THE UNIVERSITY OF MELBOURNE

Department of Economics

Working Paper Series

Aggregate Implications of Micro Asset Market Segmentation

Chris Edmond &
Pierre-Olivier Weill

Jan 2011

Research Paper Number 1117

ISSN: 0819-2642
ISBN: 978 0 7340 4470 9

Department of Economics
The University of Melbourne
Parkville VIC 3010
www.economics.unimelb.edu.au

Aggregate Implications of Micro Asset Market Segmentation*

Chris Edmond[†] Pierre-Olivier Weill[‡]

First draft: March 2009. This draft: January 2011

Abstract

This paper develops a consumption-based asset pricing model to explain and quantify the aggregate implications of a frictional financial system, comprised of many financial markets partially integrated with one another. Each of our micro financial markets is inhabited by traders who are specialized in that market's type of asset. We specify exogenously the level of segmentation that ultimately determines how much idiosyncratic risk traders bear in their micro market and derive aggregate asset pricing implications. We pick segmentation parameters to match facts about systematic and idiosyncratic return volatility. We find that if the same level of segmentation prevails in every market, traders bear 30% of their idiosyncratic risk. With otherwise standard parameters, this benchmark model delivers an unconditional equity premium of 2.4% annual. We further disaggregate the model by allowing the level of segmentation to differ across markets. This version of the model delivers the same aggregate asset pricing implications but with only one-third the amount of segmentation: on average traders bear 10% of their idiosyncratic risk.

Keywords: Asset pricing, market segmentation, idiosyncratic risk.

JEL classifications: G12.

*We thank Andrew Atkeson, David Backus, Alexandre Dmitriev, Xavier Gabaix, Greg Kaplan, Sephorah Mangin, Stijn Van Nieuwerburgh, Gianluca Violante and seminar participants at the ANU, FRB Richmond, FRB Philadelphia, NYU, Ohio State, Sydney, UNSW, UTS, Wharton, Toulouse School of Economics, the LAEF UCSB conference on Financial Frictions and Segmented Asset Markets, the 2010 Australian Conference on Quantitative Macroeconomics, the 2009 Australasian Macroeconomics Workshop, the 2009 Sydney-Melbourne Workshop on Macroeconomic Theory, and the 2007 and 2008 SED annual meetings for helpful comments and conversations. We particularly thank Amit Goyal and Turan Bali for sharing their data with us. Chris Edmond gratefully acknowledges support from the Australian Research Council, grant DP-110103457. Pierre-Olivier Weill gratefully acknowledges support from the National Science Foundation, grant SES-0922338.

[†]University of Melbourne and NBER.

[‡]UCLA and NBER.

1 Introduction

Asset trade occurs in a wide range of security markets and is inhibited by a diverse array of frictions. Upfront transaction costs, asymmetric information between final asset holders and financial intermediaries, and trade in over-the-counter or other decentralized markets that make locating counterparties difficult, all create "limits to arbitrage" (Shleifer and Vishny, 1997). A considerable empirical and theoretical literature on market microstructure has studied these frictions and conclusively finds that "local" factors, specific to the market under consideration, matter for asset prices in that market (see Collin-Dufresne, Goldstein, and Martin, 2001; Gabaix, Krishnamurthy, and Vigneron, 2007, for example). But these market-specific analyses do not give a clear sense of whether micro frictions and local factors matter in the aggregate. Indeed, by focusing exclusively on market-specific determinants of asset prices, these analyses are somewhat disconnected from traditional frictionless consumption-based asset pricing models of Lucas (1978), Breeden (1979) and Mehra and Prescott (1985). Research in that tradition takes the opposite view, that micro asset market frictions and local factors do not matter in the aggregate and that asset prices are determined by broad macroeconomic factors. The truth presumably lies somewhere between these two extremes: asset prices reflect both macro and micro-market specific factors (Cochrane, 2005).

This paper constructs a simple consumption-based asset pricing model in order to explain and quantify the macro impacts of micro market-specific factors. At the heart of our paper is a stylized model of a financial system comprised of a collection of many small micro financial markets that are partially integrated with one another. We strip this model of a fragmented financial system down to a few essential features and borrow some modeling tricks from Lucas (1990) and others to build a tractable aggregate model. In short, we take a deliberately macro approach: we do not address any particular features of any specific asset class but we are able to spell out precisely the aggregate implications of fragmentation and limits-to-arbitrage frictions.

In our benchmark model, there are many micro asset markets. Each market is inhabited by traders specialized in trading a single type of durable risky asset with supply normalized to one. Of course, if the risky assets could be frictionlessly traded across markets all idiosyncratic market-specific risk would be diversified away and each asset trader would be exposed only to aggregate risk. We prevent this full risk sharing by imposing, exogenously, the following pattern of market-specific frictions: we assume that for each market m an exogenous fraction, λ_m, of the expense of purchasing assets in that market must be borne by traders specialized in that market. In return, these traders receive λ_m of the benefit, i.e., of the dividends and resale price of assets sold in that market. We show that, in equilibrium, the parameter λ_m measures the amount of non-tradeable idiosyncratic risk. When $\lambda_m = 0$ all idiosyncratic risk can be traded and traders are fully diversified. When $\lambda_m = 1$ traders cannot trade away their idiosyncratic risk and simply consume the dividends thrown off by the asset in their specific market.

Our theoretical market setup is made tractable by following Lucas (1990) in assuming that investors can pool the *tradeable* idiosyncratic risk within a large family. In equilibrium, the "state price" of a unit of consumption in each market m is a weighted average of the marginal utility of consumption in that market (with weight λ_m) and a term that reflects the cross-sectional average marginal utility of consumption (with weight $1 - \lambda_m$). Generally, both the average level of λ_m and its cross-sectional variation across markets play crucial roles in determining the equilibrium mapping from the state of the economy, as represented by the realized exogenous distribution of dividends across markets, to the endogenous distribution of asset prices across markets. In the special case where $\lambda_m = 0$ for all markets m, the state price of consumption is equal across markets and equal to the marginal utility of the aggregate endowment so that this economy collapses to the standard Lucas (1978) consumption-based asset pricing model. The specification of λ_m, representing the array of micro frictions which impede trade in claims to assets across markets, constitutes our one new degree of freedom relative to a standard consumption-based asset pricing model.

We start by calibrating a special case of the general model where $\lambda_m = \lambda$ for all markets. We choose standard parameters for aggregates and preferences: independently and identically distributed lognormal aggregate endowment growth, time- and state-separable expected utility preferences with constant relative risk aversion $\gamma = 4$. We then use the parameters governing the distribution of individual endowments and the single λ to simultaneously match the *systematic* return volatility of a well-diversified market portfolio and key time-series properties of an individual stock's *total* return volatility (see Goyal and Santa-Clara, 2003; Bali, Cakici, Yan, and Zhang, 2005). This procedure yields segmentation of approximately $\lambda = 0.30$. We find that this model generates a sizeable unconditional equity premium, some 2.4% annual. However, as is familiar from many asset pricing models with expected utility preferences and trend growth, the model has a risk-free rate that is too high and too volatile.

Next, we extend this benchmark model by allowing for multiple types of markets, each with different amounts of segmentation, λ_m, which generates cross-sectional differences in stock return volatilities. We pick values for λ_m in order to match the volatilities of portfolios sorted on measures of idiosyncratic volatility, as documented by Ang, Hodrick, Xing, and Zhang (2006). Our main finding is that *aggregation does* *matter*: with cross-sectional variation in λ_m, we need an *average* amount of segmentation of approximately $\bar{\lambda} = 0.10$ to hit our targets, only one-third that of the single λ model. Moreover, this version of the model delivers essentially the same aggregate asset pricing implications as the single λ benchmark despite having only about one-third the average amount of segmentation. The characteristics of the micro markets in this disaggregate economy are quite distinct: some 50% of the aggregate market by value has a λ_m of approximately zero, with the amount of segmentation rising to a maximum of $\lambda_m = 0.37$ for about 2% of the aggregate market by value.

Market frictions in the asset pricing literature. Traditionally, macroeconomists have taken the view that frictions in financial intermediation or other asset trades are small enough to be neglected in the analysis: asset prices are set "as if" there were no intermediaries but instead a grand Walrasian auction directly between consumers. In particular, early contributions to the literature, such as Rubinstein (1976), Lucas (1978) and Breeden (1979), characterize equilibrium asset prices using frictionless models. The quantitative limitations of plausibly calibrated traditional asset pricing models were highlighted by the "equity premium" and "risk-free rate" puzzles of Mehra and Prescott (1985), Weil (1989) and others.

Since then an extensive literature has attempted to explicitly incorporate market frictions into an asset pricing model in an attempt to rationalize these and related asset pricing puzzles.[1] These models have tended to follow one of two approaches. On the one hand, part of the financial economics literature followed deliberately micro-market approaches, focusing on the impact of specific frictions in specific financial markets. This microfoundations approach is transparent and leads to precise implications but does not lead to any clear sense of whether or why micro asset market frictions matter in the aggregate. Moreover, these models are typically not well integrated with the standard Lucas (1978) consumption-based asset pricing framework. On the other hand are unabashedly aggregate approaches, with some financial friction faced by either some representative intermediary (see, e.g., Aiyagari and Gertler, 1999; Kyle and Xiong, 2001; Vayanos, 2005; He and Krishnamurthy, 2008a,b) or by households (see, among many others Heaton and Lucas, 1996; Chien, Cole, and Lustig, 2011; Pavlova and Rigobon, 2008). The friction "stands in" for a diverse array of real-world micro frictions on the intermediary's and households' side. Since there are large discrepancies between the

[1] See for example Aiyagari and Gertler (1991), He and Modest (1995) and Luttmer (1996, 1999) for the quantitative evaluation of asset pricing models with trading frictions. Other attempts to rationalize asset pricing puzzles retain frictionless markets but depart from traditional models by using novel preference specifications (e.g., Epstein and Zin (1989), Weil (1989, 1990), Campbell and Cochrane (1999)) and/or novel shock processes (e.g., Bansal and Yaron (2004)).

predictions of frictionless asset pricing models and the data, the calibrated friction also tends to have to be large. This approach has the advantage that the friction has macro implications, by construction, but has the disadvantage that the friction has no transparent interpretation. In particular, it is difficult to evaluate the plausibility of the calibrated friction in terms of the constraints facing real-world households and firms.

In these macro models, financial intermediaries often bear disproportionate amounts of *aggregate* risk, but this implication is inconsistent with the empirical literature on market segmentation, which emphasizes instead that intermediaries bear disproportionate amounts of "local" or *idiosyncratic* risk. Our approach takes a middle course. Starting from a model that is consistent with intermediaries bearing too much local risk, we work out the aggregation problem. With the aggregation problem solved, we can then embed our stylized model of a collection of micro markets that together form a financial system into an otherwise standard asset pricing model and examine its quantitative implications. In a sense, our model can be viewed as a multiple market version of a *limited participation* model of asset prices where agents are restricted in their ability to participate in asset trade. Important early contributions to this approach include Makoto (1995), Basak and Cuoco (1998), and Vissing-Jørgensen (2002). State of the art contributions to this literature include Gomes and Michaelides (2008), Guvenen (2009) and Chien, Cole, and Lustig (2011).

The remainder of the paper is organized as follows. In Section 2 we present our model and show how to compute equilibrium asset prices. In Section 3 we calibrate a special case of the model with a single type of market segmentation and in Section 4 we show that this model can generate a sizeable equity premium. Section 5 extends this benchmark model by allowing for multiple types of market segmentation and a non-degenerate cross-section of volatilities. Technical details and several extensions are given in the Appendix.

2 Model

The model is a variant on the pure endowment asset pricing models of Lucas (1978), Breeden (1979) and Mehra and Prescott (1985).

2.1 Setup

Market structure and endowments. Time is discrete and denoted $t \in \{0, 1, 2, ...\}$. A probability space $(\Omega, \mathcal{F}, P)$ is fixed as well as an information filtration $\{\mathcal{F}_t\}_{t=0}^{\infty}$. There are many distinct micro asset markets indexed by $m \in [0,1]$. Each market m is specialized in trading a single type of durable asset with supply normalized to $S_m = 1$. Each period the asset produces a stochastic realization of a non-storable dividend $y_{m,t} > 0$. The aggregate endowment available to the entire economy is:

$$y_t := \int_0^1 y_{m,t} S_m \, dm = \int_0^1 y_{m,t} \, dm.$$

The aggregate endowment $y_t > 0$ follows an exogenous stochastic process that we describe in detail later. Conditional on all realized aggregate variables, the endowments $y_{m,t}$ are independently and identically distributed (IID) across markets.

Preferences. We follow Lucas (1990) and use a representative family construct to provide consumption insurance beyond our market-segmentation frictions. The single representative family, which is initially endowed with the entire supply of assets, consists of many, identical, traders who are specialized in particular asset markets. The period utility for the family views the *utility* of each type of trader as perfectly substitutable:

$$U(c_t) := \int_0^1 u(c_{m,t}) \, dm,$$

where $u : \mathbb{R}^+ \to \mathbb{R}$ is a standard increasing concave utility function. Only in the special case of risk neutrality does the family view the *consumption* of each type of trader as being perfect substitutes. In general, concavity will lead the family to smooth consumption across traders in different markets. Intertemporal utility for the family has the standard time- and state-separable form, $\mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t U(c_t)\right]$, with constant time discount factor $0 < \beta < 1$. The crucial role of the representative family is to eliminate the wealth distribution across markets as an additional endogenous state variable (see, e.g., Alvarez, Atkeson, and Kehoe, 2002).

Segmentation frictions. We interpret the representative family as a partially integrated financial system. Each trader in market m works at a specialized trading desk that deals in the asset specific to that market (Figure 1 illustrates). Traders in market m are assumed to bear an exogenous fraction $\lambda_m \in [0,1]$ of the expense of trading in that market and in return receive λ_m of the benefit. The remaining $1 - \lambda_m$ of the expense and benefit of trading in that market is shared between family members.

More precisely, given segmentation parameter λ_m, the period budget constraint facing a representative trader in market m is:

$$c_{m,t} + \lambda_m p_{m,t} s_{m,t+1} + (1-\lambda_m) T_{t+1} \leq \lambda_m (p_{m,t} + y_{m,t}) s_{m,t} + (1-\lambda_m) T_t, \tag{1}$$

where $p_{m,t}$ is the ex-dividend price of a share in asset m while $s_{m,t}$ and $s_{m,t+1}$ represent share holdings in that asset, and where T_t and T_{t+1} represent, respectively, the cum-dividend value of a *family portfolio* that is brought into the period and the the ex-dividend value of the family portfolio acquired this period.

To understand the family portfolio, observe from the budget constraint (1) that a trader in market m holds directly a number $\lambda_m s_{m,t+1}$ of shares of asset m. The collection of remaining shares, $(1-\lambda_n) s_{n,t+1}$ for all $n \in [0,1]$, is collectively held by all family members in the family portfolio. The expense and benefit of trading the family portfolio is divided among family members in a manner summarized by the two terms $(1-\lambda_m) T_{t+1}$ and $(1-\lambda_m) T_t$ in the budget constraint, (1). Specifically, the term $(1-\lambda_m) T_{t+1}$ on the left-hand side means that the trader in market m is asked to contribute $1-\lambda_m$ of the expense of acquiring the family portfolio. Symmetrically, the term $(1-\lambda_m) T_t$ on the right-hand side means that the trader receives $1-\lambda_m$ of the benefit. Thus, a balanced family budget requires that:

$$\int_0^1 (1-\lambda_m) T_{t+1} \, dm = \int_0^1 (1-\lambda_n) p_{n,t} s_{n,t+1} \, dm.$$

In words, the total value of all family members' contributions to the family portfolio (the left-hand side) has to equal the total asset value of the family portfolio (the right-hand side). Defining $\bar{\lambda} := \int_0^1 \lambda_m \, dm$ we can rewrite this identity as:

$$T_{t+1} = \int_0^1 \frac{1-\lambda_n}{1-\bar{\lambda}} p_{n,t} s_{n,t+1} \, dn. \tag{2}$$

Similarly, $\int_0^1 (1-\lambda_m) T_t \, dm$ is equal to the cum-dividend value of the remaining shares brought into the period. This yields:

$$T_t = \int_0^1 \frac{1-\lambda_n}{1-\bar{\lambda}} (p_{n,t} + y_{n,t}) s_{n,t} \, dn. \tag{3}$$

2.2 Equilibrium asset pricing

A price path is an adapted sequence $p = \{p_t\}_{t=0}^{\infty}$, where $p_t : [0,1] \to \mathbb{R}^+$ is a measurable function mapping each asset $m \in [0,1]$ into its price, $p_{m,t}$, at time t. Given a price path,

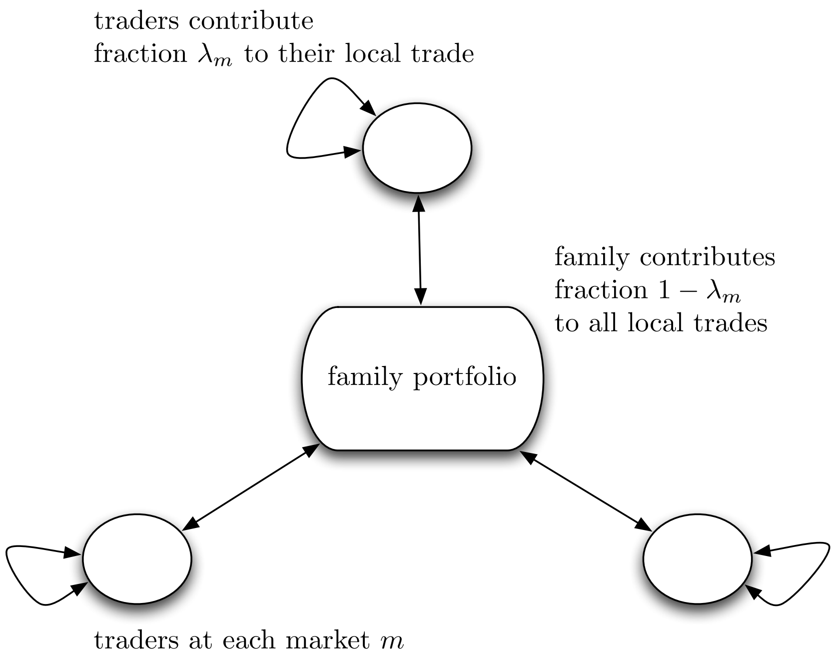

Figure 1: Market structure and segmentation frictions.

There are many markets $m \in [0,1]$. Traders at each market bear fraction λ_m of the expense of their trades and share the remaining fraction $1-\lambda_m$ of the expense with all other traders through a *family portfolio*.

the family maximizes its intertemporal utility by choosing an adapted consumption and asset holding plan, $(c, s') = \{c_t, s_{t+1}\}_{t=0}^{\infty}$, where $c_t : [0,1] \to \mathbb{R}^+$ and $s_{t+1} : [0,1] \to \mathbb{R}$ are measurable functions specifying $c_{m,t}$ and $s_{m,t+1}$ in each asset market $m \in [0,1]$. The maximization is subject to the collection of budget constraints (1), one for each $m \in [0,1]$, the accounting identities for the family portfolio, (2) and (3), and the initial asset holding condition $s_{m,0} = 1$ for all $m \in [0,1]$.

An *equilibrium* of this economy is a consumption and asset holding plan, (c, s'), and a price path, p, such that (i) (c, s') solves the family's problem given p, and (ii) asset markets clear, i.e., $s_{m,t+1} = 1$ for all $m \in [0,1]$ and $t \in \{0, 1, 2, \ldots\}$.

Equilibrium allocation. Before solving for asset prices, we provide the equilibrium allocation of consumption across markets. Substituting the accounting identities (2) and (3) into the budget constraint (1) and imposing the equilibrium condition $s_{m,t+1} = 1$, we obtain:

$$c_{m,t} = \lambda_m y_{m,t} + (1-\lambda_m) \int_0^1 \frac{1-\lambda_n}{1-\bar{\lambda}} y_{n,t} \, dn.$$

Since the realized idiosyncratic endowments $y_{n,t}$ are independent of λ_n, an application of the law of large numbers gives:

$$c_{m,t} = \lambda_m y_{m,t} + (1 - \lambda_m) y_t. \tag{4}$$

This formula is intuitive: equilibrium consumption in market m is a weighted average of the idiosyncratic and aggregate endowments with weights reflecting the degree of market segmentation. The λ_m represent the extent to which traders are not fully diversified and hence the segmentation parameters determine the degree of risk sharing in the economy. If $\lambda_m = 0$, traders are fully diversified and will have consumption equal to the aggregate endowment $c_{m,t} = y_t$ (i.e., full consumption insurance). But if $\lambda_m = 1$, traders are not at all diversified and will simply consume the dividends realized in their specific market $c_{m,t} = y_{m,t}$ (i.e., autarky).

Asset prices. To obtain asset prices, we use the first-order condition of the family's optimization problem. Let $\mu_{m,t} \geq 0$ denote the Lagrange multiplier on (1), the family's budget constraint for market m at time t. We show in Appendix A that the family's Lagrangian can be written:

$$\mathscr{L} = \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \int_0^1 \Big(u(c_{m,t}) + q_{m,t}(p_{m,t} + y_{m,t}) s_{m,t} - q_{m,t} p_{m,t} s_{m,t+1} - \mu_{m,t} c_{m,t} \Big) \, dm \right],$$

where

$$q_{m,t} := \lambda_m \mu_{m,t} + (1 - \lambda_m) \int_0^1 \frac{1 - \lambda_n}{1 - \bar{\lambda}} \mu_{n,t} \, dn \tag{5}$$

is a weighted average of the Lagrange multipliers in market m and the multipliers for other markets with weights reflecting the various degrees of market segmentation. More specifically, $q_{m,t}$ is the marginal value to the family of earning one (real) dollar in market m. The first term in (5) arises because a fraction λ_m goes to the local trader, with marginal utility $\mu_{m,t}$. The second term arises because the remaining fraction is shared among other family members, with marginal utility $\mu_{n,t}$, according to their relative contributions $(1 - \lambda_n)/(1 - \bar{\lambda})$ to the family portfolio. We will refer to $q_{m,t}$ as the *state price* of earning one real dollar in market m.

Just as equilibrium consumption in market m is a weighted average of the idiosyncratic or "local" endowment and aggregate endowment with weights λ_m and $1 - \lambda_m$, so too the state price for market m is a weighted average of the idiosyncratic multiplier and an aggregate multiplier with the same weights. To highlight this, define:

$$q_t := \int_0^1 \frac{1-\lambda_n}{1-\bar{\lambda}} \mu_{n,t} \, dn, \tag{6}$$

so that we can write $q_{m,t} = \lambda_m \mu_{m,t} + (1-\lambda_m) q_t$. If any particular market m has $\lambda_m = 0$, then the state price in that market is equal to the aggregate state price $q_{m,t} = q_t$ and is independent of the local endowment realization. If the segmentation parameter is common across markets, $\lambda_m = \lambda$ all m, then q_t is the cross-sectional *average marginal utility* and $q_t = \int_0^1 q_{m,t} dm$. More generally, q_t is not a simple average over $\mu_{m,t}$ since different markets have different relative contributions $(1-\lambda_m)/(1-\bar{\lambda})$ to the family portfolio.

The first-order conditions of the family's problem are straightforward. For each $c_{m,t}$ we have:

$$u'(c_{m,t}) = \mu_{m,t}. \tag{7}$$

After taking derivatives with respect to $s_{m,t+1}$ we find the Euler equation:

$$p_{m,t} = \mathbb{E}_t \left[\beta \frac{q_{m,t+1}}{q_{m,t}} (p_{m,t+1} + y_{m,t+1}) \right], \tag{8}$$

where the expectation is conditional on the family's information at time t. This is a standard equation, familiar from Lucas (1978), with the crucial distinction being that the stochastic discount factor (SDF), $\beta q_{m,t+1}/q_{m,t}$, is market specific.

Note that, combining the formulas for equilibrium consumption (4), market-specific state prices (5), and the pricing equation (8), we obtain a mapping from the primitives of the economy (the λ_m, $y_{m,t}$ etc) into equilibrium asset prices. In particular, it is easy to verify that the Lucas (1978) asset prices are obtained in the special case where $\lambda_m = 0$ for all m, so that $c_{m,t} = y_t$ and $\mu_{m,t} = u'(y_t)$ for all m and $q_t = \int_0^1 u'(y_t) \, dn = u'(y_t)$.

Shadow prices of risk-free bonds. To simplify the presentation of the model, we have not explicitly introduced risk-free assets. But we can still compute "shadow" bond prices under the following convention. Let $\pi_{k,t}$ denote the price at time t of a zero-coupon bond that pays one unit of the consumption good for sure at time $t+k \geq 1$, and that is held in the family portfolio. In Appendix A we show that, under appropriate conditions, these bonds would have the price:

$$\pi_{k,t} = \mathbb{E}_t \left[\beta \frac{q_{t+1}}{q_t} \pi_{k-1,t+1} \right], \tag{9}$$

with $\pi_{0,t} := 1$. Bonds are priced using the aggregate state price, q_t. In particular, the one-period shadow gross risk-free rate is given by $1/\pi_{1,t} = 1/\mathbb{E}_t\left[\beta q_{t+1}/q_t\right]$. Although the one-period SDF for bonds, $\beta q_{t+1}/q_t$, does not depend on any particular idiosyncratic endowment realization, it does depend on the *distribution* of idiosyncratic endowments and in general is *not* simply the Lucas-Breeden SDF.

3 Calibration

To evaluate the significance of these segmentation frictions, we calibrate the model.

3.1 Parameterization of the model

Preferences and endowments. Let period utility $u(c)$ be of constant relative risk aversion (CRRA) with coefficient $\gamma > 0$ so that $u'(c) = c^{-\gamma}$. The log aggregate endowment is a random walk with drift and IID normal innovations:

$$\log g_{t+1} := \log(y_{t+1}/y_t) = \log \bar{g} + \epsilon_{g,t+1}, \quad \epsilon_{g,t+1} \sim \text{IID and } N(0, \sigma^2_{\epsilon g}), \quad \bar{g} > 0. \tag{10}$$

Log market-specific endowments are the log aggregate endowment plus an idiosyncratic term:

$$\log y_{m,t} := \log y_t + \log \hat{y}_{m,t}, \tag{11}$$

so that market-specific endowments inherit the trend in the aggregate endowment. The log idiosyncratic endowment $\log \hat{y}_{m,t}$ is conditionally IID normal in the cross-section:

$$\log \hat{y}_{m,t} \sim \text{IID and } N(-\sigma_t^2/2, \sigma_t^2), \tag{12}$$

where σ_t follows the stochastic process specified below. The mean log idiosyncratic endowment, $-\sigma_t^2/2$ is chosen so that the cross-sectional average in level is normalized to one, i.e. $\int_0^1 \hat{y}_{m,t}\, dm = 1$.

Idiosyncratic endowment volatility. The cross-sectional standard deviation of the log idiosyncratic endowment, σ_t, is stochastically varying. We specify it as an AR(1) process in logs:

$$\log \sigma_{t+1} = (1-\phi)\log \bar{\sigma} + \phi \log \sigma_t + \epsilon_{v,t+1}, \quad \epsilon_{v,t+1} \sim \text{IID and } N(0, \sigma^2_{\epsilon v}), \quad \bar{\sigma} > 0. \tag{13}$$

For short, we refer to σ_t as *idiosyncratic endowment volatility*, but note that σ_t itself is an *aggregate* state variable. At any point in time, the idiosyncratic endowment volatility, σ_t, is the same in all markets m. In a frictionless model ($\lambda_m = 0$ all m), all idiosyncratic risk would be diversified away so that asset prices would be independent of σ_t. In other words, despite aggregate fluctuations in the level of idiosyncratic endowment volatility, σ_t, it would not be a priced factor. With segmentation frictions ($\lambda_m > 0$), by contrast, both the level and dynamics of σ_t will affect asset prices.

Information. For all that follows, we take family information at time t to be the filtration $\{\mathcal{F}_t\}_{t=0}^{\infty}$ generated by the volatility and the aggregate and idiosyncratic endowment processes.

Solving the quantitative model. Since $\hat{y}_{m,t} = y_{m,t}/y_t$, we can use equation (4) to write equilibrium consumption in market m as the product of the aggregate endowment, y_t, and an idiosyncratic component that depends only on the local idiosyncratic endowment realization, $\hat{y}_{m,t}$, and the amount of segmentation:

$$c_{m,t} = [1 + \lambda_m(\hat{y}_{m,t} - 1)]y_t. \tag{14}$$

Similarly, we can then use this expression for consumption and the fact that utility is CRRA to write the local state price as:

$$q_{m,t} = \theta_{m,t} y_t^{-\gamma}, \tag{15}$$

where:

$$\theta_{m,t} := \lambda_m[1 + \lambda_m(\hat{y}_{m,t} - 1)]^{-\gamma} + (1 - \lambda_m) \int_0^1 \frac{1 - \lambda_n}{1 - \bar{\lambda}}[1 + \lambda_n(\hat{y}_{n,t} - 1)]^{-\gamma}\, dn. \tag{16}$$

The SDF for market m is then:

$$M_{m,t+1} := \beta g_{t+1}^{-\gamma} \frac{\theta_{m,t+1}}{\theta_{m,t}}. \tag{17}$$

This is the usual Lucas-Breeden aggregate SDF $M_{t+1} := \beta g_{t+1}^{-\gamma}$ with a market-specific multiplicative "twisting" factor $\hat{M}_{m,t+1} := \theta_{m,t+1}/\theta_{m,t}$ that adjusts the SDF to account for idiosyncratic endowment risk.

To solve the model in stationary variables, let $\hat{p}_{m,t} := p_{m,t}/y_t$ denote the price-to-aggregate-dividend ratio for market m. Dividing both sides of equation (8) by $y_t > 0$

and using $g_{t+1} := y_{t+1}/y_t$, this ratio solves the Euler equation:

$$\hat{p}_{m,t} = \mathbb{E}_t \left[\beta g_{t+1}^{1-\gamma} \frac{\theta_{m,t+1}}{\theta_{m,t}} (\hat{p}_{m,t+1} + \hat{y}_{m,t+1}) \right], \tag{18}$$

which is the standard CRRA formula except for the multiplicative adjustment $\theta_{m,t+1}/\theta_{m,t}$. This is a linear integral equation to be solved for the unknown function mapping the state into the price/dividend ratio. In general this integral equation cannot be solved in closed form, but numerical solutions can be obtained in a straightforward manner along the lines of Tauchen and Hussey (1991).[2] We discuss these methods in greater detail in Appendix B and Appendix C below.

3.2 Calibration strategy and results

We calibrate the model using monthly postwar data (1959:1-2007:12, unless otherwise noted). Following a long tradition in the consumption-based asset pricing literature, we interpret the aggregate endowment as per capita real personal consumption expenditure on nondurables and services. We set $\bar{g} = (1.02)^{1/12}$ to match an annual 2% growth rate and $\sigma_{\epsilon g} = 0.01/\sqrt{12}$ to match an annual 1% standard deviation over the postwar sample. We set $\beta = (0.99)^{1/12}$ to reflect an annual pure rate of time preference of 1% and we set the coefficient of relative risk aversion to $\gamma = 4$.

For our benchmark calibration we assume that all markets in the economy share the same segmentation parameter, $\lambda = \lambda_m$ for all m. Given the values for preference parameters β, γ and the aggregate endowment growth process $\bar{g}, \sigma_{\epsilon g}$ above, we still need to assign values to this single λ and the three parameters of the cross-sectional endowment volatility process $\bar{\sigma}, \phi, \sigma_{\epsilon v}$.

Calibrating the idiosyncratic volatility process. The crucial consequence of market segmentation is that local traders are forced to bear some idiosyncratic risk. Thus, to explain the impact of market segmentation on risk premia, it is important that our model generates realistic levels of idiosyncratic risk. This leads us to choose the parameters of the stochastic process for idiosyncratic endowment volatility in order to match key features of the volatility of a typical stock return. To see why there is a

[2] As noted by Flodèn (2008), Kopecky and Suen (2010) and others, the Tauchen and Hussey method provides an accurate approximation so long as the underlying stochastic process is not too persistent, which turns out to be the case in our calibration. As a robustness check, Appendix C shows that our calibration results are unchanged when improving accuracy in two ways: using a larger number of quadrature nodes for the log volatility process, and choosing quadrature nodes and weight according to the modified Tauchen and Hussey formula advocated by Flodèn.

natural mapping between the two volatilities, write the gross return on stock m as:

$$R_{m,t} = g_t \frac{\hat{y}_{m,t} + \hat{p}_{m,t}}{\hat{p}_{m,t-1}}. \tag{19}$$

Thus, the volatility of $\hat{y}_{m,t}$ directly affects stock returns through the dividend term of the numerator. It also indirectly affects stock returns through the asset price, $\hat{p}_{m,t}$.

We obtain key statistics about stock return volatility from Goyal and Santa-Clara (2003). Their measure of monthly stock volatility is obtained by adding up the cross-sectional stock return dispersion over each day of the previous month.[3] In Figure 2 we show the monthly time series (1963:1-2001:12) of their measure of the cross-sectional standard deviation of stock returns, as updated by Bali, Cakici, Yan, and Zhang (2005).

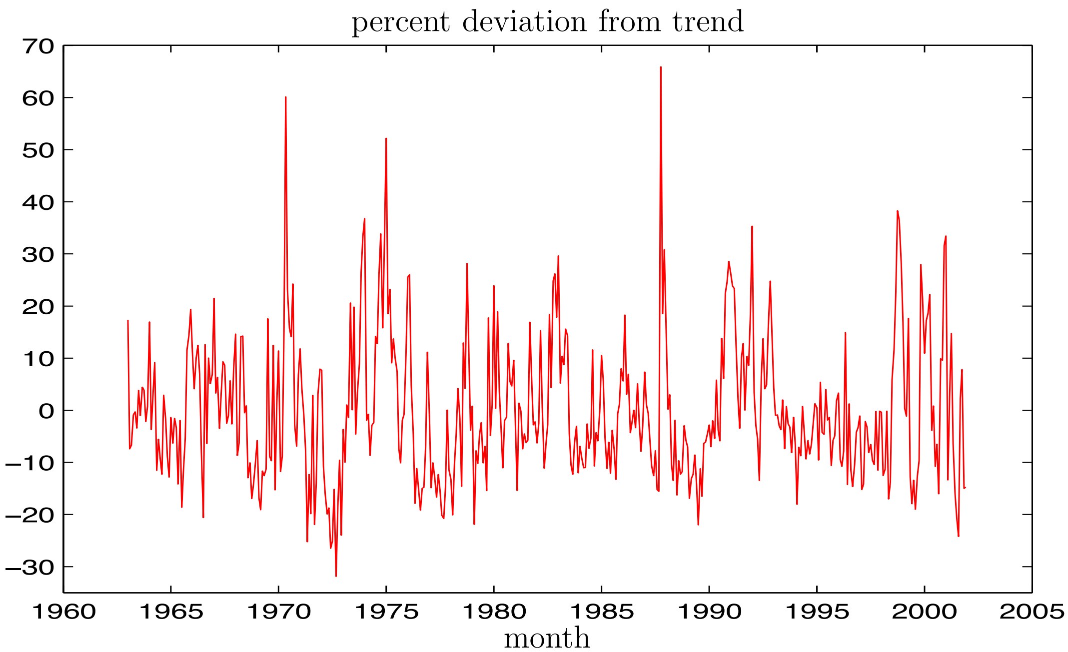

Figure 2: Cross-sectional standard deviation of stock returns, deviation from trend.

Hodrick-Prescott filtered cross-sectional standard deviation of stock returns, monthly (1963:1-2001:12). Data from Goyal and Santa-Clara (2003), updated by Bali, Cakici, Yan, and Zhang (2005).

We choose the idiosyncratic endowment volatility process so that, when we calculate the same stock return volatility measure in our model, we replicate three key features of this data: the unconditional average return volatility of 16.4% monthly, the uncon-

[3]Because they don't subtract the mean return, Goyal and Santa-Clara recognize that their measure of dispersion is biased upward. However, they note (page 979) that for "short holding periods, the impact of subtracting the means is minimal. Even for monthly excess returns, the expected squared return overstates the variance by less than one percent of its level. The advantage of this approach, of course, is that it allows us to sidestep the issue of estimating the mean stock returns".

ditional standard deviation of return volatility 4.17% monthly, and AR(1) coefficient of return volatility 0.84 monthly. We replicate these three features by simultaneously choosing the three parameters governing the stochastic process for endowment volatility: the unconditional average $\bar{\sigma}$, the innovation standard deviation $\sigma_{\epsilon v}$, and the AR(1) coefficient ϕ.

Calibrating the segmentation parameter. The segmentation parameter λ governs the extent to which local traders can diversify away the return volatility of their local asset. Thus, λ determines the extent to which the volatility factor, σ_t, has an impact on asset prices and creates systematic variation in asset returns. This leads us to identify λ using a measure of systematic volatility, specifically the 4.16% monthly standard deviation of the real value-weighted return of NYSE stocks from CRSP.

To understand precisely how the identification works, recall first what would happen in the absence of market segmentation, $\lambda = 0$. Then, we would be back in the Lucas-Mehra and Prescott model with IID lognormal aggregate endowment growth. As is well known (see, e.g., LeRoy, 2006), this model can't generate realistic amounts of systematic volatility. Specifically, with $\lambda = 0$ the return from a diversified market portfolio is $g_t(1+\bar{p})/\bar{p}$ where $\bar{p} = \beta\mathbb{E}[g^{1-\gamma}]/(1-\beta\mathbb{E}[g^{1-\gamma}])$ is the constant price/dividend ratio for the aggregate market. With our standard parameterization of the preference parameters and aggregate endowment growth, $(1+\bar{p})/\bar{p} \approx 1.0058$ so that the monthly standard deviation of the diversified market portfolio return is approximately the same as the monthly standard deviation of aggregate endowment growth, about 0.29% monthly as opposed to 4.16% monthly in the data.

In contrast with the $\lambda = 0$ case, when $\lambda > 0$ idiosyncratic endowment volatility creates systematic volatility. Indeed, because of persistence, a high idiosyncratic endowment volatility this month predicts a high idiosyncratic endowment volatility next month. Thus in every market m local traders expect to bear more idiosyncratic risk, and because of risk aversion the price/dividend ratio $\hat{p}_{m,t} = p_{m,t}/y_t$ has to go down everywhere. Because this effect impacts all stocks at the same time, it endogenously *creates* systematic return volatility. Clearly, the effect is greater if markets are more segmented and traders are forced to bear more idiosyncratic risk: a larger λ will result in a larger increase in systematic volatility.

Calibration results. The calibrated parameters are listed in Table 1. In our benchmark calibration, the level of λ is 0.31. That is, 31% of idiosyncratic endowment risk is non-tradeable. In terms of portfolio weights, we also find that $\lambda = 0.31$ implies that, in a typical market m, a trader invests approximately 31% of his wealth in the local asset and the rest in the family portfolio.[4] Table 2 shows that with these parameters, the benchmark model matches the target moments exactly. The calibrated parameters of the cross-sectional endowment volatility process are an unconditional average of $\bar{\sigma} = 0.32$, an AR(1) coefficient of $\phi = 0.78$ monthly, and an innovation standard deviation of $\sigma_{\epsilon v} = 0.21$ monthly.

4 Quantitative examples

Aggregate statistics. In comparing our model to data, we adopt the perspective of an econometrician who observes a collection of asset returns but who ignores the possibility of market segmentation. In particular, to make our results comparable to those typically reported in the empirical asset pricing literature, we focus on the properties of the aggregate market portfolio. We define the gross market return $R_{M,t+1} := (p_{t+1} + y_{t+1})/p_t$ where $p_t := \int_0^1 p_{m,t}\, dm$ is the *ex-dividend* value of the market portfolio and y_t is the aggregate endowment. We define the shadow gross one-period risk-free rate by $R_{f,t} := \mathbb{E}_t[\beta q_{t+1}/q_t]^{-1}$ where q_t is the aggregate state price that determines the price of risk-free bonds, as in (9). Implicitly, bonds are priced as if they are traded in their own frictionless "$\lambda = 0$" market, but the pricing of such bonds takes into account $\lambda > 0$ in other asset markets.[5]

We then calculate the *unconditional* equity risk premium $\mathbb{E}[R_{M,t+1} - R_{f,t}]$ and similarly for other statistics. We call p_t/y_t the price/dividend ratio of the market.

4.1 Results

Equity premium. With these definitions in mind, Table 3 shows our model's implications for aggregate returns and price/dividend ratios. We report annualized monthly statistics from the model and compare these to annualized monthly returns and to annual price/dividend ratios. (We use annual data for price/dividends because of the pronounced seasonality in dividends at the monthly frequency.) The benchmark model produces an annual equity risk premium of 2.4% annual as opposed to about 5.4% annual in the postwar NYSE CRSP data. Clearly this is a much larger equity premium than is produced by a standard Lucas-Mehra and Prescott model. For comparison, that model with risk aversion $\gamma = 4$ and IID consumption growth with annual standard

[4] The derivation of portfolio weights for the family is given in Appendix A below.

[5] An alternative approach would be to define a market-specific risk-free rate $R_{f,m,t} := \mathbb{E}_t[\beta q_{m,t+1}/q_{m,t}]^{-1}$ for market m and then an average risk-free rate by $R_{f,t} := \int_0^1 R_{f,m,t}\, dm$. In this interpretation, $R_{f,m,t}$ measures the local real risk free opportunity cost of funds in market m. In our benchmark with $\lambda_m = \lambda$ for all m, both approaches give the same average risk-free rate $R_{f,t}$.

deviation of 1% produces an annual equity premium of about 0.04%. The segmented markets model with $\lambda = 0.31$ is able to generate an equity premium some two orders of magnitude larger.

Why is there a large equity premium? Relative to standard consumption-based asset pricing models with time-separable expected utility preferences, our model delivers a large equity premium. Is it a direct consequence of our strategy of picking λ in order to match systematic return volatility? No, since model risk premia are generated by *covariances*: no matter how much return volatility you feed into a model, the equity risk premia will be zero if the model's SDF is not negatively correlated with return variation (see Cochrane and Hansen, 1992, for a forceful argument).

What, then, is the equity premium from the point of view of aggregate consumption? In our model, if one computes the unconditional average equity premium using the model generated market return and the Lucas-Breeden SDF $\beta g_{t+1}^{-\gamma}$ instead of the true model SDF, the equity premium is on the order of 0.04% (4 basis points) annual rather than the 2.4% annual in the benchmark model. While aggregate consumption does not command a big risk premium, the volatility factor does. To see this, consider the premium implied by the SDF, $\beta q_{t+1}/q_t$, where q_t is the aggregate state price that determines the (shadow) price of risk-free bonds. In general this is given by equation (6) but with a single common λ it reduces to $q_t = \int_0^1 \mu_{m,t}\, dm = \int_0^1 c_{m,t}^{-\gamma}\, dm$, the cross-sectional average marginal utility. In our benchmark model, this SDF implies an equity premium of 2.05% annual. This comes from the convexity of the marginal utility function: a high realization of σ_t makes equilibrium consumption highly dispersed across markets so that average marginal utilities are high. At the same time, a high σ_t depresses asset prices in every market, so the return on the market portfolio is low.

Level of the risk-free rate. Although the benchmark model delivers reasonable implications for the level of the equity premium, it is not so successful on other dimensions. The level of the risk-free rate is very high as compared to the data. In the model the risk free rate is about 8% annual while in the data it is more like 2%. As emphasized by Weil (1989), this is a common problem for models with expected-utility preferences. In short, attempts to address the equity premium puzzle by increasing risk aversion also tend to raise the risk-free rate so that even if it's possible to match the equity *premium*, the model may well do so at absolute *levels* of returns that are too high. This effect comes from the relationship between real interest rates and growth in a deterministic setting with expected utility: high risk aversion means low intertemporal elasticity of substitution so that it takes high real interest rates to compensate for high aggregate growth. With risk, there is an offsetting precautionary savings effect that could, in principle, pull the risk-free rate back down to more realistic levels. But in our calibration this precautionary savings effect is quantitatively small: raising λ from zero (the Mehra and Prescott case) to $\lambda = 0.31$ (our benchmark) lowers the risk-free rate by about 1% annual.

Volatility of the risk-free rate. In the data, the risk-free rate is smooth and the volatility of the equity premium reflects the volatility of equity returns. In the benchmark model, the risk-free rate is too volatile, about 5.6% annual as opposed to 1.2% annual in the data.

Yield curve. With IID lognormal aggregate growth and CRRA utility, the average yield curve in a standard asset pricing model is flat. By contrast, our model generates an increasing and concave average yield curve, as shown in Figure 3. This shape comes from the relationship between the aggregate state price q_t and aggregate volatility σ_t. Since σ_t has positive serial correlation but is not a random walk, its first difference is negatively serially correlated. This negative serial correlation is inherited by the one-period bond pricing SDF, $\beta q_{t+1}/q_t$, and, as is well known, this has the desirable implication that the average yield curve is upward sloping (see Backus and Zin, 1994, for example).

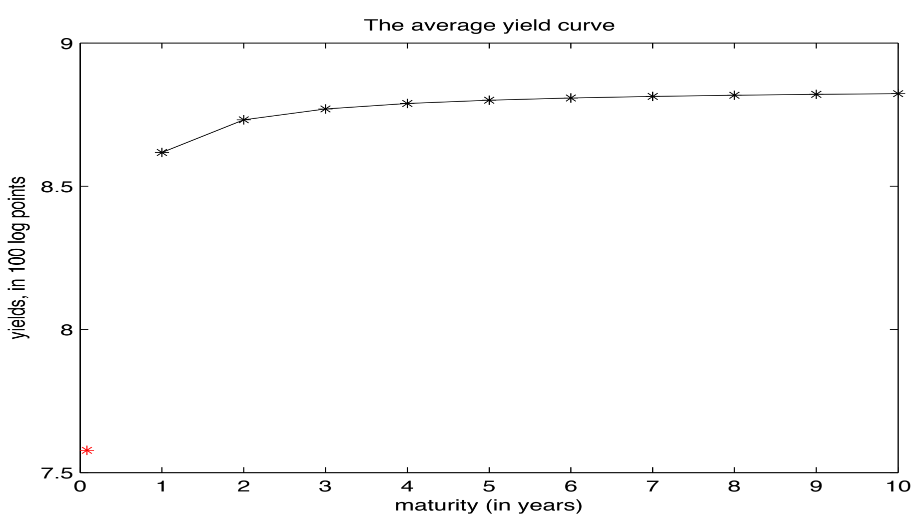

Figure 3: Average yield curve for the benchmark model.

The star point on the left is the average yield on a one-month zero coupon bond, $12\mathbb{E}\left[\log(R_f)\right]$. Note that, because the risk free rate is so volatile and because $\log(\cdot)$ is concave, this yield turns out to be about 1% lower than the average risk free rate reported in Table 3.

Price/dividend ratio. The benchmark model produces an annual price/dividend ratio of about 14 as opposed to an unconditional average of more like 34 in the NYSE CRSP data. Given the large, persistent swings in the price/dividend ratio in the data, what constitutes success on this dimension is not entirely clear. The model generates too little unconditional volatility in the log price/dividend ratio, some 21% annual as opposed to 39% in the data. Also, the temporal composition of price/dividend volatility differs somewhat between the model and the data. In particular, the log price/dividend ratio has approximately no persistence at annual frequencies, $0.76^{12} = 0.04$ in the benchmark model, which is considerably lower than the corresponding annual persistence $0.99^{12} = 0.89$ in the data. That is, the unconditional volatility of the price/dividend ratio in the data comes from large, low-frequency movements whereas the unconditional volatility in the model comes from high-frequency movements.

4.2 Discussion

Constant endowment volatility. Our benchmark model has two departures from a standard consumption-based asset pricing model: segmentation and a time-varying endowment volatility. To show that both these departures are essential for our results, we solved our model with constant endowment volatility, i.e., $\sigma_t = \sigma$ for all t. For this exercise, we fix the volatility at the same level as the unconditional average from the benchmark model, $\bar{\sigma} = 0.32$, and keep the level of segmentation at the benchmark $\lambda = 0.31$. In Table 2 we show that this "constant σ" version of the model produces essentially the same amount of unconditional cross-sectional stock return volatility as in the data (suggesting that this moment is principally determined by $\bar{\sigma}$ alone), but produces relatively little systematic stock volatility. In particular, systematic stock volatility is only about 1% monthly as opposed to 4% in the data. And recall that, for our preference and aggregate growth parameters, a standard Lucas-Mehra and Prescott model would imply negligible systematic stock volatility. Thus $\lambda > 0$ is necessary but not sufficient for our model to create systematic stock volatility from idiosyncratic endowment volatility. In Table 3 we see that the model with constant σ generates an equity premium of about 0.2% annual, an order of magnitude more than in a standard model. Time variation in endowment volatility increases the equity premium further, generating another 2.2% on top of the constant σ case, for a total of 2.4% annual.

Counter-cyclical endowment volatility. Many measures of cross-sectional idiosyncratic risk increase in recessions (see, for example Campbell, Lettau, Malkiel, and Xu, 2001; Storesletten, Telmer, and Yaron, 2004). This counter-cyclicality is also a feature of the cross-sectional standard deviation of returns data from Goyal and Santa-Clara (2003). However, the stochastic process we use for the cross-sectional volatility evolves independently of aggregate growth. To see if our results are sensitive to this, we modify the stochastic process in (13) to:

$$\log \sigma_{t+1} = (1 - \phi) \log \bar{\sigma} + \phi \log \sigma_t - \eta(\log g_t - \log \bar{g}) + \epsilon_{v,t+1} \tag{20}$$

with $\epsilon_{v,t+1}$ IID normal, as before. If $\eta > 0$, then aggregate growth below trend in period t increases the likelihood that volatility is above trend in period $t + 1$ so that volatility is counter-cyclical. We identify the new parameter η by requiring that, in a monthly regression of the cross-sectional standard deviation of stock returns on *lagged* aggregate growth, the regression coefficient is -0.56 as it is in the data. The calibrated parameters for this "feedback" version of the model are also shown in Table 1. The calibrated elasticity, η, is 2.5 so aggregate growth 1% below trend tends to increase endowment volatility by 2.5%. The other calibrated parameters are indistinguishable from the benchmark parameters. Moreover, the model's implications for asset prices as shown in Table 3 are also very close to the results for the benchmark model. This suggests that, while the model can be reconciled with the counter-cyclical behavior of cross-sectional stock volatility, this feature is not necessary for our main results.

Relationship to incomplete markets models. There is a large literature in macroeconomics that analyzes the asset pricing implications of market incompleteness when households receive uninsurable idiosyncratic income shocks.[6] One might have the impression that all our model does is shift the focus of incomplete markets models: instead of analyzing the idiosyncratic labor income risk facing households, all we do is analyze the idiosyncratic income risk faced by the financial sector. We now argue that, while our segmented markets model indeed results in uninsurable shocks, it is conceptually different from standard incomplete markets models.

To see why, note that in standard incomplete markets models the intertemporal marginal rate of substitution (IMRS) of *every* household i prices the excess return of the market portfolio:

$$\mathbb{E}\left[M_i R^e\right] = 0. \tag{21}$$

As forcefully emphasized by Mankiw (1986), Constantinides and Duffie (1996) and Kruger and Lustig (2010), idiosyncratic risk has no impact on the equity premium when utility is CRRA and idiosyncratic consumption growth is statistically indepen-

[6] See Telmer (1993) and Heaton and Lucas (1996) for important early examples.

dent from aggregate consumption growth. Indeed, in that case the IMRS can be factored into $\hat{M}_i M$, where $M = \beta g^{-\gamma}$ is the standard Lucas-Breeden SDF, and $\hat{M}_i$ is an idiosyncratic component that is independent from M. Expanding the expectation in (21) we have:

$$\mathbb{E}[\hat{M}_i M R^e] = \mathbb{E}[\hat{M}_i]\mathbb{E}[M R^e] + \text{Cov}[\hat{M}_i, M R^e] = 0.$$

From independence, $\text{Cov}[\hat{M}_i, M R^e] = 0$. Using this and dividing both sides by $\mathbb{E}[\hat{M}_i] > 0$ we obtain:

$$\mathbb{E}\left[M R^e\right] = 0.$$

As shown by Kocherlakota (1996), this asset pricing equation cannot rationalize the observed equity premium.

In our benchmark model we maintain the assumption that the idiosyncratic component of dividends, $\hat{y}_m$, is statistically independent from aggregate consumption growth. Despite this, we obtain a much larger equity premium than Mehra and Prescott (1985). The reason is that in our asset pricing model the local SDF does not have to price the excess return on the *aggregate* market portfolio, as in equation (21), but instead price the excess return on the *local* asset market. The local discount factor is correlated with the local excess return (through the local endowment realization) and this makes it impossible to strip out the influence of the market-specific factor.

Specifically, instead of equation (21) we have a pricing equation of the form:

$$\mathbb{E}\left[M_m R_m^e\right] = 0, \tag{22}$$

where M_m is the local SDF and R_m^e is the local excess return. From equation (17) we can factor the local discount factor into $\hat{M}_m M$ where M is again the Lucas-Breeden discount factor and $\hat{M}_m$ is a market-specific factor. Now proceeding as above and expanding the expectation in (22) we have:

$$\mathbb{E}[\hat{M}_m M R_m^e] = \mathbb{E}[\hat{M}_m]\mathbb{E}[M R_m^e] + \text{Cov}[\hat{M}_m, M R_m^e] = 0.$$

But $\hat{M}_m$ and R_m^e depend on the same local risk factor so $\text{Cov}[\hat{M}_m, M R_m^e] \neq 0$ and we cannot factor out $\mathbb{E}[\hat{M}_m]$. Because this makes it impossible to aggregate the collection of equations (22) into (21), the standard incomplete markets logic does not apply in our model.

5 Cross-sectional volatilities

In our first set of quantitative examples we used a common amount of segmentation, λ, for all asset markets. This implies that, conditional on the aggregate state of the economy, each market m is characterized by a common amount of volatility (essentially determined by the economy-wide σ_t and λ) so that there is no *cross-sectional variation in volatility*. We now pursue the implications of the general model with market-specific λ_m and hence a non-degenerate cross-section of volatility.

Specifically, we allow for a finite number N of market *types*. In a slight abuse of notation we continue to index these market types by $m \in \{1, ..., N\}$. We assume that each market contains the same number of assets. There is a total measure ω_m of traders in market m, with $\sum_{m=1}^{N} \omega_m = 1$. The supply of assets *per trader* is normalized to 1. With this notation, then, the aggregate endowment is $y = \sum_{m=1}^{N} y_m \omega_m$.

Calibration of market-specific λ_m: strategy. In the case of a single common λ above, the value of λ was identified by matching a measure of systematic volatility, the return volatility of a well-diversified portfolio of stocks. We now need to identify a *vector* of N segmentation parameters and we do this using a closely related strategy. In particular, we identify market types with quintile portfolios of stocks sorted on measures of idiosyncratic volatility from Ang, Hodrick, Xing, and Zhang (2006). They compute value-weighted quintile portfolios by sorting stocks based on idiosyncratic volatility relative to the Fama and French (1993) three-factor pricing model in postwar CRSP data. To give a sense of this data, Ang, Hodrick, Xing, and Zhang report an average standard deviation of (diversified) portfolio returns for the *first* quintile of stocks of 3.8% monthly (as opposed to 4.2% for the market as a whole). By construction this portfolio is 20% of a simple count of stocks but it constitutes 54% of the market by value. At the other end of the volatility spectrum, the average monthly standard deviation of a well-diversified portfolio of the *fifth* quintile of stocks is 8.2% and these constitute only about 2% of the market by value.

We choose the value of λ_m for each $m \in \{1, ..., 5\}$ to match the total volatility of the m'th quintile *portfolio* in Ang, Hodrick, Xing, and Zhang. Similarly, we choose values of ω_m so that the unconditional average portfolio weight of the family in assets of market m matches the average market share for the m'th quintile portfolio from Ang, Hodrick, Xing, and Zhang. Our calibration procedure chooses these parameters simultaneously with the parameters $\bar{\sigma}, \phi, \sigma_{\epsilon v}$ of the stochastic process for cross-sectional endowment volatility. We keep the values of the preference parameters β, γ and the aggregate growth parameters $\bar{g}, \sigma_{\epsilon g}$ at their benchmark values. If there was only one market type, then this calibration procedure would coincide with the procedure used for our benchmark model above.

Calibration of market-specific λ_m: results. The calibrated parameters from this procedure are listed in Panel A of Table 4. We find that the $m = 1$ market, with the lowest idiosyncratic volatility, has a segmentation parameter $\lambda_1 = 0.01$. These assets are close to frictionless. This $m = 1$ market consists of 20% of assets by number, by construction, but it accounts for 51% of the total market by value. By contrast, the $m = 5$ market has segmentation parameter $\lambda_5 = 0.37$ but accounts for only 2% of total market value. Across markets, the segmentation parameters, λ_m, are monotonically increasing in m while the weights, ω_m, are monotonically decreasing in m. Averaging over the five markets, $\bar{\lambda} = \sum_{m=1}^{5} \lambda_m \omega_m = 0.115$. Thus this economy, which matches the same aggregate moments as the benchmark model, hits its targets with an average amount of segmentation $\bar{\lambda} = 0.115$, roughly one-third that of the single parameter benchmark $\lambda = 0.31$. This suggests that there may be a significant bias when aggregating a collection of heterogeneously segmented markets into a "representative" segmented market.

Relative to the benchmark, the model's endowment volatility process now has a substantially higher unconditional average, approximately the same time-series variation, and more persistence. Table 4 shows that with these parameters the model matches the target moments closely but not exactly. In particular, while the average idiosyncratic volatility across the 5 markets is about the same as in the data, this is achieved with slight discrepancies at the level of each market, e.g., the 1st market has volatility of 4.2% against 3.8% in the data.

Market-specific asset pricing implications. In Panel A of Table 5 we show the risk premia for each market type in the model and their empirical counterparts. In the data, the equity premium for the low volatility $m = 1$ market is 0.53% monthly (roughly 6.5% annual) whereas in the model it is 0.17% monthly. This market accounts for half of total market value. As we go to markets with higher volatility, the model predicts that risk premia *monotonically* increase, reaching 0.53% monthly for the $m = 5$ market. However, the data predicts a *hump-shaped* pattern for the cross-section of equity premia, with the premia reaching a maximum at about 0.69% monthly for the $m = 3$ market before falling to -0.53% for the 5th and most volatile market. The model fails to account for the negative risk premium of the smallest, highest idiosyncratic volatility markets.

Aggregate asset pricing implications. In Panel B of Table 5 we show the aggregate asset pricing implications of the model with market-specific λ_m. The aggregate equity premium is 2.9%, about 0.50% *higher* than in the benchmark single λ model, despite the fact that the average segmentation here is only $\bar{\lambda} = 0.115$, one-third the single λ benchmark. For comparison, the table shows the asset pricing implications for an otherwise identical single λ economy with $\lambda = \bar{\lambda} = 0.115$. The aggregation of the micro segmentation frictions across the different markets adds some 1.2% annual to the equity premium, taking it from 1.7% to 2.9%. Compared to the benchmark model, the risk-free rate has about the same level and is slightly less volatile (though still too volatile compared to the data).

6 Conclusion

We propose a tractable consumption-based model in order to explain and quantify the macro impact of financial market frictions. We envision an economy comprised of many micro financial markets that are partially segmented from one another. Because of segmentation, traders in each micro market have to bear some local idiosyncratic risk. Assets in each market are priced by a convex combination of the marginal utility of a local trader specialized in that asset (who has to bear some of the idiosyncratic risk of that asset), and the economy-wide average marginal utility (reflecting the diversification of the remaining idiosyncratic risk in a large portfolio). We calibrate the model when all markets share the same level of segmentation and show that it can generate a sizeable equity premium. We also allow segmentation to differ across markets and show that aggregation matters: we can obtain essentially the same aggregate asset pricing implication with a much smaller *average* level of segmentation.

Technical Appendix

A General model with detailed derivations

We add three features relative to the model presented in the main text: (i) for each market m there is a density $\omega_m \geq 0$ of traders, (ii) the asset supply is $S_m \geq 0$, not normalized to 1, and (iii) there are bonds in positive net supply held in the family portfolio. The total measure of traders is one, i.e.,

$$\int_0^1 \omega_m \, dm = 1. \tag{23}$$

The average segmentation parameters is then taken to be $\bar{\lambda} := \int_0^1 \lambda_m \omega_m \, dm$. Each period one share of the asset produces a stochastic realization of a non-storable dividend $y_{m,t} > 0$. The aggregate endowment available to the entire economy is:

$$y_t = \int_0^1 y_{m,t} S_m \omega_m \, dm. \tag{24}$$

As in the text, traders in market m are assumed to bear an exogenous fraction $\lambda_m \in [0,1]$ of the expense of purchasing assets in that market and in return receive λ_m of the benefit. The remaining $1-\lambda_m$ of the expenses and the benefits is borne by the family. As show in the text, this results in a sequential budget constraint of the form:

$$c_{m,t} + \lambda_m p_{m,t} s_{m,t+1} + (1-\lambda_m) T_{t+1} \leq \lambda_m (p_{m,t} + y_{m,t}) s_{m,t} + (1-\lambda_m) T_t - \tau_{m,t}, \tag{25}$$

where the new term, $\tau_{m,t}$, is a lump-sum tax levied on market m by the government. As in the main text, T_t and T_{t+1} represent, respectively, the cum-dividend value of the family portfolio brought into the period and the ex-dividend value of the family portfolio acquired this period. Proceeding as in the text, we find that T_t and T_{t+1} satisfy:

$$\begin{aligned}
(1-\bar{\lambda}) T_t &= \int_0^1 (1-\lambda_n)(p_{n,t} + y_{n,t}) s_{n,t} \omega_n \, dn + b_{1,t} + \sum_{k \geq 1} \pi_{k,t} b_{k+1,t} \\
(1-\bar{\lambda}) T_{t+1} &= \int_0^1 (1-\lambda_n) p_{n,t} s_{n,t+1} \omega_n \, dn + \sum_{k \geq 1} \pi_{k,t} b_{k,t+1},
\end{aligned}$$

where $\pi_{k,t}$ and $b_{k,t}$ denote the price and quantity of purchases of zero-coupon bonds that pay the family one (real) dollar for sure in k periods' time.

Government. The government collects lump-sum taxes from each market and issues zero-coupon bonds of various maturities subject to the period budget constraint:

$$B_{1,t} + \sum_{k\geq 1} \pi_{k,t} B_{k+1,t} \leq \sum_{k\geq 1} \pi_k B_{k,t+1} + \int_0^1 \tau_{m,t}\omega_m \, dm, \tag{26}$$

where $B_{k,t}$ denotes the government's issue of k-period bonds at time $t-1$. We choose a particular specification of lump-sum taxes that has the property of not redistributing resources across markets:

$$\tau_{m,t} = \frac{1-\lambda_m}{1-\bar{\lambda}} \left(B_{1,t} + \sum_{k\geq 1} \pi_{k,t} \left[B_{k+1,t} - B_{k,t+1} \right] \right). \tag{27}$$

Equilibrium allocations. Market clearing requires $s_{m,t+1} = S_m$ for each m and $b_{k,t+1} = B_{k,t+1}$ for each k. We plug these conditions in the market-specific budget constraints and then use the government budget constraint combined with the expressions (27) for lump-sum taxes. After cancelling common terms we get:

$$c_{m,t} = \lambda_m y_{m,t} S_m + (1-\lambda_m) \int_0^1 \frac{1-\lambda_n}{1-\bar{\lambda}} y_{n,t} S_n \omega_n \, dn.$$

First-order conditions and asset pricing. Let $\mu_{m,t} \geq 0$ denote the multiplier on the budget constraint for market m and use the market-specific budget constraints and accounting identities for the family portfolio to write the Lagrangian:

$$\mathscr{L} = \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \int_0^1 \Big\{ u(c_{m,t}) + \mu_{m,t} \mathcal{B}_{m,t} \Big\} \omega_m \, dm \right]$$

where

$$\begin{aligned}
\mathcal{B}_{m,t} =& \lambda_m (p_{m,t} + y_{m,t}) s_{m,t} \\
&+ \frac{1-\lambda_m}{1-\bar{\lambda}} \left(\int_0^1 (1-\lambda_n)(p_{n,t} + y_{n,t}) s_{n,t} \omega_n \, dn + b_{1,t} + \sum_{k\geq 1} \pi_{k,t} b_{k+1,t} \right) \\
&- \left[c_{m,t} + \lambda_m p_{m,t} s_{m,t+1} + \frac{1-\lambda_m}{1-\bar{\lambda}} \left(\int_0^1 (1-\lambda_n) p_{n,t} s_{n,t+1} \omega_n \, dn + \sum_{k\geq 1} \pi_{k,t} b_{k,t+1} \right) + \tau_{m,t} \right].
\end{aligned}$$

Now collecting terms in $\int_0^1 \mu_{m,t}\mathcal{B}_{m,t}\omega_m\, dm$ and rearranging:

$$
\begin{aligned}
&\int_0^1 \mu_{m,t}\mathcal{B}_{m,t}\omega_m\, dm \\
&= \int_0^1 \mu_{m,t}\Big\{\lambda_m(p_{m,t}+y_{m,t})s_{m,t} - c_{m,t} - \lambda_m p_{m,t}s_{m,t+1} - \tau_{m,t}\Big\}\omega_m\, dm \\
&+ \int_0^1 \mu_{m,t}\frac{1-\lambda_m}{1-\bar{\lambda}}\Big\{b_{1,t} + \sum_{k\geq 1}\pi_{k,t}(b_{k+1,t} - b_{k,t+1})\Big\}\,\omega_m\, dm \\
&+ \int_0^1 \mu_{m,t}\frac{1-\lambda_m}{1-\bar{\lambda}}\int_0^1 (1-\lambda_n)\Big[(p_{n,t}+y_{n,t})s_{n,t} - p_{n,t}s_{n,t+1}\Big]\omega_n\omega_m\, dn\, dm.
\end{aligned}
$$

Now, in the last term, we permute the roles of the symbols m and n and then interchange the order of integration:

$$
\begin{aligned}
&\int_0^1 \mu_{n,t}\frac{1-\lambda_n}{1-\bar{\lambda}}\int_0^1 (1-\lambda_m)[(p_{m,t}+y_{m,t})s_{m,t} - p_{m,t}s_{m,t+1}]\omega_m\omega_n\, dm\, dn \\
= \;& \left[\int_0^1 \mu_{n,t}\frac{1-\lambda_n}{1-\bar{\lambda}}\omega_n\, dn\right]\int_0^1 (1-\lambda_m)[(p_{m,t}+y_{m,t})s_{m,t} - p_{m,t}s_{m,t+1}]\omega_m\, dm.
\end{aligned}
$$

Next, define the weighted average of Lagrange multipliers:

$$
q_{m,t} := \lambda_m\mu_{m,t} + (1-\lambda_m)q_t, \text{ and } q_t := \int_0^1 \frac{1-\lambda_n}{1-\bar{\lambda}}\mu_{n,t}\omega_n\, dn,
$$

as in the main text. Substituting for $q_{m,t}$ and q_t we get:

$$
\begin{aligned}
\mathscr{L} = \mathbb{E}_0\Bigg[\sum_{t=0}^{\infty}\beta^t\int_0^1 \Big\{&u(c_{m,t}) + q_{m,t}(p_{m,t}+y_{m,t})s_{m,t} - q_{m,t}p_{m,t}s_{m,t+1} \\
&- \mu_{m,t}(c_{m,t}+\tau_{m,t}) + q_t\Big(b_{1,t} + \sum_{k\geq 1}\pi_{k,t}(b_{k+1,t} - b_{k,t+1})\Big)\Big\}\omega_m\, dm\Bigg].
\end{aligned}
$$

Apart from the term reflecting the presence of bonds, this is the same Langrangian as in the main text. We take derivatives (point-wise) to obtain the first-order necessary conditions reported in the main text.

Portfolio weights and returns. To streamline the exposition we return to the model used in the main text. The total value of the family portfolio is:

$$
\int_0^1 \frac{1-\lambda_m}{1-\bar{\lambda}}p_{m,t}s_{m,t+1}\, dm.
$$

Thus, in the family portfolio, asset m is represented with a weight:

$$\psi_{m,t} := \frac{\frac{1-\lambda_m}{1-\bar{\lambda}} p_{m,t} s_{m,t+1}}{\int_0^1 \frac{1-\lambda_n}{1-\bar{\lambda}} p_{n,t} s_{n,t+1} \, dn}.$$

Letting $R_{m,t+1} = (p_{m,t+1} + y_{m,t+1})/p_{m,t}$ be the return on asset m, the return on the family portfolio can be written:

$$R_{t+1} = \int_0^1 R_{m,t+1} \psi_{m,t} \, dm.$$

Now recall that trader m holds $\lambda_m p_{m,t} s_{m,t+1}$ real dollars of asset m, and the rest of his investment:

$$(1 - \lambda_m) \int_0^1 \frac{1-\lambda_n}{1-\bar{\lambda}} p_{n,t} s_{n,t+1} \, dn,$$

is in the family portfolio. Thus, the return of trader's m portfolio can be written:

$$\Psi_{m,t} R_{m,t+1} + (1 - \Psi_{m,t}) R_{t+1},$$

where:

$$\Psi_{m,t} := \frac{\lambda_m p_{m,t} s_{m,t+1}}{\lambda_m p_{m,t} s_{m,t+1} + (1 - \lambda_m) \int_0^1 \frac{1-\lambda_n}{1-\bar{\lambda}} p_{n,t} s_{n,t+1} \, dn},$$

is the portfolio weight in the local asset.

B Computational details

Information. The aggregate state is a VAR for log consumption growth and log idiosyncratic volatility:

$$\begin{aligned} \log g_{t+1} &= (1-\rho) \log \bar{g} + \rho \log g_t + \varepsilon_{g,t+1} \\ \log \sigma_{t+1} &= (1-\phi) \log \bar{\sigma} + \phi \log \sigma_t - \eta \left(\log g_t - \log \bar{g} \right) + \varepsilon_{v,t+1}, \end{aligned}$$

where $0 \leq \rho, \phi < 1$ and where the two components of innovation, $\epsilon_{g,t+1}$ and $\epsilon_{v,t+1}$, are assumed to be contemporaneously uncorrelated. The dividend in market m is:

$$\log y_{m,t} = \log y_t + \log \hat{y}_{m,t}, \tag{28}$$

where the log idiosyncratic component is conditionally IID normal in the cross section:

$$\begin{aligned} \log \hat{y}_{m,t} &\sim \text{IID across } m \text{ and } N(-\sigma_{mt}^2/2, \sigma_{mt}^2) \\ \sigma_{mt} &= \sigma_t \hat{\sigma}_m, \end{aligned}$$

for some time-invariant market specific volatility level $\hat{\sigma}_m$.

Setup. Let utility be CRRA with coefficient $\gamma > 0$ so $u'(c) = c^{-\gamma}$. Assume markets come in N different *types* $m \in \{1, \ldots, N\}$. Note that this is an abuse of notation given that we previously used m to index a single market within the $[0, 1]$ continuum. There is an equal measure of assets, $1/N$, in each market type. The total measure of traders in a market of type m is denoted by ω_m. Thus, we have the restriction:

$$\sum_{m=1}^{N} \omega_m = 1.$$

The supply of asset per trader in a market of type m is S_m, so the total supply in that market is $S_m \omega_m$. The dividend is $y_{m,t} = y_t \hat{y}_{m,t}$ where $\mathbb{E}\left[\hat{y}_{m,t} \,|\, g_t, \sigma_t\right] = 1$. Since the aggregate endowment is y_t, we need to impose the restriction:

$$\sum_{m=1}^{N} S_m \omega_m = 1.$$

The segmentation parameter in a market of type m is λ_m and the supply per trader is S_m. In equilibrium, consumption in a market of type m is given by:

$$c_{m,t} = y_t \left(A_m + B_m \hat{y}_{m,t}\right),$$

where

$$A_m := (1 - \lambda_m) \sum_{n=1}^{N} \frac{1 - \lambda_n}{1 - \bar{\lambda}} S_n \omega_n,$$

and

$$B_m := \lambda_m S_m.$$

We then have $q_{m,t} = \theta_{m,t} y_t^{-\gamma}$ where:

$$\theta_{m,t} = \lambda_m \left(A_m + B_m \hat{y}_{m,t}\right)^{-\gamma} + (1 - \lambda_m) \sum_{n=1}^{N} \frac{1 - \lambda_n}{1 - \bar{\lambda}} \mathbb{E}\left[\left(A_n + B_n \hat{y}_{n,t}\right)^{-\gamma} \,|\, g_t, \sigma_t\right] \omega_n,$$

where, by the LLN, the conditional expectation on the right–hand side calculates the cross-sectional average of $(A_n + B_n \hat{y}_{n,t})^{-\gamma}$ within type n markets. We explain below how to compute this expectation. Now let $\hat{p}_{m,t} := p_{m,t}/y_t$ be the price/dividend ratio in a type m market. This solves:

$$\hat{p}_{m,t} = \mathbb{E}_t \left[\beta g_{t+1}^{1-\gamma} \frac{\theta_{m,t+1}}{\theta_{m,t}} \left(\hat{p}_{m,t+1} + \hat{y}_{m,t+1}\right)\right]. \tag{29}$$

Approximation. Each market is characterized by 3 states: two aggregate states (g, σ) and one idiosyncratic state $\hat{y}_m$ (to simplify notation, we omit the 'log'). Given the specification above, the transition density is of the form:

$$f(g', \sigma', \hat{y}' \,|\, g, \sigma, \hat{y}) = f(g', \sigma' \,|\, g, \sigma) f(\hat{y}' \,|\, \sigma').$$

Our approximation follows Tauchen and Hussey (1991). First, we pick quadrature nodes and weights for the aggregate state: consumption growth, Q_g and W_g (column vectors of size N_g) and volatility, Q_σ and W_σ (column vectors of size N_σ).

In their original paper, Tauchen and Hussey recommended to pick these nodes and weights according to the transition density evaluated at the mean, i.e., a bivariate Gaussian density $f(g', \sigma' \,|\, \bar{g}, \bar{\sigma})$ which in the present case is the product of two independent normal densities with means $\log \bar{g}$ and $\log \bar{\sigma}$, respectively, and variances σ_g^2 and σ_v^2. Subsequent work has highlighted, however, that when the Markov chain being approximated is highly persistent, the quality of the approximation may be poor. In our calibration exercise, this problem may arise when the moment matching algorithm searches in the region where the volatility process, σ, is highly persistent (ϕ close to 1). To alleviate this concern we follow Flodèn (2008): we generate nodes and weights for σ based on a "twisted" Gaussian density with a higher standard deviation:

$$\sigma = w\sigma_v + (1-w)\frac{\sigma_v}{\sqrt{1-\phi^2}} \qquad \text{where } w = 1/2 + \phi/4. \tag{30}$$

We also use a larger number of nodes to better capture the impact of high realization of σ. See Appendix C below for further discussion of the robustness of the approximation, and Table 7 for the results of our robustness analysis.

Next, for every quadrature value of σ, we generate quadrature nodes and weights in each market type m for the log idiosyncratic state $\log \hat{y}$, according to a Gaussian density with mean $-\hat{\sigma}_m^2\sigma^2/2$ and variance $\hat{\sigma}_m^2\sigma^2$. The resulting nodes and weights column vectors have length $N_\sigma \times N_{\hat{y}}$ and we denote them by $Q^m_{\hat{y}\,|\,\sigma}$ and $W^m_{\hat{y}\,|\,\sigma}$. In these vectors of nodes and weights, we adopt the convention that "idiosyncratic endowment comes first". That is, in the quadrature node vector, idiosyncratic endowment i under volatility j is found in entry $i + N_{\hat{y}}(j-1)$.

Now, if we combine idiosyncratic endowment, aggregate volatility, and aggregate endowment growth together we obtain, for each market type m, a finite state space that we index by $n \in \{1, 2, 3, \ldots N_s\}$, where

$$N_s := N_{\hat{y}} \times N_\sigma \times N_g.$$

We adopt the convention that the state of idiosyncratic endowment $i \in \{1, \ldots, N_{\hat{y}}\}$, volatility $j \in \{1, \ldots, N_\sigma\}$, and aggregate consumption growth $k \in \{1, \ldots, N_g\}$ correspond to state:

$$n = i + N_{\hat{y}}(j-1) + N_{\hat{y}}N_\sigma(k-1).$$

In each state, the value of idiosyncratic endowment, aggregate volatility, and aggregate consumption growth can be conveniently represented with Kronecker products of the quadrature nodes:

$$\begin{aligned} V_g &= Q_g \otimes e_{N_\sigma} \otimes e_{N_y} \\ V_\sigma &= e_{N_g} \otimes Q_\sigma \otimes e_{N_y} \\ V_{\hat{y}}^m &= e_{N_g} \otimes Q_{\hat{y}\,|\,\sigma}^m, \end{aligned}$$

where e_{N_x} denotes a $N_x \times 1$ vector of ones. By construction, entry n of vector V_g contains consumption growth if the state of market m is n, and similarly for V_σ and $V_{\hat{y}}^m$. The corresponding quadrature weights are obtained as follows. We let:

$$\begin{aligned} A &= W_g \otimes e_{N_\sigma} \otimes e_{N_y} \\ B &= e_{N_g} \otimes W_\sigma \otimes e_{N_y} \\ C^m &= e_{N_g} \otimes W_{\hat{y}\,|\,\sigma}^m, \end{aligned}$$

so that the quadrature weights for the state are:

$$W^m = A. * B. * C^m$$

where $.*$ denotes Matlab coordinate-per-coordinate product.

Transition probability matrix. To implement the method of Tauchen and Hussey (1991), we define a Matlab function:

$$f^m(s' \,|\, s) = f^m(\hat{y}' \,|\, \sigma') \times f(\sigma' \,|\, \sigma, g) \times f(g' \,|\, g),$$

as well as the quadrature weighting function:

$$\omega^m(s) = \omega^m(\hat{y} \,|\, \sigma) \times \omega(\sigma) \times \omega(g),$$

which is the probability density function used above to generate the quadrature nodes and weights for market m. Letting $N_s := N_{\hat{y}} \times N_\sigma \times N_g$, the matrix formula for the transition matrix is:

$$\begin{aligned} G \;=\; & f^m(e_{N_s} V_{\hat{y}}' \,|\, e_{N_s} V_\sigma'). * f(e_{N_s} V_\sigma' \,|\, V_\sigma e_{N_s}', V_g e_{N_s}'). * f(e_{N_s} V_g' \,|\, V_g e_{N_s}') \\ & . * (e_{N_s} * W')./\left[e_{N_s}. * \omega(V_{\hat{y}}' \,|\, V_\sigma'). * \omega(V_\sigma'). * \omega(V_g')\right], \end{aligned}$$

which we then normalize so that the rows sum to 1.

Calculating cross-sectional moments. In many instances in the program we need to calculate:

$$\mathbb{E}\left[x_m \,|\, g, \sigma\right],$$

for some random variable x_m. To do this, we consider:

$$K_\sigma = \left(I_{N_g \times N_\sigma} \otimes e'_{N_{\hat{y}}}\right)\left[x_m. * W^m\right],$$

where

$$W^m = e_{N_g} \otimes W^m_{\hat{y}\,|\,\sigma}.$$

The coordinate-wise product multiplies each realization of x_m by its probability conditional on (g, σ), and the pre-multiplication adds up. We then re-Kroneckerize this in order to obtain a $N_s \times 1$ vector:

$$K_\sigma \otimes e_{N_{\hat{y}}}.$$

C Robustness of the approximation

Table 7 illustrates that our numerical results are robust to alternative parameterizations of the numerical approximations. We consider three versions of the single λ economy: the benchmark version, the version with constant σ, and the feedback version with counter-cyclical σ_t. In our default *standard* parameterization we have $N = N_g \times N_\sigma \times N_{\hat{y}} = 3 \times 9 \times 19 = 513$ quadrature nodes and weights. It also uses the "twisted" density recommended by Floden (2008) to alleviate concerns about the accuracy of the Tauchen and Hussey (1991) procedure when the σ_t process is persistent (see equation (30) in Appendix B). In our *high* precision parameterization we have $N = N_g \times N_\sigma \times N_{\hat{y}} = 5 \times 19 \times 25 = 2,375$ nodes and weights and again use the twisting recommended by Floden. In the *no twist* parameterization we use the plain Tauchen and Hussey (1991) procedure and the same configuration of nodes as in the standard parameterization. The issue of twisting does not arise in the constant σ model.

For each of these numerical approximations the table reports the calibrated parameter values, the values of the moments we target, and the implications for aggregate asset prices.

For a given model, we see that increasing the number of nodes from the standard to high parameterization has negligible effect on the results. Similarly, the twisting recommended by Floden has negligible effect. This suggests that our calibrated stochastic process is not persistent enough to cause any problems for the plain Tauchen and Hussey procedure.

Panel A: *Preferences and aggregate endowment growth.*

Parameter	Monthly value	Notes
β	0.9992	discount rate 1% annual
γ	4	coefficient relative risk aversion
$\bar{g}$	1.0017	average aggregate growth 2% annual
$\sigma_{\epsilon g}$	0.0029	std dev aggregate growth 1% annual

Panel B: *Segmentation and idiosyncratic endowment volatility.*

		Model				
Parameter	Benchmark	Constant	Feedback	Data moment		
λ	0.310	0.310	0.310	std dev diversified market portfolio return	4.16%	monthly
$\bar{\sigma}$	0.318	0.318	0.318	average cross-section std dev returns	16.40%	monthly
$\sigma_{\epsilon v}$	0.207	0	0.207	time-series std dev cross-section std dev returns	4.17%	monthly
ϕ	0.784	0	0.785	AR(1) cross-section std dev returns	0.84	monthly
η	n/a	n/a	2.513	cross-section std dev returns on lagged growth	-0.56	in monthly data

Table 1: Parameter choices.

The top panel shows our parameters for preferences and aggregate endowment growth. These parameters are kept the same in all calculations. The bottom panel shows our parameters for segmentation and the idiosyncratic endowment volatility process σ_t and the moments in the Goyal and Santa-Clara (2003) cross-sectional standard deviation of stock returns data that they are chosen to match. The Benchmark model has a single common segmentation parameter λ and time-varying idiosyncratic endowment volatility σ_t. The Constant σ model sets $\sigma_t = \bar{\sigma}$ i.e., to the Benchmark unconditional mean, for all t. The Feedback model has *counter-cyclical* endowment volatility, with feedback from aggregate growth g_t to volatility σ_t governed by the elasticity η. See the main text for further details.

Moment	Data	Benchmark	Model Constant	Feedback
std dev diversified market portfolio return	4.16	4.16	1.01	4.16
average cross-section std dev returns	16.40	16.40	16.03	16.35
time-series std dev cross-section std dev returns	4.17	4.17	0	4.17
AR(1) cross-section std dev returns	0.84	0.84	n/a	0.84
regression cross-section std dev returns on lagged growth	−0.56	n/a	n/a	−0.56

Table 2: Fit of calibrated models.

Our target moments in the US monthly postwar Goyal and Santa-Clara (2003) cross-sectional standard deviation of stock returns data and their model counterparts. The Benchmark model has a single common segmentation parameter λ and time-varying idiosyncratic endowment volatility σ_t. The Constant σ model sets $\sigma_t = \bar{\sigma}$ i.e., to the Benchmark unconditional mean, for all t. The Feedback model has *counter-cyclical* endowment volatility, with feedback from aggregate growth g_t to volatility σ_t governed by the elasticity η. See the main text for further details.

	Moment	Data	Model Benchmark	Model Constant	Model Feedback
equity premium	$\mathbb{E}[R_M - R_f]$	5.43	2.43	0.22	2.43
	$\mathrm{Std}[R_M - R_f]$	14.25	13.27	1.01	13.27
sharpe ratio	$\mathbb{E}[R_M - R_f]/\mathrm{Std}[R_M - R_f]$	0.38	0.17	0.20	0.17
market return	$\mathbb{E}[R_M]$	7.24	10.62	9.47	10.62
	$\mathrm{Std}[R_M]$	14.44	14.41	1.01	14.41
risk free rate	$\mathbb{E}[R_f]$	1.81	8.19	9.25	8.19
	$\mathrm{Std}[R_f]$	1.20	5.55	0	5.57
price/dividend ratio	$\mathbb{E}[p/y]$ (annual)	34.38	14.10	14.13	14.10
	$\mathrm{Std}[\log(p/y)]$ (annual)	38.63	20.56	0	20.56
	$\mathrm{Auto}[\log(p/y)]$ (monthly)	0.99	0.76	n/a	0.76

Table 3: Aggregate asset pricing implications of single λ model.

Aggregate asset pricing moments in postwar US data. All return data is monthly 1959:1-2007:12 and reported in annualized percent. The stock market index is the value weighted NYSE return from CRSP, and the risk-free return is the 90 day T-bill rate. We obtain real returns after deflating by the CPI from the BLS. Data on price/dividend ratios is annual 1959-2007. To annualize monthly returns we multiply by 12 and to annualize monthly standard deviations we multiply by $\sqrt{12}$. In Table 6 below, we compare annualized monthly returns to returns calculated by explicitly time-aggregating from monthly to yearly. The Benchmark model has a single common segmentation parameter λ and time-varying idiosyncratic endowment volatility σ_t. The Constant σ model sets $\sigma_t = \bar{\sigma}$ i.e., to the Benchmark unconditional mean, for all t. The Feedback model has *counter-cyclical* endowment volatility, with feedback from aggregate growth g_t to volatility σ_t governed by the elasticity η. See the main text for further details.

Panel A: *Segmentation parameters.*

	Parameter		Moment			
			Portfolio std dev		Market share	
Market m	λ_m	ω_m	Data	Model	Data	Model
1	0.010	0.514	3.83	4.18	0.535	0.538
2	0.178	0.277	4.74	4.52	0.274	0.272
3	0.264	0.128	5.85	5.72	0.119	0.118
4	0.324	0.058	7.13	7.02	0.052	0.052
5	0.365	0.023	8.16	8.08	0.020	0.020
average	0.115					

Panel B: *Idiosyncratic endowment volatility.*

Parameter		Moment	Data	Model
$\bar{\sigma}$	0.816	average cross-section std dev returns	16.40	16.46
$\sigma_{\epsilon v}$	0.198	time-series std dev cross-section std dev returns	4.17	4.18
ϕ	0.891	AR(1) cross-section std dev returns	0.84	0.85

Table 4: Market-specific segmentation: parameters and fit.

The top panel shows the five segmentation parameters λ_m and measures of traders ω_m, for $m = 1, ...5$, and the portfolio standard deviation and market share moments in the Ang, Hodrick, Xing, and Zhang (2006) data they are chosen to match. The bottom panel shows the idiosyncratic endowment volatility process parameters and the moments in the Goyal and Santa-Clara (2003) cross-sectional standard deviation of stock returns data they are chosen to match.

Panel A: *Market-specific asset pricing implications.*

	Risk premia	
Market m	Data	Model
1	0.53	0.17
2	0.65	0.23
3	0.69	0.33
4	0.36	0.44
5	-0.53	0.53

Panel B: *Aggregate asset pricing implications.*

			Model	
	Moment	Data	λ_m	$\bar{\lambda}$
equity premium	$\mathbb{E}[R_M - R_f]$	5.27	2.92	1.69
	$\mathrm{Std}[R_M - R_f]$	14.25	15.54	11.11
sharpe ratio	$\mathbb{E}[R_M - R_f]/\mathrm{Std}[R_M - R_f]$	0.38	0.17	0.14
market return	$\mathbb{E}[R_M]$	7.24	11.07	10.27
	$\mathrm{Std}[R_M]$	14.44	16.16	11.56
risk free rate	$\mathbb{E}[R_f]$	1.81	8.15	8.58
	$\mathrm{Std}[R_f]$	1.20	3.65	2.92
price/dividend ratio	$\mathbb{E}[p/y]$ (annual)	34.38	13.90	14.04
	$\mathrm{Std}[\log(p/y)]$ (annual)	38.63	32.84	23.22
	$\mathrm{Auto}[\log(p/y)]$ (monthly)	0.99	0.88	0.88

Table 5: Asset pricing implications of market-specific segmentation.

The top panel shows the market risk premia implied by the five markets $m = 1, ...5$ and their counterparts in the Ang, Hodrick, Xing, and Zhang (2006) data. These are reported as monthly percent. The bottom panel shows the aggregate asset pricing implications. The column marked λ_m refers to the model with market-specific segmentation parameters while the column marked $\bar{\lambda}$ refers to a model with a single segmentation parameter λ that is set equal to the mean $\bar{\lambda} = \sum_m \lambda_m \omega_m$ of the market-specific λ_m model.

	Annualized monthly	Aggregated to yearly
average real risk-free rate	1.81	1.81
standard deviation of real risk free rate	1.20	2.43
average real NYSE return	7.24	7.27
standard deviation of real NYSE return	14.40	13.90
equity premium	5.43	5.47
standard deviation of equity premium	14.25	13.30
average price-dividend ratio	495.18	34.38
standard deviation of log price-dividend ratio	0.56	0.34
autocorrelation of log price-dividend ratio	-0.02	0.89
average consumption growth	2.19	2.17
standard deviation of consumption growth	1.25	1.33

Table 6: Aggregate statistics in annualized monthly data and in monthly data time-aggregated to yearly.

Aggregate postwar US data. All return data is monthly 1959:1-2007:12 and reported in annualized percent. The stock market index is the value weighted NYSE return from CRSP, and the risk-free return is the 90 day T-bill rate. Real consumption growth refers to the growth of real nondurables and services consumption per capita from the BEA. The first column shows annualized statistics for monthly data. To annualize monthly returns and consumption growth, we multiply by 12, and to annualize monthly standard deviations, we multiply by $\sqrt{12}$. The second column shows statistics for yearly data, which are obtained by compounding returns and growth over the relevant time interval. The only statistics that are substantially different in this second column concern the price dividend ratio. This is because, in the first column, the dividend that enters the ratio is the dividend per month, while in the second column it is the dividend paid over the entire year.

	Benchmark			Constant		Feedback		
	Standard	High	No twist	Standard	High	Standard	High	No twist
Calibrated parameters								
λ	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31
$\bar{\sigma}$	0.32	0.32	0.32	0.32	0.32	0.32	0.32	0.32
$\sigma_{\epsilon v}$	0.21	0.21	0.21	0	0	0.21	0.21	0.21
ϕ	0.78	0.78	0.79	n/a	n/a	0.78	0.78	0.79
η	n/a	n/a	n/a	n/a	n/a	2.51	2.51	2.52
Fitted moments								
std dev diversified market portfolio return	4.16	4.16	4.16	1.01	1.01	4.16	4.16	4.16
average cross-section std dev returns	16.40	16.40	16.40	16.03	16.03	16.40	16.40	16.40
time-series std dev cross-section std dev returns	4.17	4.17	4.17	0	0	4.17	4.17	4.17
AR(1) cross-section std dev returns	0.84	0.84	0.84	n/a	n/a	0.84	0.84	0.84
cross-section std dev returns on lagged growth	n/a	n/a	n/a	n/a	n/a	-0.56	-0.56	-0.56
Asset pricing implications								
$\mathbb{E}[R_M - R_f]$	2.43	2.43	2.43	0.22	0.22	2.43	2.43	2.42
$\text{Std}[R_M - R_f]$	13.27	13.27	13.27	1.01	1.01	13.27	13.27	13.26
$\mathbb{E}[R_M - R_f]/\text{Std}[R_M - R_f]$	0.17	0.17	0.17	0.20	0.20	0.17	0.17	0.17
$\mathbb{E}[R_M]$	10.62	10.62	10.62	9.47	9.47	10.62	10.62	10.62
$\text{Std}[R_M]$	14.41	14.41	14.41	1.01	1.01	14.41	14.41	14.41
$\mathbb{E}[R_f]$	8.19	8.19	8.19	9.25	9.25	8.19	8.19	8.20
$\text{Std}[R_f]$	5.55	5.54	5.61	0	0	5.57	5.57	5.64
$\mathbb{E}[p/y]$	14.10	14.10	14.10	14.13	14.13	14.10	14.10	14.10
$\text{Std}[\log(p/y)]$	20.56	20.56	20.50	0	0	20.57	20.57	20.50
$\text{Auto}[\log(p/y)]$	0.76	0.76	0.76	n/a	n/a	0.76	0.76	0.76

Table 7: Robustness of single λ model solutions.

The *standard* precision case has $N = N_g \times N_\sigma \times N_{\hat{y}} = 3 \times 9 \times 19 = 513$ quadrature nodes and weights. The *high* precision case has $N = N_g \times N_\sigma \times N_{\hat{y}} = 5 \times 19 \times 25 = 2,375$. In both these cases, the "twisted" density recommended by Flodèn (2008) is used to alleviate concerns about the accuracy of the Tauchen and Hussey (1991) procedure when the stochastic process is persistent (see equation (30) in Appendix B). The final *no twist* case uses the plain Tauchen and Hussey procedure and same configuration of nodes as in the standard case. The issue of twisting does not arise in the constant σ model.

References

S. Rao Aiyagari and Mark Gertler. Asset returns with transaction costs and uninsurable individual risks. *Journal of Monetary Economics*, 27:309–331, 1991. 3

S. Rao Aiyagari and Mark Gertler. Overreaction of asset prices in general equilibrium. *Review of Economic Dynamics*, 2(1):3–35, 1999. 3

Fernando Alvarez, Andrew Atkeson, and Patrick J. Kehoe. Money, interest rates, and exchange rates with endogenously segmented markets. *Journal of Political Economy*, 110:73–112, 2002. 5

Andrew Ang, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang. The cross-section of volatility and expected returns. *Journal of Finance*, 61:259–299, 2006. 2, 21, 35, 36

David K. Backus and Stanley E. Zin. Reverse engineering the yield curve. NBER Working Paper 4676, 1994. 17

Turan Bali, Nusret Cakici, Xuemin (Sterling) Yan, and Zhe Zhang. Does idiosyncratic risk really matter? *Journal of Finance*, 60(2):905–929, 2005. 2, 13

Ravi Bansal and Amir Yaron. Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance*, 59(4):1481–1509, 2004. 3

Suleyman Basak and Domenico Cuoco. An equilibrium model with restricted stock market participation. *Review of Financial Studies*, 11(2):309–341, 1998. 4

Douglas Breeden. An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics*, 7:265–296, 1979. 1, 3, 4, 10, 11, 16, 20

John Y. Campbell and John H. Cochrane. By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy*, 107 (2):205–251, 1999. 3

John Y. Campbell, Martin Lettau, Burton G. Malkiel, and Yexiao Xu. Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *Journal of Finance*, 56(1):1–43, 2001. 18

Yi-Li Chien, Harold Cole, and Hanno Lustig. A multiplier approach to understanding the macro implications of household finance. *Review of Economic Studies*, 2011. Forthcoming. 3, 4

John H. Cochrane. Asset pricing: Liquidity, trading and asset prices. *NBER Reporter*, pages 1–12, 2005. 1

John H. Cochrane and Lars Peter Hansen. Asset pricing explorations for macroeconomics. In Olivier Jean Blanchard and Stanley Fisher, editors, *NBER Macroeconomics Annual*, pages 115–165. MIT Press, 1992. 16

Pierre Collin-Dufresne, Robert S. Goldstein, and J. Spencer Martin. The determinants of credit spread changes. *Journal of Finance*, 56(6):2177–2207, 2001. 1

George Constantinides and Darrell Duffie. Asset pricing with heterogenous consumers. *Journal of Political Economy*, 104:219–240, 1996. 19

Larry G. Epstein and Stanley E. Zin. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis. *Journal of Political Economy*, 99(2):263–286, 1989. 3

Eugene F. Fama and Kenneth R. French. Common risk factors in the returns on stock and bonds. *Journal of Financial Economics*, 33:131–156, 1993. 21

Martin Flodèn. A note on the accuracy of markov-chain approximations to highly persistent AR(1) processes. *Economics Letters*, 99(3):516–520, 2008. 12, 29, 31, 38

Xavier Gabaix, Arvind Krishnamurthy, and Olivier Vigneron. Limits of arbitrage: Theory and evidence from the mortgage-backed securities market. *Journal of Finance*, 62(2):557–595, 2007. 1

Francisco Gomes and Alexander Michaelides. Asset pricing with limited risk sharing and heterogeneous agents. *Review of Financial Studies*, 21(1):415–448, 2008. 4

Amit Goyal and Pedro Santa-Clara. Idiosyncratic risk matters! *Journal of Finance*, 58 (3):975–1007, 2003. 2, 13, 19, 32, 33, 35

Fatih Guvenen. A parsimonious macroeconomic model for asset pricing. *Econometrica*, 77(6):1711–1754, 2009. 4

Hua He and David M. Modest. Market frictions and consumption-based asset pricing. *Journal of Political Economy*, 103(1):94–117, 1995. 3

Zhiguo He and Arvind Krishnamurthy. A model of capital and crises. Working Paper, 2008a. 3

Zhiguo He and Arvind Krishnamurthy. Intermediary asset pricing. Working Paper, 2008b. 3

John Heaton and Deborah J. Lucas. Evaluating the effects of incomplete markets on risk sharing and asset pricing. *Journal of Political Economy*, 104:443–487, 1996. 3, 19

Narayana R. Kocherlakota. The equity premium: It's still a puzzle. *Journal of Economic Literature*, 34:42–71, 1996. 20

Karen A. Kopecky and Richard M.H. Suen. Finite state markov-chain approximations to highly persistent processes. *Review of Economic Dynamics*, 13(3):701–714, 2010. 12

Dirk Kruger and Hanno Lustig. When is market incompleteness irrelevant for the price of aggregate risk (and when is it not)? *Journal of Economic Theory*, 145(1):1–41, 2010. 19

Albert S. Kyle and Wei Xiong. Contagion as a wealth effect. *Journal of Finance*, 56: 1401–1440, 2001. 3

Stephen F. LeRoy. Excess volatility. In Steven N. Durlauf and Lawrence E. Blume, editors, *The New Palgrave Dictionary of Economics, 2nd Edition*. Palgrave Macmillan, 2006. 14

Robert E. Lucas, Jr. Asset prices in an exchange economy. *Econometrica*, 46(6):1429–1445, 1978. 1, 2, 3, 4, 9, 10, 11, 14, 15, 16, 18, 20

Robert E. Lucas, Jr. Liquidity and interest rates. *Journal of Economic Theory*, 50: 237–264, 1990. 1, 2, 5

Erzo G. Luttmer. Asset pricing in economies with frictions. *Econometrica*, 64(6): 1439–1467, 1996. 3

Erzo G. Luttmer. What level of fixed costs can reconcile consumption and stock returns? *Journal of Political Economy*, 107(5):969–997, 1999. 3

Saito Makoto. Limited participation and asset pricing. University of British Columbia, working paper, 1995. 4

N. Gregory Mankiw. The equity premium and the concentration of aggregate shocks. *Journal of Financial Economics*, 17:211–219, 1986. 19

Rajnish Mehra and Edward C. Prescott. The equity premium: A puzzle. *Journal of Monetary Economics*, 15:145–161, 1985. 1, 3, 4, 14, 15, 17, 18, 20

Anna Pavlova and Roberto Rigobon. The role of portfolio constraints in the international propagation of shocks. *Review of Economic Studies*, 75:1215–1256, 2008. 3

Mark Rubinstein. The valuation of uncertain income streams and the pricing of options. *Bell Journal of Economics*, 7(2):407–425, 1976. 3

Andrei Shleifer and Robert W. Vishny. The limits of arbitrage. *Journal of Finance*, 52 (1):35–55, 1997. 1

Kjetil Storesletten, Chris I. Telmer, and Amir Yaron. Cyclical dynamics in idiosyncratic labor market risk. *Journal of Political Economy*, 112(3):695–717, 2004. 18

George Tauchen and Robert Hussey. Quadrature based methods for obtaining approximate solutions to nonlinear asset pricing models. *Econometrica*, 59(2):371–396, 1991. 12, 29, 30, 31, 38

Chris I. Telmer. Asset-pricing puzzles and incomplete markets. *Journal of Finance*, 48 (5):1803–1832, 1993. 19

Dimitri Vayanos. Flight to quality, flight to liquidity, and the pricing of risk. Working Paper, London School of Economics, 2005. 3

Annette Vissing-Jørgensen. Limited asset market participation and the elasticity of intertemporal substitution. *Journal of Political Economy*, 110(4):825–853, 2002. 4

Philippe Weil. The equity premium puzzle and the risk-free rate puzzle. *Journal of Monetary Economics*, 24(401-421), 1989. 3, 16

Philippe Weil. Nonexpected utility in macroeconomics. *Quarterly Journal of Economics*, 105(1):29–42, 1990. 3