectly, in terms of the inevitable life-cycle pattern in the ratio of human wealth to total wealth. These life-cycle effects are of first-order importance for the question. They imply that a substantial fraction of the population don't care very much about the very shocks which drive the model, and therefore present a challenge to the asset-pricing story. Nevertheless, our main conclusion is that idiosyncratic risk matters and that life-cycle effects actually strengthen its impact.

What drives our model is an interaction between idiosyncratic and aggregate risk which goes beyond the countercyclical-volatility effect emphasized by Constantinides and Duffie (1996), Mankiw (1986) and many subsequent papers. The converse of the fact that younger agents face the most idiosyncratic risk is that older agents face the most aggregate risk. Idiosyncratic risk is difficult to transfer across generations. Aggregate risk is not. Our framework suggests that how the aggregate risk is shared, and how this interacts with the nontradeable distribution of idiosyncratic risk, is important for asset pricing. If, for instance, an equilibrium features equity ownership increasing with age, then the effect of idiosyncratic risk will be diminished relative to the Constantinides and Duffie (1996) model. If equity ownership decreases with age, the opposite will hold. Our calibration generates an intermediate case: equity ownership increases until the late working years and then declines into retirement. The asset pricing effects of idiosyncratic shocks and the countercyclical-volatility effect remain important and our model generates slightly higher risk premia than the infinite-horizon Constantinides and Duffie (1996) model.

More specifically, our model is driven by life-cycle variation in the ratio of human wealth to financial wealth and the fact that idiosyncratic risk affects the former but not the latter. There are two main forces at work. First, as an agent ages, idiosyncratic risk becomes less important to them. This happens both because they face fewer (persistent) shocks in the future and because human wealth declines as a fraction of total wealth. The countercyclical-volatility effect, therefore, becomes less important with age and tolerance for equity-holding increases. Second, because equity returns are substantially more volatile than the wage rate, age also brings with it an increased exposure to aggregate shocks, because an increasing share of an agent’s income derives from financial assets instead of human wealth. This effect eventually counteracts the first effect and, late in the working life, tolerance for equity-holding begins to decrease with age. Taken together, the two effects imply that young agents hold zero equity, retired agents hold diversified portfolios of equity and bonds, and middle-aged agents hold levered equity, issuing bonds to both the young and the old. The risk premium which supports this allocation reflects both the countercyclical-volatility risk emphasized by Constantinides and Duffie (1996), and a “concentration of aggregate risk” upon the middle-aged, alluded to by Mankiw (1986).

Constantinides, Donaldson, and Mehra (1997) (CDM) also stress the importance of life-cycle effects for the equity premium. Like us, an important feature of their model is that young agents hold zero equity, thereby concentrating aggregate risk on older agents. The reasons, however, are fundamentally different than in our framework, which gives rise to stark, testable restrictions between the two. Our model is distinguished by idiosyncratic risk *within* generations. A young agent’s choice to avoid equity is a portfolio allocation decision: equity is too risky, so they choose not to hold any. In the CDM framework, where heterogeneity only exists across generations, the driving force is consumption smoothing and how it interacts with borrowing constraints. Young agents receive a relatively meager endowment, cannot borrow or short sell equity, and therefore choose not to hold any assets whatsoever. The two models, therefore, offer starkly different interpretations of why one might see a young household choose not to hold equity. The testable restrictions are related to overall savings behavior and how important the precautionary motive is. In our model the average young household is a net saver during the first third of their lives. That is, the precautionary motive dominates the life cycle motive, and the decision to avoid equity is driven by *risk*, in our case an avoidance of the countercyclical volatility risk. The CDM framework is consistent with the same average, young household not accumulating any assets but, in contrast, viewing equity (in a shadow value sense) as an attractive investment. Which of these interpretations is more important — it seems clear to us that the world features aspects of each of them — is something we leave to future work.

Another direction for future work involves enriching our notion of idiosyncratic risk, perhaps in relation to the growing body of work suggesting an important link between private business ownership, portfolio choice and asset pricing (*e.g.*, Gentry and

Hubbard (2000), Heaton and Lucas (2000), Polkovnichenko (1999), Quadrini (1999)). While one might interpret our reduced-form process as an amalgamation of the risks faced by both labor and private business owners — *i.e.*, it seems natural that persistence and countercyclical volatility might be an important characteristic of both — it lacks a key aspect of entrepreneurship: the decision to become an entrepreneur, take on the associated risks and bear the associated costs. A better understanding of this decision is likely to enrich our knowledge of how important individual-level risks are for the pricing of aggregate risks.

References

- Abowd, J. and D. Card, (1989), On the covariance structure of earnings and hours changes, *Econometrica* 57, 411–455.
- Aiyagari, S. R., (1994), Uninsured idiosyncratic risk and aggregate saving, *Quarterly Journal of Economics* 109, 659–684.
- Aiyagari, S. R. and M. Gertler, (1991), Asset returns with transactions costs and uninsured individual risk, *Journal of Monetary Economics* 27, 311–331.
- Altonji, J. G., F. Hayashi, and L. J. Kotlikoff, (1991), Risk sharing, altruism, and the factor structure of consumption, NBER working paper number 3834.
- Altug, S. and R. A. Miller, (1990), Households choices in equilibrium, *Econometrica* 58, 543–70.
- Alvarez, F. and U. Jermann, (1999), Quantitative asset pricing implications of endogenous solvency constraints, forthcoming, *Review of Financial Studies*.
- Amerkis, J. and S. P. Zeldes, (2000), How do household portfolio shares vary with age?, Unpublished manuscript, Columbia University.
- Attanasio, O. and S. J. Davis, (1996), Relative wage movements and the distribution of consumption, *Journal of Political Economy* 104, 1227–1262.
- Attanasio, O. and G. Weber, (1992), Consumption growth and excess sensitivity to income: evidence from U.S. micro data, Unpublished manuscript, Stanford University.
- Balduzzi, P. and T. Yao, (2000), Does heterogeneity matter for asset pricing, Unpublished manuscript, Boston College.
- Boldrin, M., L. J. Christiano, and J. D. Fisher, (2001), Habit persistence, asset returns and the business cycle, *American Economic Review* 91, 149–166.
- Brav, A., G. M. Constantinides, and C. Geczy, (2000), Asset pricing with heterogeneous consumers and limited participation: Empirical evidence, Working paper, The Wharton School, University of Pennsylvania.
- Cochrane, J. H., (1991), A simple test of consumption insurance, *Journal of Political Economy* 99, 957–76.
- Cogley, T., (2000), Idiosyncratic risk and the equity premium: Evidence from the consumer expenditure survey, forthcoming, *Journal of Monetary Economics*.

- Constantinides, G. M., J. B. Donaldson, and R. Mehra, (1997), Junior can't borrow: a new perspective on the equity premium puzzle, Working paper, University of Chicago, GSB.
- Constantinides, G. M. and D. Duffie, (1996), Asset pricing with heterogeneous consumers, *Journal of Political Economy* 104, 219–240.
- Deaton, A., (1991), Savings and liquidity constraints, *Econometrica* 59, 1221–1248.
- Deaton, A. and C. Paxson, (1994), Intertemporal choice and inequality, *Journal of Political Economy* 102, 437–467.
- den Haan, W., (1994), Heterogeneity, aggregate uncertainty and the short term interest rate: a case study of two solution techniques, Working paper, University of California at San Diego.
- den Haan, W., (1997), Solving dynamic models with aggregate shocks and heterogeneous agents, *Macroeconomic Dynamics* 1, 355–386.
- den Haan, W. and A. Marcet, (1994), Accuracy in simulations, *Review of Economic Studies* 61, 3–17.
- Gentry, W. M. and R. G. Hubbard, (2000), Entrepreneurship and household saving, Unpublished manuscript, Columbia University.
- Gertler, M., (1999), Government debt and social security in a life cycle economy, *Carnegie Rochester Conference Series on Public Policy* 50, 61–110.
- Gordon, R. J., ed., (1986), *The American Business Cycle: Continuity and Change*, University of Chicago Press.
- Graham, J. R., (2000), How big are the tax benefits of debt?, *Journal of Finance* 55, 1901–1941.
- Greenwood, J., Z. Hercowitz, and G. W. Huffman, (1988), Investment, capacity utilization and the business cycle, *American Economic Review* 78, 402–18.
- Greenwood, J., Z. Hurcowitz, and P. Krusell, (1997), Long-run implications of investment-specific technological change, *American Economic Review* 87, 342–362.
- Heaton, J. and D. J. Lucas, (1996), Evaluating the effects of incomplete markets on risk sharing and asset pricing, *Journal of Political Economy* 104, 443–487.
- Heaton, J. and D. J. Lucas, (2000), Portfolio choice and asset prices; the importance of entrepreneurial risk, *Journal of Finance* 55, 1163–1198.

- Hubbard, R. G., J. Skinner, and S. P. Zeldes, (1994), The importance of precautionary motives in explaining individual and aggregate savings, *Carnegie Rochester Conference Series on Public Policy* 40, 59–126.
- Huggett, M., (1993), The risk free rate in heterogeneous-agents, incomplete insurance economies, *Journal of Economic Dynamics and Control* 17, 953–969.
- Huggett, M., (1996), Wealth distribution in life-cycle economies, *Journal of Monetary Economics* 38, 469–494.
- Jagannathan, R. and N. Kocherlakota, (1996), Why should older people invest less in stocks than younger people?, *Federal Reserve Bank of Minneapolis, Quarterly Review* 20, 11–23.
- Krusell, P. and A. A. Smith, (1997), Income and wealth heterogeneity, portfolio choice, and equilibrium asset returns, *Macroeconomic Dynamics* 1, 387–422.
- Krusell, P. and A. A. Smith, (1998), Income and wealth heterogeneity in the macroeconomy, *Journal of Political Economy* 106, 867–896.
- Lucas, D. J., (1994), Asset pricing with undiversifiable risk and short sales constraints: Deepening the equity premium puzzle, *Journal of Monetary Economics* 34, 325–341.
- Mace, B. J., (1991), Full insurance in the presence of aggregate uncertainty, *Journal of Political Economy* 99, 928–56.
- MaCurdy, T. E., (1982), The use of time series processes to model the error structure of earnings in a longitudinal data analysis, *Journal of Econometrics* 18, 83–114.
- Mankiw, N. G., (1986), The equity premium and the concentration of aggregate shocks, *Journal of Financial Economics* 17, 211–219.
- Marcet, A. and K. J. Singleton, (1999), Equilibrium assets prices and savings of heterogeneous agents in the presence of portfolio constraints, *Macroeconomic Dynamics* 3, 243–277.
- Mehra, R. and E. Prescott, (1985), The equity puzzle, *Journal of Monetary Economics* 15, 145–61.
- Mrkaic, M., (1997), Estimation of incomplete markets models with panel data – the case of the permanent income hypothesis, Working Paper, Carnegie Mellon University.
- Polkovnichenko, V., (1999), Heterogeneity and proprietary income risk: implications for stock market participation and asset prices, Manuscript, Northwestern University.

- Quadrini, V., (1999), The importance of entrepreneurship for wealth concentration and mobility, *The Review of Income and Wealth* 45, 1–40.
- Ramchand, L., (1999), Asset pricing in international markets in the context of agent heterogeneity and market incompleteness, forthcoming, *Journal of International Money and Finance*.
- Ríos-Rull, J. V., (1994), On the quantitative importance of market completeness, *Journal of Monetary Economics* 34, 463–496.
- Ríos-Rull, J.-V., (1996), Life cycle economies and aggregate fluctuations, *Review of Economic Studies* 63, 465–489.
- Sarkissian, S., (1999), The cross-country consumption dispersion and the world business cycle, Unpublished manuscript, McGill University.
- Storesletten, K., (1999), Sustaining fiscal policy through immigration, forthcoming, *Journal of Political Economy*.
- Storesletten, K., C. I. Telmer, and A. Yaron, (2000a), Consumption and risk sharing over the life cycle, Unpublished working paper, Carnegie Mellon University.
- Storesletten, K., C. I. Telmer, and A. Yaron, (2000b), The welfare cost of business cycles revisited: Finite lives and cyclical variation in idiosyncratic risk, forthcoming, *European Economic Review*.
- Telmer, C. I., (1993), Asset pricing puzzles and incomplete markets, *Journal of Finance* 48, 1803–1832.
- Telmer, C. I. and S. E. Zin, (1995), Approximate aggregation with incomplete markets, GSIA working paper no. 1995-E5, forthcoming *Journal of Economic Dynamics and Control*.
- Weil, P., (1992), Equilibrium asset prices with undiversifiable labor income risk, *Journal of Economic Dynamics and Control* 16, 769–790.
- White, H., (1980), A heteroskedasticity-consistent covariance matrix estimator and a direct test of heteroskedasticity, *Econometrica* 48, 817–38.
- Zhang, H., (1997), Endogenous borrowing constraints with incomplete markets, *Journal of Finance* 52, 2187–2209.

A Data Appendix

Additional details regarding the data selection procedure, specific characteristics of the data and the estimation procedure based on equations (1) and (4) are as follows.

Selection Criteria

A household is selected into a panel if the following conditions are met: (i) the head of household is a male, (ii) there are no changes in family structure except for the number of children, (iii) total earnings are positive in each of the sample years, (iv) total earnings growth rates are no larger than 20 and no lower than 1/20 in any consecutive years, (v) the households was not part of the Survey of Economic Opportunity.

Individual Data

The individual earnings data are based on the 1969-1992 Family Files of the PSID. We use the PSID's Individual Files to track individuals across the different years of the Family Files. The definition of earnings includes wage earnings by head of household plus female wage earnings plus total transfers to the household. Total transfers include unemployment insurance, workers compensation, transfers by non-household family members, and several additional (minor) categories. Total earnings are then deflated to common 1968 dollars using the CPI index. Earnings are then converted to per-household-member rates by dividing total earnings by family size.

3-year repeated panels

Agents are selected into a panel if they meet the selection criteria above for the two years following the base year of the panel. We start with the 1969 PSID file (and therefore follow agents through PSID files from 1970 and 1971) which we denote panel 1968, since each PSID file provides information on the previous year's income. Our last 3-year panel, which we denote as the 1989 panel, starts with the 1990 PSID file and ends with 1992 PSID file. Hence, we end up with 22 3-year repeated panels.

The table below reports summary statistics of the 3-year repeated panels we use in our estimations.

Year	# of Households	Age of Head Household		Year	# of Households	Age of Head Household	
		Mean	Std. Deviation			Mean	Std. Deviation
1968	1623	46.71	14.32	1979	2094	43.35	15.11
1969	1663	46.42	14.44	1980	2121	43.53	15.34
1970	1702	45.94	14.77	1981	2128	43.81	15.13
1971	1737	45.44	15.03	1982	2122	43.81	14.90
1972	1792	44.78	15.09	1983	2163	43.87	14.82
1973	1820	44.41	15.20	1984	2213	43.64	14.90
1974	1875	44.24	15.13	1985	2242	43.77	14.68
1975	1927	44.16	15.27	1986	2235	44.34	14.72
1976	2001	43.75	15.38	1987	2273	44.43	14.55
1977	2030	43.58	15.50	1988	2284	44.65	14.40
1978	2045	43.76	15.28	1989	2342	44.40	14.08

Aggregate Data

The aggregate GNP levels used to define the indicator functions described in (4) were derived by merging GNP data from Gordon (1986) for the years 1910-1958 with CITIBASE data for the years 1959-1992. These were converted to 1968 dollars using the CPI and to per-capita terms by dividing by total population figures from CITIBASE.

The NBER based business cycles are derived from the monthly NBER definitions of contractions and expansions. We converted these monthly definitions into yearly ones. Our criteria, corresponding to the NBER definition, is to define years of recession, as those years for which the majority of the months are contractionary. For cases where contractions spanned less than 12 months but more than 6 months over two calendar years, we designated the first year as a recession. Our results are robust to close alternatives.

The ‘Unemployment’ definition of cycles is based on annual unemployment rates. This is constructed from two sources. For the years 1910-1970 we use Historical Statistics of the United States (page 135). For 1971-1992, we used the Economic report of the president (February 1999, page 376). To convert these rates into contractionary years we used information about the levels of unemployment as well as direction of changes. Recessions are defined as years for which unemployment rate was greater than 7.5%. If the rate was greater than 7.5% but fell more than 3% relative to the previous year, it was not counted as a recession year. In addition years for which the unemployment rate had risen by more than 3% were defined as recession years.

Figures

Panels A and B of Figure 1 are based on decomposing the data into *age* and *cohort* components following Deaton and Paxson (1994) and Storesletten, Telmer, and Yaron (2000a). Let $\sigma_{h,t-h}^2$ be the cross sectional variance of u_{it}^h across agents of age h and cohort $c = t - h$ (i.e., those born in year c). Using the $H \times T$ observations on $\sigma_{h,t-h}^2$, where T are the number of panels we have, we run the following regression

$$\sigma_{h,t-h}^2 = a_c + b_h + e_{h,c}$$

where a_c and b_h are cohort and age dummy coefficients.

Figure 1-A, plots the estimated $\{b_h\}_{h=1}^H$, where they are scaled so the graph passes through the unconditional (across cohorts) variance of u_{it}^{40} . In Storesletten, Telmer, and Yaron (2000a) we show how to cast this regression as a system of moment conditions depending on the primitive observation u_{it}^h , and hence to account for standard errors and use the moment conditions to estimate a homoskedastic analog to equation 2.

Figure 1-C, provides a scatter-plot of the estimated $\{a_c\}_{c=1}^{H+T-1}$ against the corresponding fraction of recessions that each cohort has worked through. In Figure 1-B,

we plot the variance and the linearly-detrended mean of each panel. That is we plot against time the linearly detrended $\mu_t \equiv E[y_{i,t}^h|t]$ and $\sigma_t^2 \equiv E[(y_{i,t}^h)^2|t]$.

Monte-Carlo Simulation

We simulate 500 economies with the income process described by (2) with an aggregate shock following the calibrated economy described in Appendix B. Each such economy is simulated with 3000 agents per cohort and for 24 periods. We drop the first 100 periods and end with 24 panels for each of the simulated economies. Let $cv_{s,i}$ be the cross sectional variance for the i -th panel, $i = 1, 2, \dots, 24$ of the s^{th} simulated economy, $s = 1, 2, \dots, 500$. Equation (7) correspond to the population variance,

$$cv_{s,i} = \sum_{h=1}^H \varphi_h \left(\sigma_\varepsilon^2 + \sum_{j=0}^{h-1} \rho^{2j} (I_{i-j} \sigma_E^2 + [1 - I_{i-j}] \sigma_C^2) \right), \quad (31)$$

where φ_h are population shares and the variances are based on the pooled cross-section.

We define for each simulated economy s , based on the 24 panels, two statistics:

$$\begin{aligned} Z_{1,s} &= \frac{\max_i(cv_{s,i})}{\min_i(cv_{s,i})} \\ Z_{2,s} &= \max_i |cv_{s,i} - cv_{s,i-1}| \end{aligned}$$

The reported p-values in the text correspond the proportion of $\{Z_{1,j}\}_{j=1}^{500}$ and $\{Z_{2,j}\}_{j=1}^{500}$ that are larger than the point estimate of these statistics in our 24 PSID panels.

Longitudinal Panel

For comparison, we extract a 24 year panel in which we impose the same selection criteria as that imposed in the repeated 3 year panels above. We obtain a sample of 610 households with a mean and standard deviation of 38.9 and 10.54 respectively for the age of the head of households in the 1968 panel. By 1991 the only change is in the mean age which is 61.90. Using this panel, the estimate of the persistence parameter ρ is .931 – is essentially identical to the homoskedastic estimate in Table (1).

Our estimates suggest substantially more persistence in idiosyncratic income than do those of Heaton and Lucas (1996). We do find that when we adopt their approach — estimating household-specific parameters and averaging across the estimates — we get qualitatively similar results. We estimate ρ to be 0.640 and 0.498 using their sample and our sample, respectively. The major differences between our approach and theirs are as follows. We have a longer time dimension in our panel and detrend the data somewhat differently (including life cycle and education effects). We constrain the parameters in our time series model to be the same across all households in our panel.

The moment conditions underlying our estimates condition on household age, which we find can have substantial effects. Finally, we do not explicitly allow for household-specific fixed-effects as they do, by estimating an intercept parameter, per-household. This turns out to be an important omission; we find that taking out the mean reduces our estimate of ρ to 0.750. Our feeling is that there is no easy answer here: our approach is more economical on parameters but is likely to overstate persistence by modeling deterministic cross sectional variation as a function of education alone.

We do find that our estimate of ρ is quite robust to alternative estimation procedures such as quasi differencing (i.e. basing the moment conditions on $u_{it} - \rho u_{it-1}$) and first differencing of the idiosyncratic data u_{it} – where depending on the selected years in the sample the lowest estimate of ρ was .83 but in most cases was approximately 0.9.

A.1 Estimation

Our estimation procedure has two distinct steps. In the first stage we estimate equation (1). In the second stage we estimate the system given in (4).

To recover u_{it}^h in the first stage, we need to specify $g(Y_t, x_{it}^h)$. We let $g(\cdot)$ be represented by a simple linear regression using as regressors year dummies (corresponding to Y_t) and age, age squared and education (corresponding to individual specific attributes x_{it}^h). Let the parameters of this first stage regression be summarized by θ_1 . That is $\psi_1(y_{i,t}^h, Y_t, x_{it}^h, \theta_1) \equiv y_{it}^h - \theta_1' [1, Y_t, x_{it}^h]$. We use standard Wald test to determine our first stage regressors. All of the coefficients are significant and the R^2 is 0.22.

Next, let the parameters of the system in (4) be denoted by θ_2 (that is $\sigma_E, \sigma_C, \sigma_\epsilon, \rho$). The joint system we estimate can be written compactly as

$$E \begin{bmatrix} \psi_1(y_{it}^h, Y_t, x_{it}^h, \theta_1) \\ \psi_2(y_{it}^h, Y_t, \theta_1, \theta_2) \end{bmatrix} = 0 \quad (32)$$

where ψ_1 and ψ_2 are the moment conditions corresponding to (1) and (4) respectively. The triangular structure of the moment condition allows us to get consistent estimates of θ_1 using only ψ_1 . We then estimate θ_2 using moment conditions ψ_2 . This second step incorporates the standard errors in estimating θ_1 using standard two-step GMM procedure. The additional complication that arise in our set-up is due to the overlapping structure of our repeated panels. Since these panels overlap each other by 2 years an MA(2) correction is added to the estimate of the covariance matrix associated with moment conditions ψ_2 .

Specifically define moment conditions ψ_2 to be

$$\psi_{2,i,t}^{h,0} \equiv [(u_{i,t}^{h,t})^2 - E(\sigma_\epsilon^2 + \sum_{j=0}^{h-1} \rho^{2j} (I_{t-j} \sigma_C^2 + [1 - I_{t-j}] \sigma_E^2))]$$

$$\begin{aligned}\psi_{2,i,t}^{h,1} &\equiv [u_{i,t}^{h,t} u_{i,t+1}^{h+1,t} - E(\rho \sum_{j=1}^{h-1} \rho^{2(j-1)} (I_{t-j} \sigma_C^2 + [1 - I_{t-j}] \sigma_E^2))] \\ \psi_{2,i,t}^{h,2} &\equiv [u_{i,t}^{h,t} u_{i,t+2}^{h+2,t} - E(\rho^2 \sum_{j=1}^{h-1} \rho^{2(j-2)} (I_{t-j} \sigma_C^2 + [1 - I_{t-j}] \sigma_E^2))]\end{aligned}$$

where the superscript t in u denotes the base year of the panel from which this agent is selected. By assumption $\psi_{2,i,t}^{h,j}$ is not correlated with $\psi_{2,i,t+k}^{h,l} \forall k \neq 0$, and $j, l = 0, 1, 2$. It can be easily shown that due to the overlap of the sample, for each t , $\psi_{2,i,t}^{h,0}$, $\psi_{2,i,t}^{h,1}$ and $\psi_{2,i,t}^{h,2}$ are correlated. We stack the repeated 3 moment conditions and use sample counterparts to estimate these covariance terms – where the covariance matrix is block-diagonal and each 3×3 block has non-empty off-diagonal elements.

Our results are robust to selecting fewer moment conditions, using subsets of the years above, and to a system in which the parameters of interest are exactly identified. For the exactly identified case, we also experimented with repeated panels of 4 years and the additional moment condition

$$\psi_{2,i,t}^{h,3} \equiv [u_{i,t}^{h,t} u_{i,t+3}^{h+3,t} - E(\rho^3 \sum_{j=1}^{h-2} \rho^{2(j-3)} (I_{t-j} \sigma_C^2 + [1 - I_{t-j}] \sigma_E^2))],$$

and then use the four sample counter-parts $\frac{1}{T} \sum_{t=1}^T \frac{1}{N_t} \sum_{i=1}^{N_t} \psi_{2,i,t}^{h,j}$ for $j = 0, 1, 2, 3$, as the moment conditions.

B Calibration Appendix

This appendix first describes the calibration of the no-trade (Constantinides and Duffie (1996)) economies in Section 3 and Table 2, and then goes on to describe the calibration of the economies with trade, presented in Section 4 and Table 2. It also demonstrates the sense in which our specification for countercyclical volatility — heteroskedasticity in the innovations to the idiosyncratic component of log income — is consistent with the approach used by previous authors (*e.g.*, Heaton and Lucas (1996), Constantinides and Duffie (1996)). In each case the cross sectional variance which matters turns out to be the variance of the *change* in the log of an individual’s share of income and/or consumption.

Calibration of No-Trade Economies

Aggregate consumption growth follows a two-state Markov chain, identical to that in Mehra and Prescott (1985). We parameterize the process using the same values as they did, with the mean, standard deviation and autocorrelation of aggregate consumption growth being 0.018, 0.036, and -0.14, respectively. The Constantinides and Duffie (1996) model is then ‘calibrated’ via a re-interpretation of the preference parameters of the Mehra and Prescott (1985) representative agent. Recall that we use β and γ to denote an *individual* agent’s utility discount factor and risk aversion parameters, respectively. Constantinides and Duffie (1996) construct a representative agent (their equation (16)) whose rate of time preference and coefficient of relative risk aversion are (using our notation),

$$-\log \beta^* = -\log(\beta) - \frac{\gamma(\gamma + 1)}{2} a \quad , \quad (33)$$

and

$$\gamma^* = \gamma - \frac{\gamma(\gamma + 1)}{2} b \quad , \quad (34)$$

respectively. In these formulae, the parameters a and b relate the cross sectional variance in the change of the log of individual i ’s share of aggregate consumption (y_{t+1}^2 , using Constantinides-Duffie’s notation) to the growth rate of aggregate consumption:

$$\text{Var}(\log \frac{c_{i,t+1}/c_{t+1}}{c_{it}/c_t}) = a + b \log \frac{c_{t+1}}{c_t} \quad . \quad (35)$$

All that we require, therefore, are the numerical values for a and b which are implied by our PSID-based estimates in Table 1.

Our estimates are based on income, y_{it} . Because the Constantinides-Duffie model is autarkic, we can interpret these estimates as pertaining to individual consumption, c_{it} . Balduzzi and Yao (2000), Brav, Constantinides, and Geczy (2000), and Cogley

(2000) take the alternative route and use microeconomic consumption data. While their results are generally supportive of the model, they each point out serious data problems associated with using consumption data. Income data is advantageous in this sense. In addition, our objective is just as much relative as it is absolute. That is, consumption is endogenous in the model of Section 4, driven by risk sharing behavior and the exogenous process for idiosyncratic income risk estimated in Section 2. What Table 2 asks is, “what would the Constantinides-Duffie economy look like, were its agents to be endowed with idiosyncratic risk of a similar magnitude?” Also, “how does our model measure up, in spite of its non-degenerate (and more realistic) risk sharing technology?” Using income data seems appropriate in this context. For the remainder of this appendix we set $c_{it} = y_{it}$.

We need to establish the relationship between our specification for idiosyncratic shocks and the log-shares of aggregate consumption in equation (35). Denote individual i 's share at time t as γ_{it} , so that,

$$\log \gamma_{it} \equiv \log c_{it} - \log \tilde{E}_t c_{it} \ ,$$

where the notation $\tilde{E}_t(\cdot)$ denotes the cross-sectional mean at date t , so that $\tilde{E}_t c_{it}$ is date t , per-capita aggregate consumption. For our specification, if we ignore the terms which capture cross-sectional variation due to age and education (see equation (1) in Section 2), then our estimation in Table 1 boils down to a time series model of the residuals from a regression involving only year-dummy variables. In a large cross section this will be,

$$z_{it} = \log c_{it} - \tilde{E}_t \log c_{it} \ ,$$

which have a cross-sectional mean of zero, by construction, and a sample mean of zero, by least squares. The difference between our specification and the log-share specification is, therefore,

$$\begin{aligned} \log \gamma_{it} - z_{it} &= \tilde{E}_t \log c_{it} - \log \tilde{E}_t c_{it} \\ &= \tilde{E}_t \log \gamma_{it} - \log \tilde{E}_t \gamma_{it} \ . \end{aligned}$$

The share, γ_{it} , is *defined* so that its cross-sectional mean is always unity. The second term is therefore zero. For the first term, note that in both our economy and the statistical model underlying our estimates, the cross sectional distribution is log normal, *conditional* on knowledge of current and past aggregate shocks. If some random variable x is log normal and $E(x) = 1$, then $E(\log x) = -\text{Var}(\log x)/2$. As a result,

$$\log \gamma_{it} - z_{it} = -\frac{1}{2} \tilde{V}_t(\log \gamma_{it}) \ ,$$

where \tilde{V}_t denotes the cross-sectional variance operator. Because lives are finite in our model, and because we interpret data as being generated by finite processes, this cross-sectional variance will always be well defined, irrespective whether or not the shocks are unit root processes.

The quantity of interest in equation (35) can now be written as,

$$\begin{aligned} \log \frac{c_{i,t+1}/c_{t+1}}{c_{it}/c_t} &\equiv \log \gamma_{i,t+1} - \log \gamma_{it} \\ &= z_{i,t+1} - z_{it} - \frac{1}{2} \left(\tilde{V}_{t+1}(\log \gamma_{i,t+1}) - \tilde{V}_t(\log \gamma_{it}) \right) \end{aligned} \quad (36)$$

The term in parentheses — the difference in the variances — does not vary in the cross section. Consequently, application of the cross-sectional variance operator to both sides of equation (36) implies,

$$\tilde{V}_{t+1} \left(\log \frac{c_{i,t+1}/c_{t+1}}{c_{it}/c_t} \right) = \tilde{V}_{t+1} (z_{i,t+1} - z_{it}) .$$

Ignoring the transitory shocks, the process underlying the estimates in Table 1 is:

$$z_{i,t+1} - z_{it} = (1 - \rho)z_{it} + \eta_{i,t+1} ,$$

where the variance of $\eta_{i,t+1}$ depends on the aggregate shock. For values of ρ close to unity the variance of changes in z_{it} is approximately equal to the variance of $\eta_{i,t+1}$. The left side of equation (35) is, therefore, approximately equal to the variance of innovations, $\eta_{i,t+1}$,

$$\tilde{V}_{t+1} \left(\log \frac{c_{i,t+1}/c_{t+1}}{c_{it}/c_t} \right) \approx \tilde{V}_{t+1} (\eta_{i,t+1}) .$$

For unit root shocks — which we assume for most of Section 4, this holds exactly. The estimates of σ_E and σ_C in Table 1 are therefore sufficient to calibrate the Constantinides-Duffie model.

All that remains are to map our estimates into numerical values for a and b from equation (35). Since aggregate consumption growth — the variable on the right side of equation (35) — takes on only two values (3.8 percent and -0.8 percent), computing the parameters a and b simply involves two linear equations:

$$\begin{aligned} \sigma_E^2 &= a + 0.038b \\ \sigma_C^2 &= a - 0.008b , \end{aligned}$$

The estimates in Table 1 are $\sigma_E^2 = 0.037$ and $\sigma_C^2 = 0.181$. These estimates, however, are associated with $\rho = .92$. For our unit root economies, we scale them down so as to maintain the same average unconditional variance (across age). This results in $\sigma_E^2 = 0.0095$ and $\sigma_C^2 = 0.0467$. The resulting values for a and b are $a = 0.0374$ and $b = -0.5160$.

Calibration of Models with Trade

The models in Section 4 are calibrated as follows. A period is interpreted as one year. The aggregate shock in equation (25) follows a first-order Markov chain with values $Z \in \{0.9725, 1.0275\}$. The unconditional probabilities are 0.5 and the transition probabilities are such that the probability of remaining in the current state is 2/3 (so that the expected duration of a ‘business cycle’ is 6 years). Capital’s share of output, θ from equation (25), is set to 0.40, and the average annual depreciation rate, δ , is set to match the average riskfree rate of 1.3 percent. This results in $\delta = 0.164$. The average wage rate, w , is set equal to $(1 - \theta) E(K)^\theta N^{-\theta}$. The parameter s is chosen so that the standard deviation of the risky return, r_t , is 10 percent. The secular growth rate, by which all quantities are normalized, is 1.5 percent.

In Section 4.2, the persistent component of hours worked, z_{it}^h , follows a unit root process with innovations governed by a four-state Markov chain, two states corresponding to an expansion and the other two a contraction. The conditional variances, σ_E^2 and σ_C^2 , are set to 0.0095 and 0.0467 respectively, which are taken from Table 1 and then scaled down so that the unconditional variance matches that of the $\rho = 0.92$ process (a value of 0.0281). In the case of mean-reverting shocks, we implement the estimates from Table 1 directly, by imposing the following restrictions on the approximated 4-state discrete Markov process: (i) the conditional variance in recessions and booms should be 0.181 and 0.037, (ii) the autocorrelation should be 0.92, (iii) the transition matrix Γ should be symmetric. It is infeasible to match the autocorrelation exactly, so the Markov chain we use has an autocorrelation of only 0.890. The elements of the process z are $z \in \{-2.385, 0.646\} \equiv \{z_{rl}, z_{rh}\}$ in recessions and $z \in \{-0.904, 0.467\} \equiv \{z_{bl}, z_{bh}\}$ in booms. The transition matrix is

	(z'_{rl})	(z'_{rh})	(z'_{bl})	(z'_{bh})
(z'_{rl})	.784	.016	.196	.004
(z'_{rh})	.016	.784	.004	.196
(z'_{bl})	.196	.004	.784	.016
(z'_{bh})	.004	.196	.016	.784

The parameters κ_h are chosen so as to match the PSID mean age profile in earnings.

The demographic structure is calibrated to correspond to several simple properties of the U.S. work force. Agents are ‘born’ at age 22, retire at age 65 and are dead by age 85. ‘Retirement’ is defined as having one’s labor income drop to zero and having to finance consumption from an existing stock of assets.

Risk aversion is set to match the Sharpe ratio (see Table 2), and the discount factor, β , is chosen so that the average capital to output ratio generates aggregate consumption variability of 3.7 percent. This results in $\beta = 0.69$ and an associated (average) capital to output ratio of 1.8.

The following Table illustrates the aggregate properties of our economy. The sample size for the U.S. data is chosen to be the same as that used by Mehra and Prescott (1985)

Table B1: Aggregate Moments

<i>Panel A: Population Moments of Growth Rates, Theoretical Economy</i>			
	Std Dev	Autocorrelation	Correlation with Output
Output	0.930	-0.307	1.000
Investment	1.214	-0.309	0.998
Consumption	0.038	-0.117	0.935

<i>Panel B: Sample Moments of Growth Rates, U.S. Economy, 1929-1982</i>			
	Std Dev	Autocorrelation	Correlation with Output
Output	0.062	0.561	1.000
Investment	0.358	0.225	0.143
Consumption	0.036	0.353	0.471

U.S. sample moments are based on annual NIPA data, 1929-1982. Theoretical moments are computed as sample averages of a long simulated time series.

As is discussed in the text, the production side of our economy is unrealistic. Aggregate consumption variability, however, matches the data, as does the variability of the risky asset return (discussed above). Our model does not resolve the well-known problems with production-based asset pricing models. It is best viewed in the same way one views any endowment economy: a model with realistic properties of aggregate consumption which can be supported by some, potentially unrealistic, production technology. Alternatively, one can view our economy as featuring a linear technology — commonplace in the finance literature — where we are explicit about the implications of from the production side of the model — not very commonplace in the finance literature.

C Computational Appendix

Our general solution strategy follows the work of den Haan (1994), den Haan (1997) and, in particular, Krusell and Smith (1997) and Krusell and Smith (1998). The crucial step is the specification of a finite dimensional vector to represent the law of motion for μ . Given this, each individual faces a finite-horizon dynamic programming problem. The essence of the fixed point problem is the consistency of the law of motion for μ with the law of motion implied by individual decisions. More specifically, our algorithm involves the following steps.

Algorithm

1. Approximate the distribution of agents, μ , with a finite number of moments or statistics, μ_m . The idea is to capture the information relevant for portfolio decisions in an efficient way as possible. Natural candidates are various moments of individual wealth and bond holdings. Instead, we use aggregate capital and the conditional expected equity premium ξ_t as moments.¹⁸ Note that ξ_t is in an agent's period t information set. The seemingly unconventional state variable – a conditional price – captures in an efficient way the price information subsumed in a range of equity and bond-holding moments.
2. To solve agent's dynamic programming problem it is necessary to forecast both μ'_m and ξ' . We approximate the agents' expectations for the law of motion of μ_m and ξ by

$$(\mu'_m, \xi') = \hat{G}(\mu_m, \xi, Z, Z') = A(Z, Z') \times (\mu_m, \xi) \quad (37)$$

where $A(Z, Z')$ is an $(m+1) \times (m+1)$ matrix (conditional on Z and Z'), and the entries of μ'_m on ξ are zeros (the first m rows of the $m+1$ column). The aggregate shock Z can take on two values, $Z \in \{\underline{Z}, \bar{Z}\}$, so each element in the matrix $A(Z, Z')$ above can take on four different values. Assume a particular set of values for $A(Z, Z') \forall Z, Z' \in \{\underline{Z}, \bar{Z}\}$.

3. Using the specification above, we solve the following modified version of (30):

$$\hat{V}_h(\xi, \mu_m, Z, z, \epsilon, a) = \max_{b'_{h+1}, k'_{h+1}} \{u(c_h) + \beta E [\hat{V}'_{h+1}(\hat{G}(\mu, Z, Z'), Z', z', \epsilon', k'_{h+1} \cdot R(\hat{G}(\mu, Z, Z'), Z') + b'_{h+1})]\} \quad (38)$$

¹⁸The conditional expected equity premium is defined as $\xi_t \equiv E_t\{R_{t+1}\} - q_t^{-1}$, where R_{t+1} is the return on equity in period $t+1$ and q_t is the period t price of a claim that pays one unit of the consumption good in period $t+1$. Note that, given ξ_t and conditional expectations over the future states of the world, the implicit bond price is $q_t = (E_t\{R_{t+1}\} - \xi_t)^{-1}$. We use ξ because it fluctuates substantially less than q , which implies that our approximation of the decision rules become more accurate.

subject to (29).¹⁹ The implementation of this is described below.

4. Assume an initial distribution of a large, but finite, number of agents, μ , across wealth, idiosyncratic shocks and age (we use 500 agents in each age cohort). Using the decision rules obtained in (38), simulate a long sequence of the economy (5100 periods) and discard the first 100 periods from this sequence. Note that, for each period in time, ξ must be set so that the bond market clears. That is, find a ξ^* such that $\int b'_h(\mu_m, \xi^*, Z, z, \epsilon, a) d\mu = 0$.
5. Update \hat{G} by running a linear regression of (ξ^t, μ'_m) on μ_m from the realized sequence in Step 4. If the coefficients change, use the updated \hat{G} and return to Step 3. Continue this process until convergence.
6. Evaluate the ability of \hat{G} to forecast μ'_m and ξ^t . If the goodness of fit is not satisfactory, return to Step 1 and increase the number of moments or change the functional form of \hat{G} .

Moments of μ_m and accuracy

Following Krusell and Smith (1997), we began with just the first moment, aggregate capital, $\mu_1 = \log(\bar{k})$. This variable has strong predictive power on $\log(\bar{k}^t)$ (R^2 of 0.9998), but less predictive power on ξ^t (see Table C1).

Next, we ask what other moment(s) matter for forecasting ξ^t . To this end, we collected long time series of 18 additional moments of the distribution of agents (see Table C1 for details). Of these, the moments with the largest marginal improvement of forecast accuracy of ξ^t (over the forecast including only $\log(\bar{k}^t)$) are the wealth of workers and the fraction of agents constrained in the bond market, which each improve the R^2 with on average 0.05 and 0.03, respectively. Including all the 18 moments (together with $\log(\bar{k})$) increase the forecast accuracy to 0.994. Finally, we regressed ξ^t on ξ and $\log(\bar{k})$, and found that the R^2 increased to 0.9992, with a standard deviation of forecast error less than 0.02% of ξ . Hence, ξ provides a better forecast than all the 18 moments together. Our interpretation of this finding is that the the conditional expected equity premium or, equivalently, the bond price, capture a large amount of information relevant for future bond prices. Formally, we could not reject the hypothesis that the residuals from predicting (μ'_m, ξ^t) by $\hat{G}(\mu_m, \xi, Z, Z')$ are uncorrelated with the 18 variables described above for the simulations we employ.

In summary, we include ξ as an “endogenous” moment in μ_m in order to improve the forecast of ξ^t . Note that, as the value functions explicitly incorporate ξ as a parameter, including ξ in the forecast of ξ^t come at zero computational cost.

¹⁹Note that in order to ensure that the bond market clears each period, ξ is included as an argument in the value function (Krusell and Smith (1997) use the bond price). One difference from their approach is that, as ξ enters the value function for all age groups, it does not simplify our computations to exclude ξ^t from next period value functions and rely on “approximate” future market clearing.

Dynamic Programming Problem

We now describe how the dynamic programming problem in (38) is solved.

1. First, we choose a grid for the continuous variables in the state space. That is, we pick a set of values for \bar{k} , ξ , and a . The grid points are typically chosen to lie in the stationary region of the state variables and in addition, for wealth, near the borrowing constraint and far in excess of the maximum observed wealth holdings (conditional on age). We pick 11 points for aggregate capital, 11 points for the conditional expected equity premium, and 50 points for individual wealth at each age.
2. Second, we make piecewise linear approximations to the decision rules by solving for portfolio holdings on the grid and iterating on the Euler equations.

This is done in the following way. Given the terminal condition associated with (29), the decision rules of the oldest agents (H years old) must be $b'_{H+1} = k'_{H+1} = 0$, in any state of the world. That is, the agent consumes all their wealth.

Knowing c_H , we can in turn compute b'_H and k'_H at each grid point using Euler equations of an $H - 1$ year old agent (and imposing the borrowing constraints and Kuhn-Tucker conditions):

$$\begin{aligned} u'_{H-1}(c_{H-1}) &\geq E\{u'_H(c'_H)R' \mid \mu_m, Z, z, \epsilon\} \\ qu'_{H-1}(c_{H-1}) &\geq E\{u'_H(c'_H) \mid \mu_m, Z, z, \epsilon\} \end{aligned} \quad (39)$$

Knowing b'_H and k'_H at each grid point, we then obtain a piecewise linear approximation of the decision rules by linear interpolation (outside the grid we do linear extrapolation). Computing c_{H-1} is then straightforward, and this procedure is repeated for $H - 2$ year old agents and iterated backwards until $h = 1$. Note that no further iterations are needed; given the (imperfect) expectations \hat{G} and the decision rules for $h + 1$ years old agents, the piecewise approximations are found in one single step for h years old agents.

Accuracy of approximation

The solutions on the grid points are exact by construction. To evaluate whether the interpolation between grid points gives rise to systematic Euler equation pricing errors we follow den Haan and Marcet (1994) and use simulation to construct the following moment conditions:

$$g(c, Z, z, R, q) \equiv \frac{1}{T} \sum_{t=1}^T \sum_{h=1}^H \frac{1}{N_h^*} \sum_{i=1}^{N_h^*} \left[\left(\beta \left(\frac{c_{i,t+1}^{h+1}}{c_{i,t}^h} \right)^{-\alpha} \left[R_{t+1} \frac{1}{q_t} \right] - 1 \right) \otimes z_{i,t}^h \right] \quad (40)$$

where T is the number of periods in the simulation, N_h^* is the number of *unconstrained* agents within age cohort h , and the instruments are $z_{i,t}^h = \{1, a_{i,t}^h, R_t, R_{f,t-1}\}$. We

calculate the p-value corresponding to the χ_4^2 statistic based on moment conditions $g(\cdot)$ and its covariance matrix. As in den Haan and Marcet (1994) we repeat this for $S = 300$ independent simulations and a sample size of 10000. The percent of p-values that were above (below) the upper (lower) 5% critical value were 5.4% and (4.1%) respectively. By this formal metric the Euler equation errors are quite accurate.

Table C1
Predictability

Regressors	R^2			
	(1,1)	(1,2)	(2,1)	(2,2)
$\log(\bar{k}), \xi$.999	.999	.999	.999
$\log(\bar{k})$.740	.907	.865	.922
$\log(\bar{k}), \{X_j\}_{j=1}^6$.842	.931	.909	.941
$\log(\bar{k}), \{X_j\}_{j=1}^6, \{X_j^*\}_{j=1}^6$.984	.997	.996	.996
$\log(\bar{k}), \{X_j\}_{j=1}^6, \{X_j^*\}_{j=1}^6, \{B_j\}_{j=2}^6$.986	.998	.997	.997
$\log(\bar{k}), \{X_j\}_{j=1}^6, \{X_j^*\}_{j=1}^6, \{B_j\}_{j=2}^6, fr_{bt}, fr_{st}$.987	.998	.997	.997
$\log(\bar{k}), X_1^*$.803	.951	.922	.962
$\log(\bar{k}), fr_{bt}$.802	.923	.898	.935

$X_j \equiv E[(w_{it} - E_t w_{it})^j]$, w_{it} denotes the wealth of agent i at time t , $X_j^* \equiv E[(w_{it}^* - w_{it}^*)^j]$, w_{it}^* denotes the wealth of *working* agent i at time t , $B_j \equiv E[(b_{it} - b_{it})^j]$, and b_{it} denotes the bond holdings of agent i at time t . fr_{bt} and fr_{st} denote the fraction of agents that are constrained in time t at the bond and equity market respectively. The bottom two rows represent the top two individual regressors among all regressors other than $\log(\bar{k})$ and ξ . The reported R^2 are of regressions of $\xi' = a(Z, Z') + X(Z, Z') * B(Z, Z')$

Table 1
Idiosyncratic Endowment Process: Parameter Estimates

	ρ	σ_{η}^2	σ_E^2	σ_C^2	σ_{ε}^2
A. Homoskedastic Innovations					
Estimate	0.935	0.061	–	–	0.017
Standard Error	0.008	0.004	–	–	0.005
B. Heteroskedastic Innovations					
Estimate	0.916	–	0.037	0.181	0.025
Standard Error	0.009	–	0.007	0.033	0.007
C. NBER Definition					
Estimate	0.922	–	0.039	0.161	0.022
Standard Error	0.017	–	0.012	0.047	0.011
D. Unemployment Definition					
Estimate	0.910	–	0.035	0.178	0.027
Standard Error	0.018	–	0.013	0.051	0.012

Entries are GMM estimates, based on earnings data from the Panel Study on Income Dynamics (PSID) and the age-dependent moments described in Section 2 — equations (4-6) — which are based on the process for idiosyncratic shocks, equation (2). The parameters σ_E^2 and σ_C^2 denote the variance of the innovation to the persistent component, conditional on an aggregate expansion, E, or contraction, C, respectively. Estimates in Panel B define expansion/contraction in terms of the growth rate of U.S. GNP being above/below its mean, those in Panel C use the NBER definition, and those in Panel D are based on U.S. unemployment data. Further details, including all data sources, are available in Appendix A.

Standard errors are computed using the White (1980) estimator and incorporate sampling uncertainty from the first-stage regression based upon equation (1). The p-value for a test of $\sigma_E^2 = \sigma_C^2$ is 0.041.

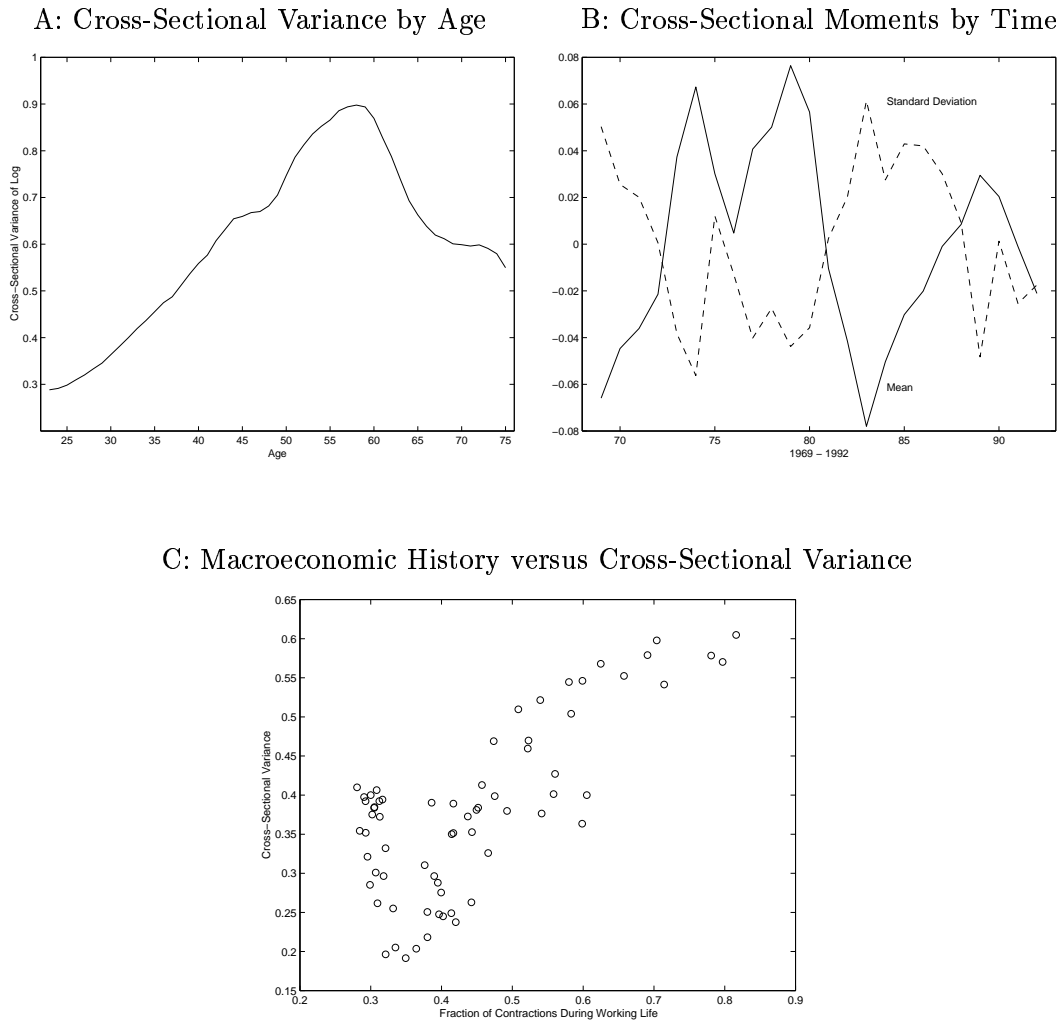
Table 2
Asset Pricing Properties

	Risk Aversion	<u>Riskfree Rate</u> Mean	Std Dev	<u>Equity Premium</u> Mean	Std Dev	Sharpe Ratio
U.S. data		1.30	1.88	6.85	16.64	41.17
U.S. data, unlevered		1.30	1.88	4.11	10.00	41.17
Models Without Trade (Constantinides-Duffie):						
No Retirement (match SR)	9.4	1.30	11.91	13.10	31.81	41.2
No Retirement (match EP)	5.3	1.30	6.43	4.11	17.32	23.7
Retirement (SR)	9.4	1.30	6.14	2.91	14.77	19.7
Retirement (EP)	5.3	1.30	3.73	1.31	9.73	13.5
Models with Trade:						
Main Model	6.7	1.30	2.79	4.21	9.71	43.3
Homoskedastic Economy	6.7	3.35	3.24	1.81	9.68	18.7
Mean-Reverting Shocks ($\rho = 0.92$)	6.7	-0.50	3.39	4.53	8.32	54.4

‘Models Without Trade’ correspond to a calibration of the Constantinides and Duffie (1996) model using the idiosyncratic risk estimates from Table 1 and the aggregate consumption moments from Mehra and Prescott (1985). Details are given in Appendix B. In rows labeled ‘match SR’ and ‘match EP,’ risk aversion is chosen to match the U.S. Sharpe ratio and the mean equity premium, respectively. Rows labeled ‘Retirement’ hold risk aversion at these levels and then incorporate retirement, defined as old agents not receiving any idiosyncratic shocks (Section 3.1). Should risk aversion be re-selected to match SR and EP, the implied values are 19.8 and 11.2, respectively. ‘Models with Trade,’ described in Section 4, enhance the definition of retirement by incorporating life-cycle savings and a life-cycle pattern in the ratio of human to total wealth. Risk aversion in the ‘Main Model’ is chosen to match the U.S. Sharpe ratio. The ‘Homoskedastic Economy’ is distinguished by the volatility of idiosyncratic shocks *not* varying with aggregate shocks. The ‘Mean-Reverting’ economy reduces the autocorrelation of idiosyncratic shocks to 0.92, holding all other parameter values constant.

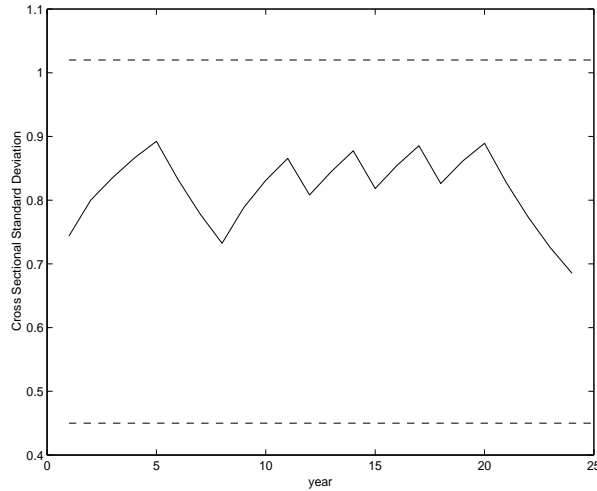
U.S. sample moments are computed using non-overlapping annual returns, (end of) January-over-January, 1956-1996. Estimates of means and standard deviations are qualitatively similar using annual data beginning from 1927, or a monthly series of overlapping annual returns. Equity data correspond to the annual return on the CRSP value weighted index, inclusive of distributions. Riskfree returns are based on the one month U.S. treasury bill. Nominal returns are deflated using the GDP deflator. All returns are expressed as annual percentages. Unlevered equity returns are computed using a debt to firm value ratio of 40 percent, which is taken from Graham (2000).

Figure 1
Cross-Sectional Variance of Earnings



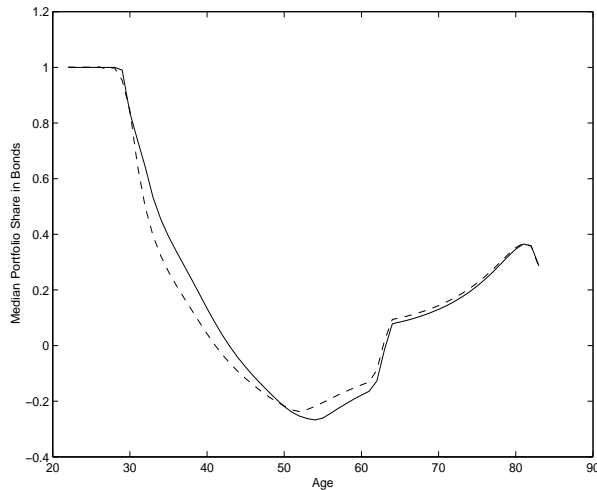
Data are taken from the Panel Study on Income Dynamics (PSID), 1968-1991. Panel A reports the cross-sectional variance by age, where ‘cohort effects’ have been removed using a dummy variable regression as in Deaton and Paxson (1994) and Storesletten, Telmer, and Yaron (2000a). Panel B reports the (linearly detrended) cross-sectional mean and the cross-sectional standard deviation of PSID earnings, for the years 1969-1992. The standard deviation is additively scaled for graphical reasons. The correlation coefficient between the two series is -0.85 . Panel B is robust to (i) alternative methods of detrending the mean and (ii) using the coefficient of variation instead of the standard deviation. Panel C plots one point per age cohort in our panel. Each point represents the cross-sectional variance for that cohort — net of age effects — versus the fraction of contractionary years during which members of that cohort were of working age.

Figure 2
Simulated Cross-Sectional Standard Deviation



The solid line is one particular realization of the overall (*i.e.*, across all generations) cross-sectional standard deviation of income from the model of Section 4. The upper line is the limit point, should all idiosyncratic shocks be drawn from the high variance (contractionary) distribution. The lower line is the analogous lower limit point. Further details are provided in section 2.4 of the text.

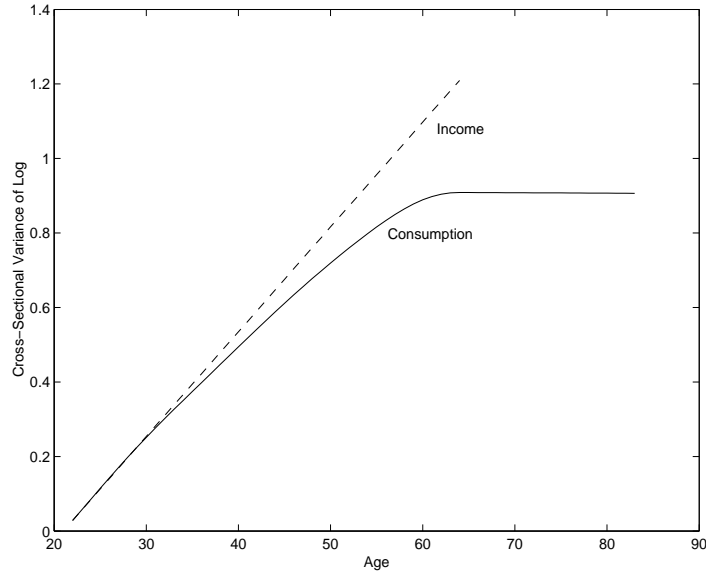
Figure 3
Mean Portfolio Share in Bonds, by Age



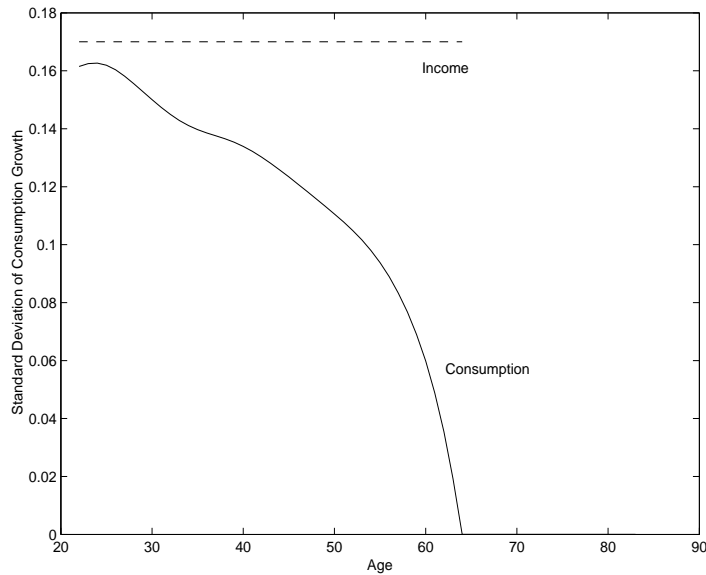
The solid line conditions on aggregate expansions. The dashed line conditions on aggregate contractions.

Figure 4 Risk Sharing Behavior

A: Cross-Sectional Variance of Log Consumption and Income



B: Cross-Sectional Standard Deviation of Growth Rates



Panel A reports the cross-sectional variance of log consumption and labor income from the economy in Section 4. We report the variance, instead of the more natural standard deviation, because it will be linear, given a unit root process (as is the case for income). Panel B reports the standard deviation of the consumption and income growth rates.