Climbing and Falling Off the Ladder: Asset Pricing Implications of Labor Market Event Risk

Lawrence Schmidt*
University of California, San Diego
(Job Market Paper)

This version: January 2015
First version: June 2013

Abstract

This paper proposes state-dependent, idiosyncratic tail risk as a key driver of asset prices. I provide new evidence on the importance of tail events in explaining the shape of the idiosyncratic distribution of income growth rates and its evolution over time. I then formalize its role within a tractable affine, jump-diffusion asset pricing framework with recursive preferences, heterogeneous agents and incomplete markets, making my results immediately applicable to a wide class of existing models for aggregate dynamics. Next, I demonstrate its importance using a calibrated model in which agents are exposed to a time-varying probability of experiencing a rare, idiosyncratic disaster. The model, whose parameters are disciplined by the data, matches the level and dynamics of the equity premium. Empirically, stock returns are highly informative about labor market event risk, and, consistent with the model’s predictions, initial claims for unemployment, a proxy for labor market uncertainty, is a highly robust predictor of returns.

*Department of Economics, UCSD, 9500 Gilman Drive, La Jolla CA 92093, lschmidt@ucsd.edu. I am particularly indebted to my advisor, Allan Timmermann, for his help and support throughout the writing of this paper. I am also grateful to participants at 4th conference of the Macro-Finance Society (especially François Gourio, the discussant) as well as the SITE conference on “Interrelations between Labor Markets and Financial Markets”, the UCSD applied, macroeconomics, and finance workshops, Brendan Beare, Jonathan Berk, Jules Van Binsbergen, John Campbell, John Cochrane, Steven Davis, Martin Eichenbaum, Marjorie Flavin, Bob Hall, James Hamilton, Simon Gilchrist, Fatih Guvenen, Bryan Kelly, Ralph Koijen, Arthur Korteweg, David Lagakos, Hanno Lustig, Serdar Ozkan, Chris Parsons, Valerie Ramey, Martin Schneider, Frank Schorfheide, Eric Swanson, Rossen Valkanov, Annette Vissing-Jorgensen, Johannes Wieland, and Motohiro Yogo for helpful discussions and comments. All errors are my own.
1 Introduction

Rare events can play an important role in asset pricing despite their relative infrequency. For example, in theoretical models with macroeconomic disasters (e.g., wars, depressions, financial crises), tail risk is an order of magnitude more important than day-to-day fluctuations in consumption, dramatically amplifying risk premia—i.e. the required rate of return on risky assets. However, aggregate tail risk is hard to quantify, precisely because these events are so rare.

This paper considers an economic environment in which tail events are cross-sectional, rather than aggregate, phenomena. Discrete events, such as transitions between jobs, can induce changes in labor income that are large, highly persistent, and largely uninsurable, effectively exposing labor market participants to idiosyncratic tail risk. I make the case for state-dependent, idiosyncratic tail risk as a key driver of asset pricing dynamics. Similar to aggregate tail events, these shocks can have a first-order effect on household welfare, and, in turn, asset prices. Unlike aggregate tail risk, large panels of earnings data imply that idiosyncratic tail risk is essentially observable, making it easier to estimate and falsify the model.

I propose a model in which predictable changes in labor market event risk induce predictable changes in risk premia. Empirically, I develop a new measure of idiosyncratic tail risk, which is constructed using new data on the cross-sectional distribution of labor income growth rates. Labor market event risk is persistent, highly cyclical, and quantitatively large. Theoretically, I embed uninsurable, idiosyncratic jumps (infrequent, large shocks) into a state of the art endowment asset pricing framework. A calibrated model shows that plausible labor market event risk can generate a large, time-varying equity premium consistent with the data, even though aggregate consumption is essentially unpredictable. In addition, I show that a number of interactions between idiosyncratic risk and stock returns are consistent with the model’s predictions and, motivated by the theory, identify a new, robust predictor of stock returns.

The basic intuition for the mechanism is as follows. As the economy contracts, the left tail of the cross-sectional distribution of income growth rates becomes fatter and the right tail becomes thinner. Large, negative shocks loom larger and lucrative outside job offers dry up. This cyclical tail risk effectively concentrates the effects of business cycles, ex post, among a small fraction of the population for whom recessions have “disastrous” consequences. As in Mankiw (1986), this concentration mechanism makes agents care much more about economic downturns than they would if idiosyncratic labor market shocks were insurable. Stocks are a very poor hedge against idiosyncratic tail risk, rationalizing a large risk premium. Moreover, investors demand higher compensation for investing in stocks when uncertainty about future idiosyncratic risk increases, generating time variation in the risk premium, return predictability, and excess volatility.
In my model, two primary ingredients, recursive preferences and persistent, idiosyncratic jump risk, combine to dramatically amplify the equity risk premium. Since idiosyncratic tail events are very large and uninsurable, investors are willing to pay a large premium to hedge against these risks, even though they are fairly unlikely. Further, given the high persistence of idiosyncratic risk in the data, investors with Epstein-Zin preferences require high compensation for investing in assets, such as stocks, whose prices tend to fall in response to bad news about future idiosyncratic risk. In my stylized model, these hedging demands are extremely important quantitatively.

The paper proceeds in four stages. First, I provide new evidence of state-dependence in the tails of the cross-sectional distribution of labor income growth rates. My analysis builds upon statistics from Guvenen et al. (2014c) which are calculated from panel administrative earnings data. While the center of the cross-sectional earnings growth distribution is insensitive to the business cycle, its tails are highly responsive. Moreover, changes in aggregate wages appear to be primarily driven by these changes in the tails.

I develop a method which allows me to estimate the higher frequency (quarterly) dynamics of idiosyncratic risk from cross-sectional moments measured at lower frequencies. Using this mixed-frequency approach, I extract an empirical proxy for the conditional skewness of the idiosyncratic income growth distribution from a large cross section of macroeconomic and financial time series. This procedure yields a quarterly index capturing the level of idiosyncratic risk at a point in time, which I also use to calibrate a parametric model with time-varying labor market event risk. The data indicate that idiosyncratic tail risk is quantitatively large, highly persistent and cyclical, and it exhibits substantial time series variation, even in periods without recessions.

Second, I propose a tractable asset pricing framework which integrates heterogeneous agents, incomplete markets, and state-dependent cross-sectional consumption moments into a Lucas (1978) endowment economy. The key mechanism is that the shape of the distribution of idiosyncratic shocks to consumption growth is linked to aggregate state variables. I solve the model with arbitrary, jump diffusion aggregate cash flow dynamics (stochastic volatility and time-varying, compound Poisson jumps), following Drechsler and Yaron (2011). Symmetrically, idiosyncratic shocks have state-dependent Gaussian and jump components. These idiosyncratic jump shocks provide an analytically tractable way of capturing infrequent, large changes in consumption. As in Constantinides and Duffie (1996), agents are ex-ante identical and ex-post heterogeneous, which preserves the analytical tractability of a representative agent economy.

To complement my solution for a fully-specified endowment-based model, I also derive an intertemporal capital asset pricing (ICAPM) representation of the stochastic discount factor from my incomplete markets economy, generalizing a recent contribution by Campbell et al. (2014). This representation reveals that, in addition to several risk factors which also appear in repre-
sentative agent models, news about contemporaneous and future idiosyncratic risk are priced risk factors. The contemporaneous covariance between returns and idiosyncratic risk measures has received virtually all of the attention in the extant literature. When idiosyncratic risk is fairly persistent, news about future idiosyncratic risk likely carries a substantially higher weight.

Third, I illustrate the quantitative importance of idiosyncratic tail risk in affecting the dynamics of risk premia within a stylized model. The novel mechanism is that agents are exposed to rare, idiosyncratic disasters, where the idiosyncratic disaster probability persistent and time-varying. While the structure of the model resembles that of the Bansal and Yaron (2004) long-run risk model, the state variables in my model are considerably less persistent and aggregate consumption growth is essentially unpredictable. Risk premia are not driven by ultra-persistent state dynamics; instead, the presence of idiosyncratic disaster risk and incomplete markets amplifies the risk premium. My model, whose key parameters are disciplined by the data, matches a number of key asset pricing moments reasonably well; the equity premium is large (6.5%) and time-varying, and stock returns are excessively volatile.

In my theoretical framework, risk premia dynamics are linked to uncertainty about future investment opportunities. Therefore, in order for idiosyncratic risk to affect time-variation in the risk premium (return predictability), it must be the case that uncertainty about future idiosyncratic risk is time-varying. This necessary condition is indeed satisfied by the data. Initial claims for unemployment—the rate of involuntary job loss in the U.S. private sector—is a good proxy for the conditional volatility of my idiosyncratic risk index. Thus, if time-variation in idiosyncratic risk is a key driver of risk premium dynamics, then changes in labor market uncertainty should be accompanied by changes in the equity risk premium.

The last section tests this implication of the theory, as well as several others, by providing new evidence about the interactions between idiosyncratic risk and stock returns. My findings are generally consistent with the theory. First, my uncertainty measure is a robust predictor of broad market returns. Over the 1967-2012 sample, initial claims produces more accurate forecasts of returns than a number of conventional state variables from the literature on return predictability, including the dividend yield, the book-to-market ratio, the earnings-price ratio, and the default yield. Furthermore, stock returns are highly informative about future labor market conditions—i.e., the level of and uncertainty about idiosyncratic risk—whereas they convey less information about aggregate consumption growth.

While I emphasize labor income throughout, my asset pricing framework provides a tractable way of pricing risks associated with the redistribution of wealth more generally. These shocks could also come from households’ idiosyncratic exposures to firms’ capital—e.g. private equity and entrepreneurial investments (Heaton and Lucas (2000) and Moskowitz and Vissing-Jorgensen
In my framework, redistribution risk varies over time and enters as a priced state variable. Moreover, this incomplete markets mechanism, which is largely absent in production-based asset pricing models, is likely to generate an amplified response to aggregate shocks. If unfavorable redistributions become more likely when productivity is low and/or uncertainty is high (e.g., because default risk is higher), the associated increase in discount rates will affect firms' incentives to invest.

**Related Literature.** This paper lies at the intersection of literatures in finance, macroeconomics, and labor economics. It relates most closely to a literature on asset pricing with incomplete markets, building on the seminal contributions of Mankiw (1986) and Constantinides and Duffie (1996). In these models, uninsurable idiosyncratic risk concentrates aggregate shocks, ex post, among a small fraction of agents, amplifying risk premia. My theoretical model embeds such a concentration mechanism within a more general, dynamic environment. In response to the early literature, some authors question the quantitative importance of the mechanism (see e.g., Lettau (2002) and Cochrane (2008)). In my model, the combination of persistent idiosyncratic jump risk and the additional hedging demands associated with Epstein-Zin preferences yields a different conclusion.

This paper also relates to an extensive literature, both empirical and theoretical, studying time-variation in the equity risk premium. A large empirical literature provides evidence that stock returns are predictable by a variety of macroeconomic and financial indicators, though there is some debate about whether this predictability can be exploited out-of-sample. I provide a new empirical proxy for the risk premium, initial claims for unemployment, and document that it outperforms a number of popular candidates from the literature, particularly at short horizons. In my theoretical model, stock returns are predictable, and the dynamics of the risk premium are calibrated to match the dynamics of this empirical proxy.

My theoretical model combines elements from two families of representative agent asset pricing models, which provide potential mechanisms to explain both the level and dynamics of the equity risk premium. One mechanism, first introduced by Rietz (1988) and Barro (2006), emphasizes the risk of rare macroeconomic disasters. The second mechanism is long-run risk, first proposed by Bansal and Yaron (2004). These models combine Epstein-Zin preferences with

---

1See also Cogley (2002), Constantinides (2002), Krebs (2003), Krebs (2004), De Santis (2005), Krebs (2007), and Storesletten et al. (2007), among others.
2Angeletos (2007) and Toda (2014a, b) appear to be the first to introduce these preferences into the incomplete markets setting.
3See, e.g., surveys in Campbell (2000) and Rapach and Zhou (2013).
4See Gourio (2012), Gourio (2013), Gabaix (2012), Wachter (2013), and Tsai and Wachter (2013). A primary critique is that key parameters are challenging to estimate (Julliard and Ghosh (2012)) particularly when the disaster probability varies over time, though Backus et al. (2011) and Kelly and Jiang (2014) propose potential identification strategies.
a highly persistent, heteroskedastic component in expected aggregate consumption growth.\textsuperscript{5} I show that predictability in the higher moments of the idiosyncratic shock distribution—the idiosyncratic disaster probability—affects asset prices in a very similar manner as the long-run risks (predictability in the mean) emphasized in the representative agent literature. However, these higher moments are easier to identify from the data.

A growing body of empirical literature documents that households face a substantial amount of idiosyncratic labor income and/or consumption risk, as well as its time variation.\textsuperscript{6} One strand focuses on job displacement risk—i.e., large, highly persistent, and uninsurable declines in income which are often linked to the extensive margin: see e.g. Krebs (2007) and Davis and Von Wachter (2011). Also important are situations where a worker voluntarily switches jobs, presumably because she receives a better outside offer, as occurs in on-the-job search models. Both types of “idiosyncratic tail events” can result in large, persistent, and uninsurable changes in income, potentially having disproportionate impacts on welfare even though their realizations only hit a small fraction of households each period.

Theoretical motivations for labor market event risk are plentiful. In search models, virtually all variation in labor income comes from transitions between jobs.\textsuperscript{7} Altonji et al. (2013) also emphasize transitions as a source of lifetime earnings dispersion. Berk et al. (2010), Berk and Walden (2013), and Lagakos and Ordoñez (2011) provide models in which firms offer partial insurance to workers, causing wages to optimally move less than one-for-one in response to productivity shocks.\textsuperscript{8} In Berk et al. (2010), financial distress can cause this insurance to break down, concentrating losses among workers who switch jobs. Further, partial insurance will imply that firm profits fall when labor market event risk is high.

Two contemporaneous working papers also study asset pricing implications of state-dependent idiosyncratic risk. Both provide empirical evidence consistent with the mechanism and calibrate theoretical models with Epstein-Zin preferences which are nested within my affine framework. Constantinides and Ghosh (2014) use data from the consumer expenditure survey (CEX) to show that the skewness of household consumption growth is cyclical, complementing earlier work by Brav et al. (2002). They solve a model with a high, time-varying risk premium in which aggregate consumption and dividend growth are i.i.d but the higher moments of household consumption growth are persistent. They also present qualitative evidence that household

\textsuperscript{5}See Bansal et al. (2012) for a discussion of multiple extensions of the long-run risk model.

\textsuperscript{6}For labor income, see Storesletten et al. (2004), Mckay and Papp (2012), Huggett and Kaplan (2013), Guvenen et al. (2014a), and Guvenen et al. (2014c), among many others. For consumption, see Deaton and Paxson (1994), Brav et al. (2002), Vissing-Jorgensen (2002), Blundell et al. (2008), Malloy et al. (2009), Ludvigson (2013), Constantinides and Ghosh (2014), and Heathcote et al. (2014).

\textsuperscript{7}Kuehn et al. (2013) and Hall (2014) both discuss interactions between labor market search and asset prices. My model gives a reduced-form way of studying asset pricing implications of search frictions in incomplete markets.

\textsuperscript{8}See Guiso et al. (2005) for empirical evidence on risk-sharing in the labor market.
skewness measures are priced in the cross-section. Importantly, whereas Constantinides and Ghosh (2014) use unconditional moments of the cross-sectional consumption growth distribution in their calibration, mine also emphasizes conditional moments—i.e., time-series dynamics.\(^9\) Whereas my income-based measures exhibit substantial persistence at business cycle frequencies, the CEX-based measures exhibit little autocorrelation.

Herskovic et al. (2014) identify a common factor in the idiosyncratic volatility of firm-level shocks and demonstrate that this common component is priced in the cross-section of stock returns. They argue that shocks to idiosyncratic firm volatility are correlated with shocks to idiosyncratic household income volatility, rationalizing a risk premium.\(^10\) They replicate these cross-sectional asset pricing patterns in a model with incomplete markets where the volatility of idiosyncratic consumption growth shocks is persistent and countercyclical.

A key difference between these papers and mine is the manner in which the idiosyncratic shock process is tied to the data. My income process parameters and state dynamics are chosen to directly target the micro-level evidence. I provide additional time-series evidence consistent with the mechanism, complementing the cross-sectional evidence in both papers. By working within a general affine environment, I uncover several new insights about the deep structure of models within this class. First, I show that, for purposes of asset pricing, all assumptions about the distribution of idiosyncratic shocks can be summarized by a cross-sectional certainty equivalent. Risk premia depend in part on the covariance between returns and this certainty equivalent, which takes an analytically tractable affine form in my model. Second, my ICAPM representation illustrates the effect of state-dependent idiosyncratic risk on agents’ hedging demands, a key driver of the quantitative performance of the model with Epstein-Zin preferences.

Moreover, while both papers present calibrated models which generate a market risk premium consistent with the data, the mechanisms are somewhat different from my paper. Herskovic et al. (2014) match the level of the risk premium on the market portfolio via aggregate consumption risk; the first order effect of uninsurable idiosyncratic risk is to help explain cross-sectional differences in risk premia across sorted portfolios.\(^{11}\) In Constantinides and Ghosh (2014), the primary source of the risk premium is low-frequency variation in discount rates which is induced

---

\(^9\)These empirical findings complement mine, which are derived from the cross-sectional distribution of income growth rates. While CEX consumption measures have a more direct link with the theory, they require the use of much smaller sample sizes, a non-trivial concern given the small sample properties of higher moment estimates—see, e.g., Kim and White (2004). Also, survey-based consumption measures are more susceptible to measurement errors relative to administrative earnings data.

\(^{10}\)Their leading measure of income risk is the change in the dispersion (measured as the difference between the 90th and 10th percentiles) of year-on-year income growth rates from Guvenen et al. (2014c), whereas I emphasize the asymmetry of the idiosyncratic income growth distribution.

\(^{11}\)The authors explain: “Because the exposure of the aggregate market portfolio to the [idiosyncratic volatility factor] is close to zero in the data, most of the equity risk premium on the market portfolio must come from the standard consumption-CAPM term.”
by very persistent changes in idiosyncratic risk. As such, the way in which idiosyncratic risk affects asset prices is quite similar to the time preference shocks in the model of Albuquerque et al. (2014).\footnote{Depending on the calibration, the half-life of a shock to the state variable governing idiosyncratic risk varies from 4.1 to 6.8 years in Constantinides and Ghosh (2014). In my calibration, the comparable half-life is 1.4 years.} In my calibration, a large risk premium obtains with substantially less persistent shocks because of predictability in aggregate dividend growth, which causes stock prices to fall in response to bad news about idiosyncratic risk.

2 The Evolution of Idiosyncratic Risk Over Time

Here, I provide new evidence about idiosyncratic labor income risk and its evolution over time. Section 2.1 describes the cross-sectional statistics from Guvenen, Ozkan, and Song (2014c, hereafter “GOS”), which are key inputs to my analysis. Then, section 2.2 combines these data with a large cross-section of macroeconomic time series to calculate an idiosyncratic skewness index. Next, section 2.3 use this index to fit a parametric model with labor market event risk, which fits the data well and provides a rich picture of idiosyncratic risk dynamics, giving a sense for its magnitude in the data. Finally, section 2.4 provides evidence that a necessary condition for idiosyncratic risk to affect risk premium dynamics (i.e. return predictability) is satisfied, namely, that uncertainty about idiosyncratic risk is time-varying. In particular, initial claims for unemployment is a very good proxy for uncertainty about labor market conditions.

2.1 The Social Security Income Data

My analysis relies upon two sources of cross-sectional information from GOS. GOS report a number of statistics for the cross-section of real income growth rates to provide nonparametric evidence on its evolution over time. These statistics are calculated from a nationally representative sample of panel earnings records for 10% of males aged 25-60 in the U.S. population from the Social Security Administration (“SSA”, 163 MM observations). The data provide uncapped (i.e. not top-coded), nominal annual earnings for each individual from 1978-2011, which are adjusted to real terms using the personal consumption expenditure deflator.\footnote{According to GOS, the earnings data “include wages and salaries, bonuses, and exercised stock options as reported on the W-2 form (box 1).” See GOS for additional details about the data and sample selection criteria.}

The first source is time-series data (annual) with the variance, skewness, and several quantiles of the cross-sectional distribution of labor income growth rates at 1 and 5-year horizons. These statistics pool all earners in the sample, giving a snapshot of the entire cross-sectional distribution of wage changes across the U.S. population as it evolves over time. Taking the 1-year skewness
measure as an example, the first data point is the skewness of all real income growth rates from 1978-1979, the second uses growth rates from 1979-1980, and so on. These data reveal substantial cyclical variation in skewness/asymmetry of the distribution, while the variance of idiosyncratic shocks is almost acyclical—a highly robust result.\footnote{GOS show that variance/scale measures are approximately constant and skewness measures are time-varying, even after controlling for cohort and life-cycle effects, age, and lagged income.} Section 2.2 uses these data to extract a quarterly skewness index from a large cross-section of macroeconomic variables.

Second, GOS provide additional statistics describing the shape of the income growth distribution after conditioning on prior earnings. GOS sort individuals into 100 different bins on the basis of lagged earnings over the prior 5 years. Within each bin and for each year in their sample, GOS calculate a variety of quantiles and moments for the cross-sectional distribution of income growth rates, again at 1 and 5-year horizons. They report averages of these statistics for each bin in expansion periods and recession periods, respectively.

Given my focus on asset pricing, the nature of idiosyncratic risk faced by relatively high earners—i.e. those who are likely to participate in financial markets—is of particular interest. My calibration targets averages of GOS’s statistics over the 91st through 95th percentiles of the lagged earnings distribution. These individuals have sufficiently high earnings that they are likely to participate in stock markets. However, labor income is still likely to be their primary source of non-housing wealth.\footnote{In 2011, GOS report that the 90th and 95th percentiles of wages are $98k and $135k per year, respectively, in 2005 dollars. GOS also emphasize the high cyclicality of the incomes of extremely high earners, particularly those in the top 1%. See also Guvenen, Kaplan, and Song (2014a). While the higher cyclicality is interesting, these individuals likely possess substantial financial wealth, making it difficult to characterize the extent to which income shocks translate to consumption. For example, existing evidence on partial insurance is unlikely to be representative of their behavior.} Appendix A.1 shows that my conclusions are relatively insensitive to the specific choice of the endpoints.

The GOS data reveal a number of insights about this risk, which I discuss in greater detail in Appendix A.1. First, high earners face a substantial degree of idiosyncratic labor income risk, even in expansions. Second, the entire income growth distribution shifts to the left in bad times; all reported quantiles are strictly lower in recessions relative to expansions. Third, the center of the distribution (e.g. the median income growth rate) is relatively insensitive to the business cycle, but its conditional tails are highly state-dependent. In recessions, the left tail of the income growth distribution becomes fatter, while its right tail shrinks, suggesting that, ex post, aggregate shocks are disproportionately borne by those who experience large shocks. Section 2.3 replicates these features using a model with idiosyncratic, state-dependent jump risk.
2.2 Conditional skewness index

GOS emphasize changes in skewness in recession years relative to expansion years. However, asset pricing tests tend to be conducted using high frequency (e.g. monthly or quarterly) data. In order to better assess the potential linkages between labor income risk and asset pricing dynamics, it is helpful to have a more continuous, higher frequency notion of idiosyncratic risk, particularly since recessions need not coincide neatly with calendar years. Here, I extract an empirical proxy for the conditional skewness of the idiosyncratic income growth distribution from a large cross section of macroeconomic time series. My skewness index is available at a quarterly frequency and is available over a longer time period, making it easier to understand its time series properties. Below, I calibrate my theoretical model to match its dynamics.

I use a mixed-frequency approach to estimate a quarterly skewness index using cross-sectional moments from the SSA data which are time-aggregated and only available on an annual basis. I begin with a quarterly, reduced form model for the data-generating process of labor income. Then, I derive the third cross-sectional moments of idiosyncratic wage changes, where wages are measured annually. This method yields a regression equation which identifies the third moment of idiosyncratic shocks in the quarterly model--my skewness index--semi-parametrically.

My quarterly model for labor income is a generalized version of the canonical permanent income model. Let $w_i^t$ and $age_i^t$ be individual $i$’s log labor income (after subtracting common shocks and a deterministic life-cycle component) and age in quarter $t$, respectively. I assume

$$
w_i^t = \underbrace{\xi_i^t}_{\text{permanent component}} + \underbrace{\alpha_i + \beta_i age_i^t}_{\text{profile heterogeneity}} + \underbrace{\rho(L) \cdot \eta_i^t}_{\text{state dependent transitory shock}} + \underbrace{\epsilon_i^t}_{\text{state independent transitory shock}} \quad (1)
$$

$$
\xi_i^t = \xi_i^{t-1} + \eta_i^t, \quad \eta_i^t | y_t \sim F_{\eta}(\eta; y_t), \quad \epsilon_i^t \perp y_t, \quad (\alpha_i, \beta_i) \overset{iid}{\sim} G(\alpha, \beta), \quad (2)
$$

$$
E[(\eta_i^t)^3 | y_t] = a + b' y_t \equiv a + x_t \quad (3)
$$

where $\eta_i^t$ is a shock that is independently distributed over time conditional on the aggregate state, a finite-dimensional, observable random vector, $y_t$. $\xi_i^t$ is a permanent component to wages. As I discuss below, my theoretical model requires that $\xi_i^t$ be a random walk. An alternative would be for $\xi_i^t$ to follow a persistent, stationary process such as an AR(1). I impose the random walk restriction throughout, noting that estimates of the AR(1) parameter in other studies are generally close to 1.\textsuperscript{16} $\alpha_i$ and $\beta_i$ allow for heterogeneity in income levels and growth rates and are drawn from a bivariate distribution with mean zero and finite third moments.

\textsuperscript{16}GOS estimate an version of this model with an AR(1) persistent component, albeit with different distributional assumptions, and obtain an annual autocorrelation coefficient of 0.979. An alternative specification yields an estimate of 0.999. However, introducing profile heterogeneity could potentially lead to lower estimates.
I also allow for a transitory component in labor income. The first term, which can depend on current and past $\eta^i_t$ via the lag polynomial $\rho(L)$, allows permanent shocks to have additional temporary effects on measured income. For example, large negative realizations of $\eta^i_t$ could be accompanied by unemployment spells, leading to temporary interruptions in the flow of labor income. The second term, $\epsilon^i_t$, is a mean zero transitory component that is stationary and independent of the aggregate state. While it is straightforward to also allow for state dependence in the distribution of $\epsilon^i_t$, I maintain this assumption for parsimony.

Finally, I assume that the third moment of $\eta^i_t$ is an affine function of $y_t$. A sufficient condition is that the cumulant-generating function (the log of the moment-generating function) of $\eta^i_t$ is linear in $y_t$. Most distributions in theoretical asset pricing models satisfy this condition, since it often leads to exponential affine solutions for prices, facilitating analytical tractability. Two leading examples are compound Poisson processes with time-varying arrival intensities and gamma random variables with time-varying shape parameters. Their cumulant-generating functions are affine in these time-varying parameters. Independent sums of these variables also satisfy this property. For example, if $\eta^i_t$ has a compound Poisson distribution with intensity $\lambda_t = \lambda_0 + \lambda_1 y_t$, my regression recovers $\lambda_1$ up to a constant of proportionality. This distribution provides an analytically tractable way to represent infrequent, large shocks (“jump risk”).

Next, I derive a mixed-frequency representation which enables me to estimate $b'y_t$—my quarterly skewness index, $x_t$. Define $W^i_{A,t} = \sum_{j=0}^{3} \exp(W^i_{t-j})$ and $w^i_{A,t} \equiv \log W^i_{A,t}$. $W^i_{A,t}$ is a trailing four quarter moving average of labor income, which is only observed in the fourth calendar quarter of each year. In this notation, the year-on-year change in annual log wages is $w^i_{A,t} - w^i_{A,t-4}$. Finally, $\mathcal{F}_t$ is a filtration capturing aggregate information up to time $t$, including $\{y_{t-j}\}_{j=0}^{\infty}$.

**Proposition 1.** Let $\{w^i_{t}\}_{t=0}^{\infty}$ be generated as in (1-3). Then, for $k \geq 4$,

$$E\left[\left(w^i_{A,t} - w^i_{A,t-k} - E[w^i_{A,t} - w^i_{A,t-k} | \mathcal{F}_t]\right)^3 | \mathcal{F}_t\right] \approx c_k + b'\phi_k(L; \rho)y_t, \quad (4)$$

where the coefficients $c_k$ and the known lag polynomial $\phi_k(L; \rho)$ are defined in Appendix A.2.1.

The proof makes use of a simple log-linear approximation for time-aggregated wages, which replaces an arithmetic mean with a geometric mean. Simulation results, which are available upon request, demonstrate that approximation errors are negligible for the parametric model in Section 2.3, particularly over longer horizons. Henceforth, I ignore these errors.

---

17 Signing bonuses could generate similar effects for large positive shocks.

18 My aggregation result goes through if, given the aggregate state, the third central moment of $\epsilon^i_t$ is constant.

19 Otherwise, one could still motivate my specification using standard linear projection arguments.

Proposition 1 says that, when the third central moment of the permanent shock $\eta_i^t$ is affine in $y_t$, I can recover $b$ semi-parametrically with a regression. Given annual data on third central moments of the cross-section—I use the 1 and 5-year measures from GOS—and a model for $\rho(L)$, $b$ is the vector of slope coefficients from a regression of the third moment on a constant and $\phi_k(L; \rho)y_t$. Further, I can pool information from different horizons $k$, since the coefficients $b$ are the same for all $k$. My estimation assumes $\rho(L)$ has a restricted MA(1) structure: $\rho(L) = \rho \cdot [1 + L]$. This specification implies that permanent shocks can have additional temporary effects that last about 6 months, though I find similar results with different lag lengths.

Next, I must specify the state variables $y_t$ to include, as well as an estimation method. I explore two approaches. One option is to estimate $b$ directly via GMM (OLS when $\rho$ is known, non-linear least squares otherwise). The length of the sample precludes the estimation of a large number of coefficients, so the risk of overfitting is nontrivial. Thus, such an approach works well only if the dimension of $y_t$ is relatively low. Appendix A.2.2 adopts this approach, estimating univariate and bivariate specifications involving several theoretically-motivated regressors.

My preferred approach estimates (4) using statistical methods which are designed to provide optimal forecasts in a data-rich environment—situations where the number of predictors is large (potentially much larger) relative to the sample size for the forecast target. These methods enable the researcher to exploit the rich information in a large cross-section of predictors while substantially reducing the risk of overfitting the data. Relative to principal components methods, these techniques are more efficient in the presence of irrelevant factors—that is, factors that explain cross-sectional variation among predictors that are uncorrelated with target variable. I adapt the Three-Pass Regression Filter (3PRF) method of Kelly and Pruitt (2014) to extract the optimal linear predictor from a large number of macroeconomic and financial time series.

I include 109 quarterly time series in the vector $y_t$, all of which are available from 1960-2013. I begin with 97 monthly macroeconomic and financial time series considered in Bernanke et al. (2005). I augment these series with 12 additional variables from the literature on stock return predictability, which are updated regularly by Goyal and Welch (2008). I then construct quarterly series by averaging the underlying monthly series within each calendar quarter. To allow for potential lead-lag relationships, I additionally include several lags and (weighted and unweighted) moving averages of each variable. Further details are in Appendix A.2.3.

Table 1 documents the goodness-of-fit of the overall index, which is constructed using the 3PRF, as well as other indices which are constructed using different subsets of the 109 variables in $y_t$.21 I obtain the data from Global Insight and transform them as in Bernanke et al. (2005) to ensure stationarity. Of the 120 variables included in Bernanke et al. (2005), Wu and Xia (2014) identify a subset of 97 variables which are available through the end of 2013. Wu and Xia (2014) also verify that the findings of Bernanke et al. (2005) are replicable using the subset with available data.

---

21I obtain the data from Global Insight and transform them as in Bernanke et al. (2005) to ensure stationarity. Of the 120 variables included in Bernanke et al. (2005), Wu and Xia (2014) identify a subset of 97 variables which are available through the end of 2013. Wu and Xia (2014) also verify that the findings of Bernanke et al. (2005) are replicable using the subset with available data.
These subindices are constructed using an alternative method for estimating linear models in a data-rich environment: forecast combination methods—e.g. Timmermann (2006). The 3PRF uses a series of regressions to optimally combine the information from each univariate model, whereas the forecast combination approach takes a weighted average of the univariate models, using the inverse of the mean-squared error (IMSE) as weights.\(^{22}\)

Looking at the first row of Table 1, the overall performance of my skewness index is quite strong. A single factor extracted from the 109 variables achieves \(R^2\)’s of 70% and 81% at the 1-year and 5-year horizons, respectively. My estimate of \(\rho\), which governs transitory risk, is 0.45, suggesting that a 10% decline in permanent income in quarter \(t\) is associated with an additional 4.5% transitory decline in income in quarters \(t\) and \(t + 1\). For additional discussion of the role of state-dependent transitory risk, which much easier to see within the context of the univariate and bivariate specifications explored there, I refer the interested reader to Appendix A.2.2.

The next row of Table 1 reports the performance of an alternative index, which is constructed using IMSE weights. The \(R^2\)’s decline somewhat to 66% and 72% at the 1-year and 5-year horizons, respectively, which is perhaps unsurprising because the combination methods trades some robustness for efficiency. Interestingly enough, however, the time-series properties of skewness index from the IMSE combinations are virtually identical to those of the 3PRF estimates. The correlation between the two quarterly indices, which is reported in the last column, is 99.6%.

Figure 1 plots the estimated quarterly skewness index, \(x_t\), obtained using the 3PRF, and the alternative combination estimate. Both measures are visually indistinguishable from one another and highly cyclical, peaking in expansions and bottoming out during (or immediately after) recessions. Note however that skewness dynamics appear to considerably richer than the two-state (expansion and recession) Markov process typically assumed in the empirical literature on estimating earnings process.\(^{23}\) Moreover, the measures are quite persistent, exhibiting substantial variation at business cycle frequencies. The first-order autocorrelation, when expressed as a monthly number, is around 96% for each measure.

In addition to the overall skewness indices, Figure 1 plots fitted values from the 40 best-performing univariate regression models, which are shaded from light to dark according to goodness-of-fit. Darker lines indicate better fit. These univariate forecasts are highly correlated with one another and generally track the overall indices quite closely, indicating a strong factor

---

\(^{22}\)Kelly and Pruitt (2014) prove that the 3PRF is consistent as the number of predictors and time series observations go to infinity. The efficiency gains associated with the 3PRF (which estimates a cross-sectional regression for each time series observation) require a large number of predictors. With the smaller number of variables for the subindices, the combination methods (which estimate fewer parameters) perform better.

\(^{23}\)In unreported results, I find that a quarterly NBER recession indicator has little explanatory power for the time aggregated skewness measures. However, such a result could partially reflect a timing mismatch between NBER recession dates and labor market peaks and troughs, which can lag the business cycle.
structure in the data. Note that the estimation sample begins in 1978, while the overall and univariate skewness measures track one another quite closely prior to 1978, suggesting that the identified factor is a genuine feature of the data, rather than an artifact of over-fitting.

Turning back to Table 1, I find that the skewness index loads most heavily on measures of real activity, providing strong empirical support for the concentration mechanism in Mankiw (1986). Indices constructed using two subcategories have a correlation of 96% with the overall index. The first, real output and income, includes a number of industrial production indices and measures of total household income. The second category, employment and hours, primarily reflects information about the extensive margin in the labor market (employment growth, the unemployment rate, and the distribution of unemployment durations). Next, indices constructed from more forward-looking measures—real inventories, orders, and unfilled orders—have an 89% correlation with the overall index. All achieve $R^2$'s which are slightly inferior, but generally comparable with the performance of overall combination forecast.

After these initial measures, my skewness index loads most heavily on financial variables. An index constructed using realized stock returns and the Goyal and Welch (2008) predictors has a 69% correlation with the overall index. Moreover, this subindex outperforms any of the combination indices (including the overall measure) in tracking 5-year idiosyncratic skewness. Indices constructed using money and credit quantity aggregates and interest rates perform reasonably well and have correlations of 64% and 62%, respectively, with the overall index.

Perhaps more surprisingly, an index constructed using aggregate consumption measures does not capture the variation in idiosyncratic skewness relative to other measures of real activity. The consumption-based subindex achieves a modest $R^2$’s of 29% and 5% at the 1-year and 5-year horizons, respectively, though the consumption-based index maintains a 50% correlation with the overall skewness index. Much of the disconnect in performance relative to other economic activity measures is certainly due to the substantial measurement errors in high frequency consumption data. In addition, such a result could reflect temporal instabilities in lead-lag relationships between aggregate consumption and other measures of real activity.

Of the remaining categories, none of the subindices is capable of explaining very much of the variation in idiosyncratic skewness. Particularly striking is the lack of explanatory power of average hourly earnings, which provides further evidence for a link between the extensive margin and the distribution of idiosyncratic shocks in the labor market.

---

24Recall that, so as to use the same data as in Bernanke et al. (2005), I construct quarterly series by averaging the monthly measures over each quarter. In Appendix A.2.2, I estimate univariate regressions using quarterly NIPA data on real consumption of nondurables and services. Using these data, aggregate consumption works somewhat better. However, its performance is inferior to that of other univariate indicators, such as real compensation growth, employment growth, or a measure of overall profitability of corporations.
2.3 Parametric model with labor market event risk

The analysis in the previous section was semi-parametric, allowing me to make statements about the time-series behavior of idiosyncratic skewness without distributional assumptions. To close a quantitative asset pricing model, such assumptions are required. This section adopts a more parametric approach, fitting a model with labor market event risk to the data. My discussion here gives a heuristic view of the model. Appendix A.3.1 discusses distributional assumptions, parameters, and calibration method in detail. I also show that the model matches the variation in cross-sectional skewness from Section 2.1 and the time-series dynamics in Section 2.2 well.

Maintaining the assumptions on the labor income process from above, state dependence manifests itself via the distribution of the permanent shock $\eta_i_t$. I assume that $\eta_i_t$ is the sum of three, mutually independent shocks: two state-dependent jump (compound Poisson) shocks with time-varying intensities and a normally-distributed state-independent neutral shock which hits every period. The first jump component is a good shock with a positive mean, capturing infrequent, large upward adjustments in consumption—“climbing the ladder”. These changes could come as a result of a promotion or the arrival of an attractive outside offer. The second jump component is a bad shock with a negative mean, capturing infrequent, large downward adjustments—“falling off the ladder”—likely driven by events such as job loss.

I incorporate state dependence by allowing the probability of experiencing a large shock (jump intensity) to vary over time. Empirically, I tie these intensities to my skewness index from Section 2.2. I assume that, as the skewness index increases, large positive and negative shocks become more and less likely, respectively. If the data are generated according to my parametric model, my regression procedure would provide a consistent estimate of these intensities up to location and scale. Finally, as is common in the asset pricing literature, I assume that the distribution of jump sizes is does not vary over time, which facilitates analytical tractability.

Figure 2, Panel A plots the evolution of the fitted probabilities of good and back shocks, respectively, over time. The probability of receiving a large shock is about 8% per year. According to the fitted model, bad shocks are almost always more likely than good shocks, though the difference between the two probabilities is relatively small in expansions. As the economy moves into a recession, the probability of receiving a bad shock increases substantially, while the probability of a good shock goes almost to zero. These probabilities remain elevated in the early part of the post-recession recovery, then revert back to lower levels.

The fitted good intensity goes slightly negative during the financial crisis. For purposes of calculating the model-implied moments, I truncate the bad intensity so that the sum of the two intensities is always constant. The dashed line shows the untruncated path of the bad
probability. These estimates suggest that, during the financial crisis, idiosyncratic risk reached unprecedentedly high levels relative to the rest of the period where the index is available. Therefore, from an incomplete markets perspective, the Great Recession could easily be considered a “disaster” in spite of the relatively moderate observed decline in aggregate consumption.25

Figure 2, Panel B plots the several quantiles of model-implied distributions of year-on-year income growth rates. These estimates condition on the observed trajectory of the skewness index. To emphasize changes in higher moments, I subtract the median from each quantile—i.e., I normalize it to zero. At a 1-year horizon, the central quantiles barely move at all. The interquartile range is essentially unchanged, while the extreme quantiles (2.5, 5, 95, 97.5) are highly state dependent. The 5th percentile varies between -40 and -50 log points, and the the 2.5th percentile fluctuates even more, ranging between -55 and -80 log points. These extreme quantiles move up and down together, increasing in expansions and falling significantly in recessions.

The magnitudes associated with the state-dependent shocks are extremely large. Incorporating Jensen’s inequality, the average decline in wages from a large negative shock is about 49%! The 5th percentile of wages25 move forward, recall that these estimated magnitudes are for declines in pre-tax labor earnings for a single individual. Associated declines in household consumption are likely to be smaller. In my quantitative model, I assume that less than 25% of income declines pass through to consumption, on the low end of estimates in Blundell et al. (2008) and Heathcote et al. (2014).

Figure 3 characterizes the densities of the permanent component of year-on-year changes in wages in expansions and recessions, respectively. To generate the figures, I randomly sample with replacement from the observed values of $x_t$ in expansion quarters and recession quarters, respectively, then plot the densities of year-on-year changes in the permanent component of wages (i.e., I strip out transitory shocks). I also allow for a time variation in the average logarithmic growth rate of labor income, which is assumed to be an affine function of the change in aggregate private sector real compensation, choosing the slope and intercept to exactly match the median in expansions and recessions from Table 9. I use a log scale on the vertical axis in order to better show the changes in the tails.26

Dashed vertical lines indicate the average change in log wages in expansions and recessions, respectively. Note that, despite the location shift, the densities are essentially indistinguishable from one another in the center of the distribution. However, the behavior of the tails is radically

25 Using NIPA data, I calculate that the peak-to-trough decline in quarterly real consumption of nondurables and services was approximately 1.6%, which was spread over a 4 quarter period. Some caution is necessary when drawing this conclusion, because my fitted 5-year skewness measures are more negative than their observed values in the GOS data during the crisis.

26 On this scale, the pdf of a normal distribution is a quadratic.
different. In expansions, the density is relatively close to symmetric, though the left tail is slightly fatter than the right tail. As the economy moves from an expansion to a recession, the left tail becomes fatter, while the right tail shrinks substantially. This evidence again suggests that observed changes in aggregate wages are almost exclusively driven by changes in the tails.

In conclusion, the evidence in this section suggests that the business cycle has a particularly strong impact on tails of the conditional distribution of idiosyncratic labor income growth distribution. This result suggests that aggregate shocks are far from equally distributed across households; instead, state-dependent shocks appear to be disproportionately borne by those who receive very large positive or negative shocks. I am able to replicate these features with a simple model where labor income is subject to idiosyncratic “jump risk”. Moreover, my observable proxy for idiosyncratic risk is highly persistent. Section 3 develops the asset pricing implications of a model which allows for all of these features.

2.4 Uncertainty about idiosyncratic risk is countercyclical

In my theoretical model, the risk premium is constant when shocks to the state vector are homoskedastic (Proposition 3). Idiosyncratic risk can only affect risk premium dynamics if uncertainty about future idiosyncratic risk is time-varying. In this section, I provide evidence that my skewness index, denoted by $x_t$, is heteroskedastic, and I identify three variables which capture uncertainty about future idiosyncratic risk.

My preferred uncertainty measure is initial claims for unemployment insurance ($\text{Claims}_t$). I divide the number of claims filed in each month by the size of the workforce (from the BEA) in the prior month to obtain a stationary measure. An individual is only eligible for UI if he/she becomes “unemployed through no fault of his/her own” (e.g. laid off).\textsuperscript{27} Thus, the normalized series may be interpreted as the rate of involuntary job loss in the cross-section of employed individuals in the private sector. Earlier, I demonstrated that my skewness index is most closely linked with measures of real activity and employment growth. Initial claims is a leading indicator of future labor market conditions (Barnichon and Nekarda (2012)), which is available on a very timely basis and is subject to little measurement error.

While initial claims plausibly proxies for expected future labor market conditions, it is reasonable to ask why it should proxy for uncertainty. One would expect more layoffs when aggregate productivity is low. An elegant justification for a link between aggregate productivity and labor market uncertainty comes from Ilut et al. (2014). Using establishment-level Census data, they find strong evidence that firm-level hiring and firing decisions respond more strongly to bad

\textsuperscript{27}Source: http://www.edd.ca.gov/unemployment/Eligibility.htm.
news relative to good news about productivity. Ilut et al. (2014) show that, when this condition holds, the conditional volatility of aggregate employment growth is higher in bad times (i.e. when average firm productivity is low), even if all productivity shocks are iid. The same condition also implies that cross-sectional dispersion of employment growth is countercyclical.

Second, filing a claim for unemployment insurance benefits is time-consuming. Individuals must fill out a lengthy application, and there is a waiting period while the UI benefits office verifies the reason for the separation with employers. If a worker is fairly certain that he/she will be able to find a job quickly, such a process may not be worth the effort. In contrast, when uncertainty about one’s future job prospects is high, the expected benefit from filing a claim is higher.

I also consider two alternative uncertainty measures. The first is, $x_{t-1}$, the lagged skewness index. If, for example, $x_t$ follows a square-root process, then its conditional mean and volatility are perfectly correlated with one another. The argument from Ilut et al. (2014) also applies to $x_t$. The second, $Vol_t$, is a measure of cross-sectional volatility of employment growth across states. I construct the cross-sectional measure using quarterly, seasonally-adjusted, state-level employment growth data from Hamilton and Owyang (2011).28

Figure 4 plots the three (standardized) measures, initial claims for unemployment ($Claims_t$), the skewness index ($x_t$), and $Vol_t$ over the time period for which both series are available. $x_t$ and $Claims_t$ are strongly negatively correlated and highly cyclical. Further, $Claims_t$ slightly leads $x_t$; bivariate Granger-causality tests provide strong evidence that $Claims_t$ Granger-causes $x_t$ when 2 or more lags are included in the VAR. The business-cycle frequency movements of all three series are quite similar, spiking sharply in recessions, though the cross-sectional measure differs somewhat during the double-dip recession of 1982.

Next, I show that all three measures are good proxies for time series uncertainty about future idiosyncratic risk. Table 2 presents the results from heteroskedasticity tests which are similar to the ARCH test of Engle (1982). I estimate AR(p) or VAR(p) models using $x_t$ and $Claims_t$, calculate the residuals, then regress absolute or squared residuals on each uncertainty measure. I whether slope coefficients in the second stage regression are zero.29 I also repeat the analysis using residuals for $Claims_t$. Rows correspond with different specifications for the conditional mean. The third column reports the first stage $R^2$’s for each conditional mean model, which are generally quite high. The remaining columns report Newey-West $t$-statistics on the slope coeffi-

---

28I am grateful to James Hamilton for making the data available. I extend the data to the present by aggregating the monthly, seasonally-adjusted series which are now provided by the Bureau of Economic Analysis. For each state, I estimate $\Delta emp_s,t = \alpha_s + \beta_s \Delta emp_t + \sigma_s u_{s,t}$, where $emp_t$ is the cross-sectional average employment growth. My uncertainty measure is the cross-sectional volatility of the fitted residuals: $Vol_t \equiv \sqrt{\frac{1}{S} \sum_{s=1}^{S} \hat{u}_{s,t}^2}$.

29The standard errors, as currently calculated, are conditional on the estimated first stage coefficients.
cient for each uncertainty measure.\textsuperscript{30} \( R^2 \)'s from these second-stage regressions are in brackets.

The results are qualitatively identical regardless of the specification considered. Increases in initial claims, decreases in the skewness index, and increases in cross-sectional employment dispersion are highly significant predictors of the volatility of both sets of residuals. The statistical tests also appear to be fairly insensitive to the use of absolute of squared residuals. In terms of \( R^2 \), lags of the skewness index and initial claims have roughly the same degree of explanatory power for \( x_t \) residuals. Turning to the bottom panel, \( \text{Claims}_{t-1} \) is also a good proxy for the volatility of its own residual, whereas the skewness index has less explanatory power. \( \text{Vol}_t \) also captures initial claims residual volatility reasonably well.

3 Asset Pricing Framework

In this section, I embed an endowment-based asset pricing model with heterogeneous agents and incomplete markets within a general affine, jump-diffusion framework. My general setup closely resembles the model in Toda (2014b), which builds heavily upon Constantinides and Duffie (1996). I place more structure on the stochastic environment, similar to Drechsler and Yaron (2011) and Eraker and Shaliastovich (2008), which leads to approximate analytical solutions. Section 3.1 describes the structure of the model. Section 3.2 presents the general equilibrium conditions, Section 3.3 presents analytical solutions to a log-linearized model, and Section 3.4 presents an ICAPM characterization of its stochastic discount factor. See Appendix B.1 for a more formal discussion of the model and the associated equilibrium.

3.1 General Structure

Time, indexed by \( t \), is discrete and there are an infinite number of periods. There is a continuum of infinitely-lived agents, indexed by \( i \in I = [0, 1] \). Agents choose consumption and savings to maximize lifetime utility, with identical Epstein and Zin (1989) and Weil (1989) preferences:

\[
U^i_t = \left( (1 - \delta)(C_t^i)^{1-1/\psi} + \delta(E_t[U^i_{t+1}]^{1-\gamma}) \right)^{1/(1-1/\psi)},
\]  

\textsuperscript{30}I use 4 lags, though results are insensitive to this choice.
where $\psi$ governs the elasticity of intertemporal substitution (EIS) and $\gamma$ is the coefficient of relative risk aversion.\(^{31}\) Each agent receives an endowment which evolves according to

$$\frac{C_t^i}{C_{t-1}^i} = \exp(\eta^i_t) \frac{C_t}{C_{t-1}}, \quad \Rightarrow \quad \Delta c_t^i = \Delta c_t + \eta^i_t, \quad E[\exp(\eta^i_t) | F_t] = 1 \quad (6)$$

where $C_t$ is the aggregate endowment and $\eta^i_t$ is an idiosyncratic shock which redistributes the aggregate endowment across agents.\(^{32}\) Agents can also invest in $K$ other financial assets in zero net supply, paying dividends ($D_{kt}$). As in Constantinides and Duffie (1996), shocks are structured so that no-trade is an equilibrium—i.e. agents choose to consume their endowments. These financial assets do not affect allocations, but they are priced by the model.

The endowment process (6) emerges naturally as a special case of the permanent income model from section 2.2 with no profile heterogeneity or transitory shocks. Ruling out profile heterogeneity is essentially without loss of generality, because differences across agents in profile heterogeneity are isomorphic to different initial endowments.\(^{33}\) While the elimination of transitory risk is a substantial departure from the data, estimates in Blundell et al. (2008) suggest that households smooth away virtually all transitory income risk. Further, existing representative agent results suggest that, when the EIS $> 1$, agents’ willingness to substitute over time means that transitory dynamics generally play a relatively minor role in affecting risk premia.\(^{34}\)

My general model for aggregate dynamics, formalized in Assumption 1 in Appendix B.1, summarizes the state of the economy by a $(L \times 1)$ vector, $y_{t+1}$. $y_{t+1}$ follows the stationary VAR,

$$y_{t+1} = \mu_y + F_y y_t + G_{y,t}^2 y_{t+1} + J_{y,t+1} \quad ,$$

whose innovation has both Gaussian and jump components. The variance-covariance matrix for the Gaussian shocks and the arrival intensities for the jump shocks can both depend on $y_t$.

Aggregate consumption and dividend growth are linear combinations of $y_{t+1}$: $\Delta c_{t+1} = S^c_{t+1} y_{t+1}$ and $\Delta d_{k,t+1} = S^k_{t+1} y_{t+1}$. This setup is quite flexible, encompassing the majority of cash flow dynamics in endowment-based asset pricing models—e.g. the Bansal and Yaron (2004) long-run risk model and the Wachtler (2013) time-varying disaster model.\(^{35}\)

\(^{31}\)Krebs (2007) and Toda (2014b) show how the assumption of infinitely-lived agents may be relaxed. Allowing for a constant probability of death each period is isomorphic to lowering the discount rate $\delta$.

\(^{32}\)As before, I denote levels with capital letters and logs with lower case letters, e.g. $c_t = \log(C_t)$.

\(^{33}\)If agents can write contingent claims on the aggregate state and have no borrowing constraints, they can effectively smooth out any deterministic differences in the growth rate of the endowment process.

\(^{34}\)See, e.g. Bansal et al. (2010) and Dew-Becker and Giglio (2013).

\(^{35}\)The disaster model obtains if I assume that both $S_a$ and $S_h$ have a common exposure to one of the Poisson jump components, whose probability varies over time.
Next, I parameterize the distribution of the idiosyncratic shock $\eta_{t+1}^i$ given $y_{t+1}$. $\eta_{t+1}^i$ is the sum of a finite number of independent Gaussian and jump components whose higher moments depend on $y_{t+1}$. As in Constantinides and Duffie (1996) and Storesletten et al. (2004), the volatility of the Gaussian component is potentially time-varying, and, as in my calibrated model of section 2.3, the poisson arrival rate for each jump component can depend on $y_{t+1}$. I also need to subtract off a location adjustment (a function of $y_{t+1}$) to ensure that the idiosyncratic shocks are truly idiosyncratic—i.e. satisfy the conditional moment restriction in (6). For a more precise statement of the shock structure, as well as a technical assumption about agents’ information, see Assumption 2 in Appendix B.1.

The most restrictive assumption required by the model is that the distribution of $\eta_{t+1}^i$ is identical across agents—i.e. all individuals face the same level of idiosyncratic risk. In the data, the idiosyncratic risk of unemployed individuals could be quite different from that of employed individuals. I rule out individual-specific state dependence. The model can capture unemployment-type events in reduced-form, except that it collapses all of the effects of unemployment into a single shock rather than allow for multiple shocks which unfold over time. It also abstracts away from heterogeneity in risks between individuals (e.g. skill heterogeneity, life cycle effects, etc.) and across industries/occupations, which would require a more complicated model.

Finally, I impose an affine structure on the distribution of shocks in the model, which is summarized by Assumption 3. For example, the idiosyncratic jump intensities are assumed to be affine in $y_{t+1}$, as was the case in section 2.3. These final restrictions ensure that, after performing the Campbell and Shiller (1988) approximation, the model generates valuation ratios which are exponential affine in the state vector $y_t$, leading to approximate analytical solutions.

3.2 Equilibrium Conditions

From Epstein and Zin (1989), equilibrium requires that, for any asset return $\tilde{R}_{t+1}$, each agent’s consumption profile satisfies the Euler equation:

$$1 = E_t \left[ \delta^\theta \left( \frac{C_{t+1}^i}{C_t^i} \right)^{-\frac{\theta}{\gamma}} (R_{c,t+1}^i)^{-(1-\theta)\tilde{R}_{t+1}} \right] = E_t \left[ \delta^\theta \left( \frac{C_{t+1}^i}{C_t^i} \right)^{-\gamma} \left( \frac{WC_{t+1} + 1}{WC_t} \right)^{-(1-\theta)} \tilde{R}_{t+1} \right], \tag{8}$$

where $\theta = \frac{1-\gamma}{1-1/\psi}$, $R_{c,t+1}^i \equiv \frac{W_{t+1}^i + C_{t+1}^i}{W_t}$ is the return of an (non-traded) asset delivering an arbitrary agent’s consumption stream, and $WC_t$ is the wealth-consumption ratio. The Constantinides and Duffie (1996) mechanism combines two key assumptions: permanent, proportional idiosyncratic shocks and homothetic preferences. Since agents’ endowments are i.i.d. in growth

---

36 A similar simplifying assumption is common in the literature on rare macroeconomic disasters.
rates, they always have the same first-order conditions. Then, \( WC_t \) is identical across agents (I suppress \( i \) superscripts), the wealth distribution does not enter the state space, and the marginal rate of substitution of an arbitrary household is a valid stochastic discount factor.

Plugging (6) into (8) yields

\[
1 = E_t \left[ \delta^\theta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} \left( \frac{WC_{t+1} + 1}{WC_t} \right)^{-1-\theta} \exp(-\gamma \cdot \eta_{t+1}) \right] \equiv E_t[M_{t+1} \tilde{R}_{t+1}],
\]

so that the pricing kernel, which I will denote by \( M_{t+1} \), may be decomposed into the product of the two terms in brackets. The first is the standard pricing kernel from representative agent models with Epstein-Zin preferences, which only depends on aggregate quantities.\(^{37}\) The second term incorporates idiosyncratic consumption risk, which, since it is undiversifiable and uninsurable to the agent, also affects risk premia.

The change in each individual agent’s endowment is the result of a compound lottery. The first stage draws the aggregate state, which determines the higher moments of the distribution from which the idiosyncratic shocks, \( \eta_{t+1} \), are drawn in the second stage. With the sole exception of the consumption claim, \( \eta_{t+1} \) is independent of \( \tilde{R}_{t+1} \). I refer to assets satisfying this restriction as “financial assets”. I price them by projecting out idiosyncratic risk:

\[
1 = E_t \left[ \delta^\theta \exp\left[ -\gamma(\Delta c_{t+1} + \nu_{t+1}) \right] \left( \frac{WC_{t+1} + 1}{WC_t} \right)^{-1-\theta} \tilde{R}_{t+1} \right] \equiv E_t[M_{t+1} \tilde{R}_{t+1}],
\]

\( \nu_{t+1} \) is the log of an expected utility maximizer’s certainty equivalent for the second stage lottery—i.e. the distribution of \( \exp(\eta_{t+1}) \) given \( y_{t+1} \). \( \nu_{t+1} \) converts agents’ (identical) preferences over the higher moments of the second stage lottery into the same units as aggregate consumption.

With recursive preferences, the presence of uninsurable risk can have two, often complementary, effects on risk premia (expected excess returns) relative to the representative agent model. The first is a direct effect, coming from a cross-sectional correlation between the certainty equivalent and returns. Covariance with the certainty equivalent is priced in exactly the same manner as covariance with aggregate consumption. (11) shows that the certainty equivalent is higher in

\[^{37}\text{The equilibrium wealth-consumption ratio will differ from that of a representative agent model with the same aggregate dynamics, since idiosyncratic risk affects the value of the consumption claim.}\]
states in which odd moments of the cross-sectional distribution of $\eta_{t+1}^i$ are higher, and vice versa for even moments. All else constant, assets which perform well when idiosyncratic skewness is unexpectedly low provide a valuable hedging benefit, lowering investors' required rate of return. When the idiosyncratic shock distribution is fat-tailed and highly negatively skewed, as in the data, $\nu_{t+1}$ can be much more volatile than $\Delta c_{t+1}$, amplifying the risk premium.

The second amplification mechanism is an indirect effect which comes from investors' hedging demands (long run risk). This effect is quite transparent in the ICAPM representation in section 3.4 below. When the EIS ($\psi$) is greater than 1 and $\gamma > 1$, investors have a preference for the early resolution of uncertainty and may be willing to pay a premium for assets which offer a hedge against unfavorable news about the higher moments of future idiosyncratic shocks. Such a preference strengthens investors' hedging demands, changing the term in the pricing kernel involving the wealth-consumption ratio.

If agents have CRRA preferences—as is the case in Constantinides and Duffie (1996) and the vast majority of extant literature on asset pricing with incomplete markets—only contemporaneous covariances are priced, and the indirect effect is zero. My estimates in section 5.2 below suggest that the direct effect is reasonably small in the data, which is perhaps unsurprising given that stock returns and labor market measures tend to be leading and lagging business cycle indicators, respectively. However, there is scope for the indirect effect to be quite large. Stock returns are highly informative about future idiosyncratic risk, and my skewness index is quite persistent. Both features combine to generate large hedging demands, generating a large risk premium with moderate risk aversion. Recursive preferences are thus crucial for my quantitative results.

3.3 Solution

Next, I solve a log-linearized version of the model. Since my specification of aggregate dynamics has many properties that have been studied elsewhere, I focus on the incremental effects from incomplete markets. The resulting expressions for asset prices are quite similar, the testable implications for the co-movement of aggregate variables and asset prices can be quite

\[ \text{It plugs in a Taylor series expansion of } \exp(-\gamma \cdot \eta_{t+1}^i) \text{ around zero before taking the (cross-sectional) expectation. The projected kernel will not price assets whose payoffs depend on } \eta_{t+1}^i \text{ properly.} \]

Along these lines, Cochrane (2008) summarizes a common view which dismisses uninsurable idiosyncratic risk as a key driver of the equity premium: “In order to generate risk premia, then, we need the distribution of idiosyncratic risk to vary over time...It needs to widen unexpectedly, to generate a covariance with returns, and so as not to generate a lot of variation in interest rates. And, if we are to avoid high risk aversion, it needs to widen a lot...Slow, business cycle-related variation in idiosyncratic risk $[\nu_{t+1}^{\nu_{t+1}} \text{ in my notation}]$ will give risk to changes in interest rates, not a risk premium.”

The only requisite approximation is the standard Campbell and Shiller (1988) log-linearization, which features in the representative agent solution. Adding idiosyncratic risk does not necessitate additional approximations. For additional details and discussion, see Eraker and Shaliastovich (2008) and Drechsler and Yaron (2011).
different. For brevity, many technical details may be found in Appendix B.4.

I linearize the return on the consumption claim around a constant log wealth-consumption ratio \( \bar{wc} \).\(^{41}\) Then, the log of the one period pricing kernel approximately equals

\[
m_{t+1}^i = \theta \log \delta - (1 - \theta) \kappa_c - \gamma \Delta c_{t+1} - (1 - \theta) (\rho_c wc_{t+1} - wc_t) - \gamma \cdot \eta_{t+1}^i.
\]

As in (9), the representative agent pricing kernel is augmented by an additional term capturing idiosyncratic risk. I then price financial assets by projecting out idiosyncratic risk, yielding

\[
m_{t+1} = \theta \log \delta - (1 - \theta) \kappa_c - \gamma (\Delta c_{t+1} + \nu_{t+1}) - (1 - \theta) (\rho_c wc_{t+1} - wc_t),
\]

an expression involving the certainty equivalent \( \nu_{t+1} \). Lemma 2 in the Appendix shows that, under my distributional assumptions, the certainty equivalent \( \nu_{t+1} \) is an affine function of \( y_{t+1} \).

While (13) will correctly price financial assets, the solution method for assets whose payoffs depend on \( \eta_{t+1} \), namely the consumption claim, is somewhat different.

Proposition 2 gives my key result, namely the model has an affine solution.

**Proposition 2.** Let Assumptions 1-3 hold. The log-linearized model satisfies

(i) \( wc_t = A_0 + A'_t y_t \),

(ii) \( pd_k = A_{0,k} + A'_{0,k} y_t \), for \( k = 1, \ldots, K \).

where \( A_0, A_{0,1}, \ldots, A_{0,K} \) are scalars and \( A, A_1, \ldots, A_K \in \mathbb{R}^K \).

While further details are in Appendix B.4.2, a brief outline of the solution method is as follows. I guess (and later verify) that the log of the wealth-consumption ratio is an affine function of \( y_t \).\(^{42}\) I solve for \( A_0 \) and \( A \) using the Euler equation for the consumption claim and the method of undetermined coefficients. Given my restrictions on the law of motion for the state vector, I can evaluate the Euler equations analytically.\(^{43}\) Since the Euler equations must hold for each \( y_t \),

\(^{41}\) Denoting continuously compounded returns by lowercase letters (e.g. \( \bar{r}_{c,t} = \log R^i_{c,t} \)), \( \bar{r}_{c,t+1} \approx \kappa_c + \Delta c_{t+1} + \rho_c wc_{t+1} - wc_t \), with linearization constants \( \rho_c \equiv \frac{\exp(\bar{wc})}{\exp(\bar{wc})} < 1 \) and \( \kappa_c \equiv \log(\exp(\bar{wc}) + 1) - \frac{\exp(\bar{wc})}{\exp(\bar{wc}) + 1} \). Analogously, the returns on the dividend claims approximately satisfy \( r_{k,t+1} \approx \kappa_k + \Delta d_{k,t+1} + \rho_{kd} p_{kd,t+1} - p_{kd,t} \), where \( p_{kd,t} \) is the log price-dividend ratio for the the \( k \)th dividend stream. The linearization constants are the same, except that long-run values of the price-dividend ratios replace \( \bar{wc} \).

\(^{42}\) Since lagged values of \( \eta^i \) cannot help forecast future values of \( y_{t+1} \) and \( y_t \) is a first order Markov process, the wealth-consumption ratio will only depend on the aggregate state, \( y_t \).

\(^{43}\) An expression for the the conditional moment generating function of \( y_{t+1} \) given \( y_t \) is given in Lemma 3 in Appendix B.4.2. This expectation is an exponential affine function of \( y_t \).
in the state space, I get a system of $L + 1$ nonlinear equations which pin down the coefficients. Analytical solutions are available in special cases, but, in general, the system must be solved numerically. A similar procedure yields solutions for valuation ratios for the other risky assets.

Plugging the affine form into the projected kernel (13) and subtracting terms known at $t$ yields

$$m_{t+1} - E_t(m_{t+1}) = -\left[\gamma S_c' + \gamma \frac{\partial v_{t+1}}{\partial y_{t+1}} + (1 - \theta)\rho_c A'\right][y_{t+1} - E_t(y_t)] \equiv -\Lambda'[y_{t+1} - E_t(y_t)], \quad (14)$$

a multi-factor CAPM-like formula. $\Lambda$ captures the sensitivity of investors’ intertemporal marginal rate of substitution to shocks to the vector of aggregate state variables. The first two terms in $\Lambda$ capture news about contemporaneous (short run) consumption risk. The former captures preferences over the first moment of the cross-sectional distribution of consumption growth, while the latter captures preferences over its higher moments. The third term in $\Lambda$ captures investors’ hedging demands (long run risk), incorporating the indirect effect.

Conditional on the prices of risk ($\Lambda$), and the dividend-price ratio coefficients ($A_{0,k}$ and $A_k$), the representative agent solutions in Drechsler and Yaron (2011) go through with almost no modifications. For example, the vast majority of the excellent discussion in Drechsler and Yaron (2011) describes the model conditional on the valuation ratios and, as such, is directly applicable here. Thus, my discussion is quite brief. However, these ratios—the key objects governing risk premia and the transformation between the physical and risk-neutral measures—differ from those obtained in the absence of idiosyncratic risk.

Proposition 3 gives the solution for the equity premium, which is derived in Appendix B.4.3.

**Proposition 3.** Let Assumptions 1-3 hold. The risk premium for the $k^{th}$ risky asset is

$$\log(E_t[R_{k,t+1}]) - r_{f,t+1} = [S_k + \rho_k A_k]'G_{y,t}G_{y,t}'\Lambda + \lambda_{y,t}'\Omega_k, \quad (15)$$

where $\lambda_{y,t}$ is the vector of jump intensities for $J_{y,t+1}$ and $\Omega_k$ is defined in Appendix B.4.3.

The first term reflects the covariance between the Gaussian innovation to returns and the pricing kernel. The second term is the difference between the expected value of the jump component of returns under the physical measure and its expected value under the risk-neutral measure, which reflects compensation for jump risk. The terms are additive because Gaussian and jump shocks are independent. For additional discussion, see Drechsler and Yaron (2011), section 3.3.3.

If $G_{y,t}$ or $\lambda_{y,t}$ vary over time, then the equity premium can also be time-varying. An immediate corollary is that a variable will only predict excess returns if it is correlated with uncertainty about shocks to the aggregate state. Above, I presented evidence that this condition holds in
the data. When the distribution of idiosyncratic shocks is particularly negatively skewed (i.e. large negative shocks are more likely), uncertainty about future skewness is generally high.

Finally, Appendix B.2 provides affine expressions which can be used to price a risky payment, such as a single dividend $D_{k,t+h}$, as of time $t$. These expressions can be used to derive the term structure of (real or nominal) interest rates as well as the term structure of risk premia—see, e.g. Lettau and Wachter (2007) and Van Binsbergen et al. (2012). Understanding the pricing of risky cash flows at different points in time can help to clarify the mechanics of the model, and, in some cases, generate additional testable predictions. See Appendix B.2 for further details.

3.4 ICAPM Representation

Campbell (1993) proposes an alternative method to derive the pricing kernel which substitutes out consumption growth, therefore relying only on returns data. The result is an intertemporal capital asset pricing model (ICAPM), which can be implemented empirically when the return of aggregate wealth is observable. Even if the return on wealth is unobservable, such a representation highlights the key sources of priced risk in a model. In addition to allowing for incomplete markets, I generalize Campbell et al. (2014) to allow for general affine jump-diffusion dynamics in the aggregate state vector, $y_t$.

Proposition 4, proved in Appendix B.4.4, provides an ICAPM representation of the pricing kernel. Define $R_{c,t+1} = E[R_{c,t+1}^i | y_{t+1}]$ and $r_{c,t+1} = \log R_{c,t+1}$.

**Proposition 4.** Let Assumptions 1-3 hold. Then, the pricing kernel satisfies:

$$m_{t+1} - E_{t} m_{t+1} = -\gamma (\nu_{t+1} - E_t[\nu_{t+1}]) + (1 - \gamma) \left[ N_{FIR,t+1} + N_{DR,t+1} - \gamma N_{CF,t+1} + \frac{1}{2} N_{UNC,t+1} \right],$$

where $\nu_{t+1} \equiv \frac{1}{1 - \gamma} \log E_{t+1}[\exp(1 - \gamma) \eta_{t+1}^i | y_{t+1}]$,

$$N_{FIR,t+1} \equiv \left[ E_{t+1} - E_t \right] \sum_{j=1}^{\infty} \rho_t^j \nu_{t+1+j}^i, \quad N_{DR,t+1} \equiv \left[ E_{t+1} - E_t \right] \sum_{j=1}^{\infty} \rho_t^j r_{c,t+1+j},$$

$$N_{CF,t+1} \equiv \left[ E_{t+1} - E_t \right] \sum_{j=0}^{\infty} \rho_t^j \Delta c_{t+1+j}, \quad N_{UNC,t+1} \equiv \left[ E_{t+1} - E_t \right] \sum_{j=1}^{\infty} \rho_t^j \vartheta_{t+j},$$

and $\vartheta_t$ is defined in Appendix B.4.4.

Relative to the representative agent model, idiosyncratic risk adds two news terms to the pricing
kernel, which are likely to be positively correlated in practice. As discussed above, the first term
captures the “direct effect”, news about contemporaneous idiosyncratic risk. Agents dislike
assets that underperform when the certainty equivalent, $\nu_{t+1}$, is unexpectedly low.

The second term provides compensation for news about the future trajectory of idiosyncratic
risk—the indirect effect. $\nu^*_t$ is a certainty equivalent, but the associated power $(\gamma - 1)$ is lower,
reflecting the fact that $r_{c,t+1}$ is also exposed to the idiosyncratic shock $\eta_{c,t+1}$. Given the high
persistence of my skewness measure, this term is likely to be substantially larger in magnitude
than the contemporaneous term. The additional hedging demands associated with this second
term provide the primary amplification mechanism in my quantitative exercise below.

The first two representative agent terms reflect the differential pricing of news about future
discount and cash flow growth rates, respectively. Within a homoskedastic representative agent
model, only these terms are present. All else constant, investors’ intertemporal hedging motives
make them willing to offer a discount for stocks that positively covary with discount rate news.
The opposite is the case with cash flow news. The decomposition, which is due to Campbell and
Vuolteenaho (2004), also implies that cash flow news carries a risk price which is $\gamma$ times larger
than discount rate news. Intuitively, discount rate shocks are transitory and cash flow shocks
are permanent, making the latter more important to an investor with a long time horizon.

Equation (16) indicates that the price of risk on $N_{FIR,t+1}$ is $\gamma - 1$, one unit smaller than the
coefficient on cash flow news. For standard choices of $\gamma$, this implies that the cross-sectional price
of risk for news about future higher moments of consumption growth is much closer in absolute
value to the price of risk for cash flow news (that is, news about the mean of consumption
growth) than discount rate news. Moreover, if the cross-sectional certainty equivalent $\nu^*_t$ is
more persistent and/or volatile than aggregate consumption growth, this term can play a very
important quantitative role in amplifying risk premia.

Finally, in the presence of stochastic volatility and/or jumps, there is a final representative agent
term which captures news about state variables governing the higher moments of aggregate
shocks. The Jensen’s inequality term $\vartheta_t$ is high when uncertainty is high. All else constant,
risk averse agents are willing to pay a premium for assets which hedge against increases in
uncertainty. Thus, the price of risk on $N_{UNC,t+1}$ is negative. For additional discussion, I refer
the interested reader to Campbell et al. (2014).

Empirical implementations of the ICAPM include $r_{c,t+1}$, which is assumed to be observable,
as an element of the state vector $y_{t+1}$. Under my assumptions, $\vartheta_t$ and $\nu^*_t$ are affine function
functions of $y_t$ and $y_{t+1}$, respectively. Each news terms above is a linear combinations of VAR

\footnote{Popular choices for $\gamma$ in the theoretical literature with Epstein-Zin preferences often range between 5 and 15.}
residuals. Thus, the ICAPM provides an alternative approach for deriving the prices of risk $\Lambda$, which can potentially be more robust to misspecification of the dynamics of $\Delta c_{t+1}$.

4 Quantitative Model

Section 3 integrates my incomplete markets mechanism into a general, jump diffusion model for aggregate cash flows. Here, I work with a standard specification for aggregate risk to highlight the amplification in risk premia from incomplete markets. The novel mechanism is that agents are exposed to rare, idiosyncratic disasters, and the idiosyncratic disaster probability is time-varying. Despite its simplicity, the stylized model is matches key asset pricing moments well, without relying on low-frequency variation in state variable dynamics.

4.1 Setup

I perturb the representative agent model so that agents are exposed to idiosyncratic, uninsurable event risk. For parsimony, throw away all state-independent sources of risk. Further, while my empirical results provide evidence of state dependence in both tails of the idiosyncratic risk distribution, the stylized model only emphasizes downside risk to keep the model as transparent as possible. As such, the novel mechanism in the model is the inclusion of a time-varying probability of uninsurable idiosyncratic disasters within an otherwise standard endowment economy.

My model for aggregate dynamics adopts the general structure of the Bansal and Yaron (2004) long-run risk model. Aggregate cash flows evolve according to

\[
\begin{align*}
\Delta c_{t+1} &= \mu_c + \phi_c x_t + \sigma_c \epsilon_{c,t+1} \\
\Delta d_{t+1} &= \mu_d + \phi_d x_t + \sigma_d \epsilon_{d,t+1},
\end{align*}
\]

where $d_{t+1}$ is the log dividend on the market portfolio. Asset pricing dynamics are driven by two persistent state variables. The first is $x_t$, a small but persistent component governing expected cash flow growth. The second, $\sigma_t^2$, captures the conditional variance of shocks in the economy.

The persistent component, $x_t$, plays a second role in my model. $x_t$ also controls the higher moments of idiosyncratic shocks to consumption. Agents are exposed to a single jump component, $J_{\eta,t+1}$, and its Poisson intensity $\lambda_{\eta,t+1}$—the personal disaster probability—is $\lambda_0 - \lambda_1 x_t$. $\lambda_1$ is

---

45Analytical expressions for these linear combinations are available upon request.

46State independent risk primarily affects the risk-free rate and thus has little effect on excess returns. Adding such a component is similar to changing the discount rate $\delta$. 

28
positive, so that personal disasters become more likely in states when expected cash flow growth is low. As above, I normalize $x_t$ to have mean zero and variance one, so that $\lambda_0$ and $\lambda_1$ can be interpreted as the unconditional disaster probability, and $\lambda_1$ is the sensitivity of the conditional disaster probability to a 1 standard deviation change in $x_t$.

The personal disaster magnitude (the jump size) is normally distributed with mean $\mu_b$ and standard deviation $\sigma_b$. This choice is inconsistent with the calibration in Section 2.3, which assumed exponentially-distributed jumps. In the interest of conservatism, I choose the thin-tailed normal shocks. In my calibration, $\mu_b$ is a large, negative number, and $\sigma_b^2$ is very large relative to uncertainty about aggregate consumption. Analogously with rare macroeconomic disasters, these infrequent labor market events are associated with extremely high marginal utilities, so they have a large impact on asset prices despite their relative infrequency.

The state variables evolve according to

$$
x_{t+1} = \rho x_t + \sigma_t \epsilon_{x,t+1} \\
\sigma_{t+1}^2 = 1 - \rho + \rho \sigma_t^2 + \sigma_t \epsilon_{\sigma,t+1}
$$

(19) (20)

where I normalize $E[x_t] = 0$ and $E[\sigma_t^2] = 1$. In Bansal and Yaron (2004), the shock to $\sigma_t^2$ is homoskedastic. I allow $\sigma_t^2$ to follow a square-root process to be consistent with evidence from Table 2. This restriction also guarantees that $\sigma_t^2$ is nonnegative in the continuous time limit.

Relative to Bansal and Yaron (2004), I allow for a somewhat richer correlation structure among the residuals in the model. The covariance matrix for the shocks is

$$
E_t[\epsilon_{t+1} | \epsilon_{t+1}'] = \begin{bmatrix}
\varphi_c^2 & \pi_c \varphi_c^2 & 0 & 0 \\
\pi_c \varphi_c^2 & \varphi_c^2 + \varphi_d^2 + \pi_x^2 \varphi_x^2 + \pi_\sigma^2 \varphi_\sigma^2 & \pi_x \varphi_x^2 & \pi_\sigma \varphi_\sigma^2 \\
0 & \pi_x \varphi_x^2 & \varphi_x^2 & \chi \varphi_x \varphi_\sigma \\
0 & \pi_\sigma \varphi_\sigma^2 & \chi \varphi_x \varphi_\sigma & \varphi_\sigma^2
\end{bmatrix}
$$

(21)

The Bansal and Yaron (2004) specification assumes that $\pi_c$, $\pi_x$, $\pi_\sigma$, and $\chi$ are zero. Bansal et al. (2012) allow $\pi_c \neq 0$, which permits for a nonzero covariance between consumption and dividend innovations. Analogously, $\pi_x$ and $\pi_\sigma$ allow for a nonzero correlation between news about dividend growth and the state vector. $\chi$ permits a nonzero correlation between $x_t$ and $\sigma_t^2$ innovations, which has considerable support in the data.
4.2 Amplification

Under these assumptions, $A_x$, the sensitivity of the wealth consumption ratio—the key determinant of hedging demands in the pricing kernel—to $x_t$ innovations is

$$A_x = \left[ \frac{1 - \frac{\psi}{\phi}}{1 - \rho_x \rho_c} \right] \left[ \phi_c + \rho_x \frac{\partial \nu^*_t}{\partial x_t} \right],$$

(22)

where $\nu^*_{t+1} \equiv \frac{1}{1-\gamma} \log E_t[\exp((1-\gamma)\eta^t_i|x_{t+1})]$ and $\rho_c$ is the log-linearization constant. Assuming that $\gamma > \psi > 1$, as is standard in the long-run risk literature, the first term in brackets is positive. The second term captures the sensitivity of the household’s flow utility to changes in $x_t$. The first piece, $\phi_c$, comes from the predictability of aggregate consumption growth, while the second piece involving the cross-sectional certainty equivalent, $\nu^*_t$, comes from predictability of the higher moments of idiosyncratic consumption growth shocks. Assuming that $\lambda_1 > 0$ and $\rho_x > 0$, the contribution from this term is also positive. Disasters states are associated with extremely high marginal utility, so this second term dominates in my model.

From inspection of (22), the amplification mechanism associated with idiosyncratic risk becomes quite clear. Marginal utility becomes more sensitive to $x_t$ as the cross-sectional certainty equivalent becomes more sensitive to $x_t$. Households face more risk, so their hedging demands against future increases in that risk are larger. In addition, as in the representative agent case, the sensitivity also increase as $x_t$ becomes more persistent and/or aggregate consumption becomes more predictable. Though the expression for the coefficient on $\sigma^2_t$ is messier, a similar amplification is present. Under my preference configuration, the price of volatility risk is negative. Idiosyncratic risk makes it larger in magnitude relative to the representative agent case.

Turning to the price dividend ratios, one can also show

$$A_{x,m} = \frac{1}{1 - \rho_x \rho_m} \left( \phi_x - \phi_c - \rho_x \left[ \gamma \left( \frac{\partial \nu_t}{\partial x_t} - \frac{\partial \nu^*_t}{\partial x_t} \right) + \frac{1}{\psi} \frac{\partial \nu^*_t}{\partial x_t} \right] \right).$$

(23)

The contribution from the first two terms in parentheses come from the representative agent solution, whereas the last term (in brackets) comes from incomplete markets. The first term within the brackets in (23), which is generally positive, compares the sensitivity of the certainty equivalents of two invididuals, where one is more risk averse than the other, to changes in the higher moments of $\exp(\eta^t_i)$ given $y_{t+1}$. Analogously, $\nu^*_{t+1} \equiv \frac{1}{1-\gamma} \log E_t[\exp((1-\gamma)\eta^t_i|x_{t+1})]$ is a certainty equivalent for the same lottery when the risk aversion parameter is $\gamma - 1$. When $\gamma \geq 2$, the two partial derivatives will have the same sign, and $\frac{\partial \nu^*_t}{\partial x_{t+1}} \geq \frac{\partial \nu^*_t}{\partial x_{t+1}}$.
markets analogue to the $-\frac{\phi}{\psi}$ term coming from the representative agent solution.

Inspection of the bracketed term in (23) reveals one of the potentially counterintuitive implications of the incomplete markets model. Idiosyncratic risk generally increases the sensitivity of the pricing kernel to shocks to the aggregate state vector. However, idiosyncratic risk pushes in the opposite direction for the price dividend ratios. All else constant, for a given degree of dividend predictability, returns will tend to fall less in response to bad news relative to the representative agent case.

What is the intuition behind this term? In the representative agent model, when aggregate consumption becomes more risky, standard calibrations assume that the dividend claim becomes even riskier. This gives agents an incentive to shift away from dividends, causing the price-dividend ratio to fall. In my model, given the interactions between aggregate and idiosyncratic risk, idiosyncratic risk also increases when dividends becomes more risky. All else constant, this additional channel makes the dividend claim—which is not exposed to idiosyncratic shocks—appear more favorable than it otherwise would, increasing the precautionary savings demand for financial assets. The last term captures the strength of this precautionary savings motive. So, the price-dividend ratio will be less responsive to changes in the state variables than it would be in the representative agent model.

Taking stock, within the context of this stylized model, idiosyncratic risk increases the sensitivity of household marginal utility (the pricing kernel) to news about the state variables, but it reduces the sensitivity of returns. My quantitative exercise suggests that the increased sensitivity of the pricing kernel is more important. However, one can find parameter configurations where consumption is substantially riskier than dividends, causing stock prices to increase in response to bad news, which is counterfactual (see, e.g. the impulse responses in Figure 6). In order to generate a substantial risk premium, dividends must be fairly predictable.\textsuperscript{48}

\section*{4.3 Calibration}

Table 3 provides an overview of the parameters in the quantitative model, along with the calibrated values. While the number of parameters is larger than Bansal and Yaron (2004), many key parameters are tied directly to the data. The model is quarterly, to match the frequency of the analysis in section 2. See Appendix B.3 for further details.

\textsuperscript{48}Constantinides and Ghosh (2014) calibrate a model with a relatively similar structure, except that idiosyncratic risk is driven by a single variable which follows a square root process. In their calibration, aggregate consumption and dividends are i.i.d. When idiosyncratic risk is sufficiently persistent, they generate a large equity premium even in the absence of predictability. This occurs because the level and volatility of idiosyncratic risk are perfectly correlated, so agents' preference for an early resolution of uncertainty causes prices to fall in response to bad news about future idiosyncratic risk.
I begin with the parameters governing idiosyncratic shocks. The disaster probability parameters are chosen to correspond with the state-dependent component of large negative shocks from the calibrated model in section 2.3. Turning to the distribution of jump sizes (idiosyncratic disaster magnitudes), I choose parameters with an eye towards conservatism. I do not assume that income shocks translate one-for-one into consumption shocks, since stockholders have some means with which to smooth their consumption over time, and households often have more than one earner. I combine my calibrated parameters from section 2.3 with an assumption about the elasticity of consumption growth with respect to permanent income growth. I set $\mu_b = -18\%$ and $\sigma_b = 11.5\%$, values which translate to an elasticity of about 23%.\(^{49}\)

I choose the level of consumption predictability $\phi_c$ so that an agent’s consumption is i.i.d. conditional on not receiving a jump shock. This choice implies that, unlike the Bansal and Yaron (2004) calibration, aggregate consumption growth is essentially unpredictable. $\phi_c$ exactly offsets the location adjustment ($F_\eta$) which is subtracted off to ensure proper aggregation. Given this restriction, the only source of predictability in $\Delta c_{t+1}$ is the conditional expectation of the jump shock—a restriction which approximately holds in the income data (see Figure 3).\(^{50}\) If, instead, I were to set $\phi_c = 0$, $\Delta c_{t+1}$ is a random walk. However, the location adjustment would counterintuitively imply that, for all individuals who do not receive jump shocks, the distribution of consumption growth shifts to the right as the personal disaster probability increases.

I choose the persistence parameters, $\rho_x$ and $\rho_{\sigma}$, to match the first order dynamics of my skewness index and initial claims, respectively. I estimate the AR(1) parameters using a regression which is adjusted for finite-sample bias as in Bauer et al. (2012).\(^{51}\) Analogously, the correlation between the AR(1) innovations, $\chi$, is estimated directly from the bias-corrected regression residuals. I choose the volatility of $\sigma_t^2$ to match the Kelley’s skewness of initial claims in the data.

Table 4 compares my assumptions about state variable dynamics and cash flow predictability with two popular calibrations of the long run risk model, Bansal and Yaron (2004, BY) and Bansal, Kiku, and Yaron (2012, BKY), both of which generate large risk premia in line with the data.\(^{52}\) The top panel compares the persistence coefficients, which are expressed as monthly autocorrelations. In my model, the half life of an $x_t$ shock is 1.4 years, which is considerably shorter than the half lives of 2.7 and 2.3 years in the BY and BKY calibrations, respectively. $\sigma_t^2$

\(^{49}\) Blundell et al. (2008), Table 7 estimates a 22.5\% elasticity for male earnings. All other estimates are higher.

\(^{50}\) Wachter (2013) makes a similar assumption in a model with rare microeconomic disasters.

\(^{51}\) The OLS coefficient in an AR(1) model suffers from a downward finite sample bias, which can be nontrivial when the dependent variable is fairly persistent. Bauer et al. (2012) develop an algorithm which corrects for this bias and show that it improves the ability of an estimated affine term structure model to fit the data.

\(^{52}\) My key objective in making this comparison is to illustrate the amplification associated with incomplete markets. While the first order autocorrelations of my skewness index and initial claims are indeed lower than standard choices of $\rho_x$ and $\rho_{\sigma}$ in the long run risk literature, this need not imply that idiosyncratic risk and labor market uncertainty feature important sources of low frequency variation.
is also less persistent. $\sigma_t^2$ shocks have a half life of 2.9 years versus 4.4 years in the BY model. The BKY model emphasizes extremely low frequency movements in volatility, so their choice of $\rho_o$ implies that a $\sigma_t^2$ shock has a half life of 57.7 years.

The next panel reports the volatility of expected consumption and dividend growth, expressed as an annualized percentage. Given my focus on idiosyncratic risk, I deliberately shut off almost all consumption predictability, so the volatility of annualized expected consumption growth is 15 basis points. Predictable variation in aggregate consumption is the primary source of the equity premium in BY and BKY, so the volatility of expected consumption growth is considerably higher (2% and 1.5%, respectively).

Given the lower persistence of $x_t$ in my model, dividends are more predictable at short horizons but less predictable at longer horizons. The volatility of the conditional mean of dividend growth is 10% when expressed as an annualized rate, as compared with 6.1% and 3.7% in the BY and BKY calibrations. As such, I report a measure of the overall level of dividend predictability, which is the change in the expected discounted sum of future dividend growth associated with a 1 standard deviation increase in $x_t$. I calculate these sums with 0%, 5%, and 10% annual discount rates. Regardless of the discount rate, overall dividend predictability in my model is roughly comparable with BY and somewhat higher relative to BKY.

### 4.4 Performance

Table 5 demonstrates the ability of the quantitative model to match a number of key asset pricing moments. Data moments are taken from BKY, who calculate statistics using annual time series of real returns and cash flow growth rates from 1930-2008. I refer the reader to their paper for further details about the underlying data sources. Next, I use the model to simulate 50,000 annual time series of the same length, then report a number of quantiles of the finite sample distribution of the calibrated model. These quantiles can also be interpreted as robust standard errors for the model-implied moments.

Most importantly, my model generates a large and time-varying equity premium of about 6.5% per year. It easily replicates the excess volatility puzzle; the volatility of the market return is 10% larger than that of dividend growth. The addition of incomplete markets leads to a fairly volatile real interest rate, whereas long run risk models with complete markets tend to exhibit too little volatility. The model also matches the level of the price-dividend ratio almost exactly, though it understates its volatility (a shortcoming of the BY and BKY models as well). The price dividend ratio also exhibits a lower degree of autocorrelation relative to the data, which is unsurprising given that the state variables in my model are not very persistent.
Looking at the cash flow moments, the biggest differences between the model and the data are the first order autocorrelations of consumption and dividend growth, which are significantly lower and higher than the corresponding values in the data, respectively. The former is by construction, given that I deliberately shut off almost all consumption predictability to highlight the amplification coming from incomplete markets.

The autocorrelation of dividend growth deserves more discussion. In my model (and the BY and BKY models), the leading term in the equity premium is \((1 - \theta)\sigma_t^2 \phi^2 A_x A_{x,m}\), which is the covariance between returns and the hedging demand for \(x_t\) shocks. The addition of idiosyncratic disaster risk makes \(A_x\) large and positive. However, dividends need to be riskier than consumption in order for \(A_{x,m}\) to be positive. This ensures that, consistent with the data, valuation ratios are procyclical. I achieve this by assuming that dividends are fairly predictable over short to medium-term horizons, increasing the autocorrelation of model-implied dividend growth. As discussed above, the overall level of dividend growth predictability in my model is comparable with Bansal and Yaron (2004).

Table 6 highlights the incremental contribution from incomplete markets by comparing the asset pricing moments from my model with those obtained from a comparable representative agent model. The Markets column indicates whether the relevant moment is obtained from the incomplete markets or representative agent version of the model. It reports the average of the moments in Table 5 of a long simulation of 1 million quarters. I shut off idiosyncratic risk by setting \(\lambda_0\) and \(\lambda_1\) equal to zero. In addition, I raise the rate of time preference considerably to 0.996, which generates a risk-free rate of about 2\%. In addition, I demonstrate the effects of shutting down several dimensions of risk which are embedded in my baseline calibration.

In the representative agent version of my baseline specification, the equity premium is generated by three distinct channels. The first is a contemporaneous covariance between consumption and dividend growth innovations. Second, since \(\phi_c \neq 0\), there is a small amount of predictability in aggregate consumption growth. Third, there is stochastic volatility about aggregate consumption innovations, as well as shocks to the state vector. The Baseline column of Table 6 indicates that these channels combine to generate an equity premium of 3.2% per annum. Thus, the addition of incomplete markets doubles the risk premium in the baseline specification.

Looking at several of the other asset pricing moments, some general patterns emerge. The risk-free rates are considerably more volatile in the incomplete markets model, and price-dividend ratios are somewhat more autocorrelated. Returns are even more volatile in the representative agent versions of the model. This follows from \((23)\), the expression for \(A_{x,m}\). Holding dividend predictability constant, the additional precautionary savings motive associated with incomplete markets.

\(^{53}\)In some specifications, price-dividend ratios approach infinity if I try to match the observed risk-free rate.
markets reduces the sensitivity of the price-dividend ratio to changes in $x_t$.

The baseline model allows for a contemporaneous correlation between dividend growth innovations and shocks to $x_t$ and $\sigma^2_t$ via the parameters $\pi_x$ and $\pi_\sigma$. The next column, titled “No Covariance”, sets both parameters to zero. Eliminating these covariances reduces the incomplete markets equity premium by 1%, whereas it leads to a substantially smaller reduction in the representative agent equity premium. This restriction reduces the volatility of returns and cash flow growth and generates a mild reduction in the autocorrelation of dividend growth.

The next column reduces the importance of stochastic volatility relative to the baseline model. While shocks to $x_t$ and $\sigma^2_t$ continue to be heteroskedastic, the terms involving $\varphi_c$ and $\varphi_d$ in the covariance matrix (21) are now assumed to be homoskedastic. This restriction affects both risk premia symmetrically, reducing both by about 0.7%.

The final column of Table 6, “IID Cons”, keeps the reductions on the role of stochastic volatility from the “Less CF Vol” column and, in addition, sets $\varphi_c = 0$. These restrictions imply that aggregate consumption is i.i.d. Again, the associated reduction in the risk premium of about 0.5% is the same across models. With i.i.d. aggregate consumption, the incomplete markets model generates a risk premium of 5.3% versus a 2% risk premium with complete markets.

5 Interactions between Idiosyncratic Risk and Stock Returns

In this section, I test several necessary conditions implied by my general model. I discuss the implications of my incomplete markets model for return predictability, and I consider the dynamic interactions between proxies for idiosyncratic risk and asset returns.

5.1 Labor market uncertainty predicts returns

A primary objective is to study the ability of an asset pricing model with incomplete markets to generate large, time-varying equity premia. In this section, I test a necessary condition for such a model by considering the ability of my preferred labor market uncertainty measure, initial claims, to predict returns in the data. I also consider its covariance with and compare its forecasting power with other leading predictor variables from the extant literature. I find that initial claims for unemployment outperforms essentially all of the univariate predictors at short horizons (3 months to 1 year) and the vast majority of variables at a 2 year horizon. Moreover, I find that common components of the associated univariate forecasts track labor market conditions. Many variables which are motivated as proxies for aggregate consumption
risk also contain important information about idiosyncratic risk.

In this analysis, I emphasize initial claims relative to the other two measures discussed above. I do so for a number of reasons. First, initial claims is available at a higher frequency, does not require the estimation of any parameters, and is less likely to be prone to measurement errors. Second, as discussed above, Granger-causality tests suggest that initial claims leads the skewness index, suggesting that initial claims is likely to outperform in a predictive setting. Finally, in contrast to the other two measures, initial claims has substantial explanatory power for uncertainty about future skewness as well as its own innovations.

Since a wealth of potential predictors have been suggested in the literature, I focus on a subset of 12 monthly variables considered in Goyal and Welch (2008), which are compiled and updated regularly by Ivo Welch. As the vast majority of these variables are quite standard in the literature, I refer the reader to Goyal and Welch (2008) for detailed descriptions of variable construction, as well as references to the original studies which proposed each variable.

In addition to the univariate predictors, I summarize the predictive content of all 12 variables by taking equal-weighted combinations of the fitted values from a univariate regression of 1 year-ahead excess returns on each predictor. I emphasize these combination forecasts in lieu of estimating multivariate models because the finite sample properties of these forecasts are much more desirable, and, as emphasized by Goyal and Welch (2008), estimation error is a first-order concern within this context. Indeed, these combinations generally outperform all but the best univariate models in-sample, and Rapach et al. (2010) demonstrate that combinations perform much better out-of-sample.

I produce three combination forecasts. The first is an equal weighted combination of the univariate forecasts from each of the variables over the entire sample period: 1928-2012. The second begins the estimation in 1967, the first period for which initial claims data are available. Finally, I orthogonalize each of the predictors with respect to initial claims, then form combinations of the fitted values from univariate regressions of returns on these orthogonalized predictors.

Table 7 presents a number of pairwise correlations between initial claims, each of the predictors, and a measure of employment declines over the next three months—a simple measure of labor market conditions. Initial claims has a 38% correlation with future employment declines and both of the combination forecasts (67% and 58%, respectively). It is even more strongly correlated with the dividend yield (74%), the book-to-market ratio (76%), and the default yield (69%). It is also positively correlated with the T-bill rate (41%), the long term yield (41%) on government bonds, and the inflation rate (24%), which is primarily driven by the period in the 1970s where both inflation and labor market uncertainty were elevated. More surprising
is the negative correlation with the earnings-price ratio (-46%), which appears to be driven by differences in low frequency variation between the two measures.\footnote{An even stronger negative correlation (-72\%) arises between the dividend yield and the earnings price ratio.}

Next, I report the pairwise correlations between each of the equity premium combination forecasts and our predictor variables. The first of the combination forecasts is most strongly correlated with the dividend yield (72%), the book-to-market ratio (67%), and the default yield (62%). All three measures are highly correlated with initial claims, suggesting that they are all capturing a common macroeconomic risk factor. Figure 5 overlays initial claims with the dividend yield, as well as the first of the combination forecasts. These measures are highly correlated with one another; spikes or troughs in initial claims are generally accompanied by similar movements in one or both of the other risk premium measures.

Turning to the second combination model for the 1967-2012 sample, the combination forecast is most strongly correlated with the default yield (71%) and the term spread (63%). While the individual pairwise correlations change a lot, the two combination forecasts are fairly highly correlated with one another (77%), consistent with time variation in the implied risk premium from the combinations being somewhat more robust to estimation error relative to univariate models. The dividend yield and book-to-market ratio track this combination forecast less closely. During this period, the pairwise correlation between initial claims and the combination forecast is still higher than any of the remaining univariate predictors.

Finally, when I form a combination forecast using the orthogonalized predictors, the resulting series loads most heavily on the term spread, inflation, and the yield curve. Orthogonal components of the dividend yield, the book-to-market ratio, and the default yield, variables which are most highly correlated with initial claims, are much less strongly correlated with these combination forecasts. Note that this combination forecast, despite being uncorrelated with initial claims, captures information about the conditional mean of employment growth. The combination forecast which is constructed using the orthogonalized predictors has a 41\% correlation with future cuts in employment, which is actually higher than the pairwise correlation between employment cuts and initial claims (38\%).

Table 8 summarizes the forecasting performance of each of the predictor variables for cumulative returns. I report the $R^2$ and the $t$-statistic on $\beta_h$ from the following predictive regression:

$$\sum_{j=1}^{h} r_{t+j} \equiv r_{t:t+h} = \alpha_h + \beta_h x_t + u_{t:t+h},$$

where $r_t$ is the log return on a given portfolio, $x_t$ is the predictor variable, and $h$ is the forecast.
horizon. Rows correspond with different predictors, while columns correspond with different portfolios and forecast horizons. I consider forecasts of the log excess return on the CRSP value-weighted index, as well as the Fama and French (1993) SMB portfolio.\textsuperscript{55} I consider forecast horizons ($h$) of 3, 12, and 24 months, though results are similar at other horizons. My sample period is 1967-2012. In order to make an apples-to-apples comparison, when looking at the other predictors, I limit my attention to the period for which initial claims data are available.

The results in Table 8 suggest that initial claims for unemployment is a powerful, highly robust predictor of broad market returns (left columns). At a three month horizon, initial claims achieves an $R^2$ of 2.4%. Initial claims outperforms every one of the Goyal and Welch (2008) predictors, and its performance is comparable with the first and third combination forecasts. The only other statistically significant univariate predictor is the term spread, which achieves an $R^2$ of 1.76%. At a 1 year horizon, the $R^2$ is 6.4%, which is statistically significant. Only the term spread performs better with an $R^2$ of 8%. At a 2 year horizon, claims performs a bit worse, though the magnitude of the $R^2$ is still reasonably high. The combination forecasts perform extremely well at the 1-2 year horizons.

A couple of other points are worth noting about the left panel of Table 8. First, the 1967-2012 sample period is a tough one for the Goyal and Welch (2008) variables. Many of the most frequently emphasized predictors, including the dividend yield, book-to-market ratio, and the default yield fail to achieve statistical significance. Stock market realized volatility is statistically significant at longer horizons, though the associated magnitudes are quite small. Inflation achieves significance, though its sign is (arguably) wrong. Second, the second combination forecast outperforms all other models by a wide margin at all horizons. This is not surprising, given that I am taking an average of fitted values from 12 univariate regressions, all of whose coefficients are estimated using data from the period over which evaluation takes place.

Turning to the right panels, I find that initial claims is an even stronger predictor of the excess return on the Fama and French (1992) SMB portfolio. The $R^2$ values are 4.3%, 10.8%, and 9.5% at 1 quarter, 1 year, and 2 year horizons, respectively. This performance is better than any of the Goyal and Welch (2008) predictors or any of the combination forecasts at all horizons. The term spread, which performed the best at predicting the market return, has essentially no predictive content for the SMB portfolio. Further, initial claims is the only predictor which is statistically significant at the 95% level at the 2 year horizon. These results suggest that small stocks may be disproportionately exposed to deterioration in labor market conditions, causing their risk premia to increase more when labor market uncertainty is high relative to larger stocks.

\textsuperscript{55}Results are qualitatively similar for the HML portfolio, though the statistical evidence is much weaker. None of the variables (including the combinations) forecast HML well at short horizons, though I find weak evidence that initial claims forecasts HML at long horizons.
In additional unreported tests, I include initial claims in bivariate regressions along with each of the Goyal and Welch (2008) predictors. In virtually all of the specifications, initial claims remains a positive, statistically significant predictor. These results are available upon request.

5.2 Market returns are informative about future labor market conditions

When investors have Epstein-Zin preferences, an asset’s risk premium depends on the covariance between its return and news about both contemporaneous and future idiosyncratic risk. In addition, agents are willing to pay a premium to hedge against labor market uncertainty shocks. In this section, I explore the covariance structure between market return innovations and my proxies for the level of and uncertainty about idiosyncratic risk. Empirically, I find that while return innovations have little predictive content for contemporaneous measures, they are highly informative about future labor market conditions.

To demonstrate this relationship in as parsimonious of a way as possible, I estimate model-free impulse response functions. My method closely relates to the local projection method of Jordà (2005). Jorda’s method uses direct forecasts to estimate impulse responses at longer horizons, as opposed to iterating on a (potentially misspecified) one-period model for the evolution of the state vector. However, I identify the shocks via different means, using an argument from Lamont (2001) which is frequently used to construct portfolios whose returns are informative about innovations in economic state variables: factor-mimicking or economic tracking portfolios.

For a given observable variable $y_t$, my definition of the impulse response is

$$ E[y_{t+k}r_{t+1} - E[r_{t+1}] = v, \mathcal{F}_t] - E[y_{t+k}r_{t+1} - E[r_{t+1}] = 0, \mathcal{F}_t] $$

for different values of $k$. Given a set of conditioning variables, $z_t$, I decompose $y_{t+k}$ as

$$ y_{t+k} = \text{proj}(y_{t+k}|z_t) + \left[ E_t(y_{t+k}) - \text{proj}(y_{t+k}|z_t) \right] + \epsilon_{t:t+k} = \beta'_k z_t + \xi_k + \epsilon_{t:t+k}, \quad (25) $$

where $\xi_{t:t+k}$ is a term reflecting potential misspecification of the conditional mean of $y_{t+k}$ and $\epsilon_{t:t+k}$ is the “true” innovation. I additionally assume that, as is the case in my general theoretical framework, the conditional mean of returns takes the linear form $r_{t+1} = \gamma' z_t + v_{t+1}$, where $v_{t+1}$ has mean zero. Then, the impulse response equals $E[\epsilon_{t:t+k}|v_{t+1} = v]$. One obtains consistent estimates of $\xi_k + \epsilon_{t:t+k}$ and $v_{t+1}$ by taking the residuals from regressions of $y_{t+k}$ and $r_{t+1}$ on $z_t$, respectively. Given these residuals, I estimate $\text{proj}(\epsilon_{t:t+k}|v_{t+1} = 1) \equiv \alpha_k$ by regressing $\hat{\xi}_k + \hat{\epsilon}_{t:t+k}$ on $\hat{v}_{t+1}$. Inference is straightforward, since the estimate of $\alpha_k$ from this two step procedure is identical to the coefficient on $r_{t+1}$ from a regression of $y_{t+k}$ on $z_t$ and $r_{t+1}$.

This approach works because $v_{t+1}$ has mean zero and is independent of $\xi_k$, so misspecification of the conditional mean adds noise to the dependent variable $(\hat{\xi}_k + \hat{\epsilon}_{t:t+k})$ of the second stage
regression. As long as I have estimated the return innovation correctly, I need not have specified the mean of \( y_{t+k} \) correctly. The advantage of such an approach is that, in contrast with macroeconomic time series, returns are almost serially uncorrelated. While the conditional mean of returns does vary over time, this variation is second order compared with its highly volatile unforecastable component. However, the use of a direct estimation method places practical constraints on the maximum lag length which can be considered.

Figure 6 shows the estimated impulse response functions to market excess returns for six different macroeconomic variables over twelve quarters. The vector \( z_t \) includes 4 lags of the target variable, the dividend yield, initial claims for unemployment, the term spread, and the 3 month T-bill rate. For purposes of identification, it is more important that the variables \( z_t \) capture the conditional mean of returns, as opposed to the target variables. Including lags of the target helps to reduce noise in the estimation of the news terms, though, consistent with my identification argument, the results are insensitive to the inclusion of one or more lagged terms.

The top left panel shows the responses of real aggregate consumption growth (real consumption of nondurables and services from the National Income and Product Accounts). The estimated response is positive and significant for the first few quarters, though it quickly trails off to zero at longer horizons. Note however that the associated magnitudes are quite small. A 1 standard deviation (+8.5%) quarterly return innovation is associated with a cumulative consumption response of only about 30 basis points, which is 15% of the standard deviation of annual consumption growth. Note that the absence of a response after the first year is inconsistent with the presence of a highly persistent component in expected consumption growth. However, my regression-based test could have low power to detect news about an extremely persistent component if its variance is sufficiently small.

Next, I consider the response of my conditional skewness index. Since \( x_t \) is normalized to have an unconditional variance equal to one, the response is measured in standard deviation units. The response is small and insignificant on impact, peaking at about 1/3 of one standard deviation 3 to 4 quarters after the return innovation is observed. The response turns statistically insignificant around 7-8 quarters later. My point estimates are slightly negative, though insignificant, in the last four quarters. Such a result is consistent with a transitory component in idiosyncratic skewness which subsequently reverses itself.\(^{56}\) The magnitude of the skewness response is quite substantial; the cumulative response over the first two years is 1.35 standard deviations.

Finally, I plot the response of my labor market uncertainty proxy, initial claims for unemployment, to return innovations. The response is hump-shaped and unambiguously negative.

\(^{56}\)If I reestimate the regressions with only lags of the targets in \( z_t \) (which would be valid if returns were unpredictable), these negative point estimates disappear.
A positive return innovation is associated with a decrease in future labor market uncertainty, where the news is most informative about initial claims 6-18 months in the future. Here, the magnitudes are fairly substantial, given that the high persistence of initial claims.

6 Conclusion

This paper presents evidence for the quantitative importance of idiosyncratic tail events as an important driver of variation in risk premia over time. The vast majority of theoretical research on time varying risk premia exclusively emphasizes risks associated with the level of aggregate consumption over time. My analysis suggests that risks associated with redistribution of consumption across across agents can be just as important, if not more important, than aggregate consumption risks. I view this paper’s contribution as a “proof of concept”; there remains plenty of room for additional work.

Labor market event risk is likely to provide a novel mechanism for the amplification of aggregate shocks. If the uninsurability of labor market shocks causes discount rates to rise much more sharply in response to bad news than they would if markets were complete, firms’ incentives to invest are likely to be substantially distorted. For example, a recent literature emphasizes the link between uncertainty and economic growth. In representative agent models, uncertainty affects risk premia indirectly (e.g., by changing the distribution of aggregate consumption). In my model, as well as Herskovic et al. (2014), uncertainty has an additional, direct effect on preferences when it is liked with the distribution of idiosyncratic shocks. My model can be easily embedded in a production setting, and I plan to explore these interactions in future work.

In the data, aggregate and idiosyncratic risks are tightly linked with one another. While my general model easily accommodates the study of these interactions, I deliberately downplay risks associated with aggregate consumption so as to highlight the potential of the incomplete markets mechanism. My model simply takes labor market event risk and its relationship with aggregate shocks as an exogenous input. A richer model would endogenize these interactions, enabling it to address a larger number of policy questions.

My estimates of the distribution of idiosyncratic shocks are intended to provide an order of magnitude for the degree of tail risk agents face via the labor market. Given recent improvements in the quality of panels of earnings records, one should be able to pin down these distributions fairly precisely. Its tails are effectively observable given the cross-sectional sample sizes available. This feature make the key parameters of the incomplete markets model much easier to estimate relative to those governing aggregate tail risk.
References


Tsai, Jerry, and Jessica A. Wachter, 2013, Rare booms and disasters in a multi-sector endowment economy, Working Paper.


Figure 1: Conditional skewness estimates from combination and univariate models

This figure plots estimates of quarterly conditional third central moments obtained by estimating (4) using a large cross-section of macroeconomic and financial variables. Dark blue lines plot the combined forecasts, which average over all univariate models using the 3PRF and inverse mean-squared-error combination weights, respectively. Thinner lines plot the implied indices from the 40 best-fitting univariate models (as measured by the total sum of squared residuals). Lines are color-coded to correspond with overall goodness-of-fit, where the darkest lines indicate the best fit.
Panel A: Quarterly poisson intensities for large shocks

Panel B: Implied cross-sectional distributions of annual wage growth (F12M / L12M)

Figure 2: Fitted dynamics of idiosyncratic distributions

Panel A plots the poisson intensities for good and bad shocks from the estimated model for the income process. Panel B plots the difference between the median and several quantiles of the model-implied distributions of year-on-year changes in income.
Figure 3: Log densities of permanent component of year-on-year income growth in expansions and recessions

This figure plots the log of the densities of year-on-year changes in permanent income ($\phi_4(L; 0)\eta_t$) in expansions and recessions, respectively. Dashed vertical lines correspond with the average change in log wages in expansions and recessions, respectively. See the text for further details.
Figure 4: Co-movement of labor market uncertainty proxies

This figure plots the co-movement of initial claims for unemployment, as a fraction of private payroll employment, my idiosyncratic skewness index, and a measure of cross-sectional employment growth volatility across U.S. states. Series are standardized to have mean zero and unit variance.
Figure 5: Co-movement of initial claims with representative equity premium forecasts

This figure plots the co-movement of initial claims for unemployment, expressed as a fraction of private payroll employment, with two measures of the equity risk premium: the market dividend-price ratio and an equal-weight combination of univariate forecasts from the Goyal and Welch (2008) predictors. All series are standardized to have mean zero and variance 1.
Figure 6: Model-free impulse responses to market excess return innovations

This figure plots model-free impulse responses of key macroeconomic variables to market excess return innovations. The impulse response is the slope coefficient on the market return, \( r_{m,t+1} \), from a univariate regression of \( y_{t+k} \) on a vector of predictors, \( x_t \), and \( r_{m,t+1} \). The vector \( z_t \) includes \( y_t \), the dividend yield, initial claims for unemployment, the term spread, and the 3 month T-bill rate. Shaded regions are pointwise 95% confidence bands, calculated using Newey-West standard errors, where the number of lags equals the horizon minus 1.
<table>
<thead>
<tr>
<th>Category</th>
<th>Number of Variables</th>
<th>$R^2$: 3$^{rd}$ moment 1 year</th>
<th>$R^2$: 3$^{rd}$ moment 5 year</th>
<th>Correlation w/ overall index</th>
</tr>
</thead>
<tbody>
<tr>
<td>All variables (3PRF)</td>
<td>109</td>
<td>70.1</td>
<td>80.8</td>
<td>100.0</td>
</tr>
<tr>
<td>All variables (Combination)</td>
<td>109</td>
<td>66.4</td>
<td>72.2</td>
<td>99.6</td>
</tr>
<tr>
<td>Real output and income</td>
<td>18</td>
<td>66.6</td>
<td>68.8</td>
<td>96.3</td>
</tr>
<tr>
<td>Employment and hours</td>
<td>24</td>
<td>57.7</td>
<td>63.3</td>
<td>95.8</td>
</tr>
<tr>
<td>Real inventories, orders, and unfilled orders</td>
<td>5</td>
<td>56.8</td>
<td>65.1</td>
<td>89.1</td>
</tr>
<tr>
<td>Stock returns and predictability state variables</td>
<td>14</td>
<td>51.4</td>
<td>74.4</td>
<td>68.9</td>
</tr>
<tr>
<td>Money and credit quantity aggregates</td>
<td>6</td>
<td>53.9</td>
<td>66.0</td>
<td>63.9</td>
</tr>
<tr>
<td>Interest rates</td>
<td>11</td>
<td>30.7</td>
<td>47.0</td>
<td>62.1</td>
</tr>
<tr>
<td>Housing starts and sales</td>
<td>7</td>
<td>11.2</td>
<td>13.4</td>
<td>49.7</td>
</tr>
<tr>
<td>Consumption</td>
<td>4</td>
<td>28.9</td>
<td>5.1</td>
<td>48.6</td>
</tr>
<tr>
<td>Consumer expectations</td>
<td>1</td>
<td>7.0</td>
<td>17.6</td>
<td>34.0</td>
</tr>
<tr>
<td>Price indices</td>
<td>15</td>
<td>7.6</td>
<td>16.9</td>
<td>25.6</td>
</tr>
<tr>
<td>Average hourly earnings</td>
<td>2</td>
<td>3.2</td>
<td>7.2</td>
<td>19.8</td>
</tr>
<tr>
<td>Exchange rates</td>
<td>2</td>
<td>12.5</td>
<td>15.1</td>
<td>12.5</td>
</tr>
</tbody>
</table>

**Table 1:** Goodness-of-fit statistics and variance decomposition of skewness index

This table presents the results from estimating equation (4) for different subsets of $y_t$. The dependent variables are 1 and 5-year 3$^{rd}$ central moments of the cross-sectional distribution of income growth rates for 1978-2011 from GOS. Columns report $R^2$ values at each horizon, as well as the correlation between the implied skewness measure and the overall measure in the top row. Parameters are estimated so as to minimize the sum of squared residuals for both 1 and 5-year measures. The specification in the first line uses the 3PRF, while the remaining specifications construct a linear index as a weighted average of the univariate forecasts from each variable in the category, using the inverse of the mean-squared error as weights. Variables and categories are listed in Appendix A.2.3. The parameter ($\rho$) governing the state dependent, temporary shock is 0.45.
<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Mean specification</th>
<th>First stage $R^2$</th>
<th>Absolute residuals $\text{Claims}_{t-1}$</th>
<th>$\text{x}_{t-1}$</th>
<th>$\text{Vol}_t$</th>
<th>Squared residuals $\text{Claims}_{t-1}$</th>
<th>$\text{x}_{t-1}$</th>
<th>$\text{Vol}_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skewness ($x_t$)</td>
<td>AR(1)</td>
<td>75.0</td>
<td>3.50</td>
<td>-4.31</td>
<td>3.35</td>
<td>2.63</td>
<td>-3.65</td>
<td>-4.31</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[19.1]</td>
<td>[18.1]</td>
<td>[10.6]</td>
<td>[16.8]</td>
<td>[17.6]</td>
<td>[18.1]</td>
</tr>
<tr>
<td></td>
<td>AR(2)</td>
<td>80.6</td>
<td>2.73</td>
<td>-3.21</td>
<td>3.19</td>
<td>2.43</td>
<td>-3.08</td>
<td>-3.21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[12.6]</td>
<td>[12.9]</td>
<td>[8.9]</td>
<td>[12.8]</td>
<td>[13.7]</td>
<td>[12.9]</td>
</tr>
<tr>
<td></td>
<td>AR(4)</td>
<td>80.8</td>
<td>2.69</td>
<td>-3.22</td>
<td>3.07</td>
<td>2.44</td>
<td>-3.15</td>
<td>-3.22</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[12.6]</td>
<td>[13.2]</td>
<td>[8.4]</td>
<td>[12.4]</td>
<td>[14.1]</td>
<td>[13.2]</td>
</tr>
<tr>
<td></td>
<td>VAR(1)</td>
<td>75.3</td>
<td>3.22</td>
<td>-4.09</td>
<td>3.39</td>
<td>2.55</td>
<td>-3.57</td>
<td>-4.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[17.5]</td>
<td>[17.3]</td>
<td>[10.1]</td>
<td>[14.6]</td>
<td>[17.2]</td>
<td>[17.3]</td>
</tr>
<tr>
<td></td>
<td>VAR(2)</td>
<td>88.0</td>
<td>3.73</td>
<td>-3.06</td>
<td>2.85</td>
<td>3.30</td>
<td>-3.27</td>
<td>-3.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[12.7]</td>
<td>[8.6]</td>
<td>[5.6]</td>
<td>[12.4]</td>
<td>[10.0]</td>
<td>[8.6]</td>
</tr>
<tr>
<td></td>
<td>VAR(4)</td>
<td>92.2</td>
<td>2.16</td>
<td>-2.35</td>
<td>3.15</td>
<td>2.23</td>
<td>-2.06</td>
<td>-2.35</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[2.8]</td>
<td>[3.5]</td>
<td>[6.4]</td>
<td>[2.6]</td>
<td>[3.8]</td>
<td>[3.5]</td>
</tr>
<tr>
<td>Claims$_t$</td>
<td>AR(1)</td>
<td>87.8</td>
<td>6.74</td>
<td>-3.42</td>
<td>4.08</td>
<td>3.94</td>
<td>-2.85</td>
<td>-3.42</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[21.4]</td>
<td>[6.7]</td>
<td>[27.4]</td>
<td>[10.1]</td>
<td>[3.1]</td>
<td>[6.7]</td>
</tr>
<tr>
<td></td>
<td>AR(2)</td>
<td>89.4</td>
<td>4.93</td>
<td>-2.98</td>
<td>4.25</td>
<td>2.70</td>
<td>-2.38</td>
<td>-2.98</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[23.6]</td>
<td>[5.4]</td>
<td>[27.6]</td>
<td>[9.5]</td>
<td>[1.9]</td>
<td>[5.4]</td>
</tr>
<tr>
<td></td>
<td>AR(4)</td>
<td>89.5</td>
<td>4.99</td>
<td>-2.95</td>
<td>4.25</td>
<td>2.65</td>
<td>-2.34</td>
<td>-2.95</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[22.8]</td>
<td>[4.9]</td>
<td>[27.6]</td>
<td>[9.3]</td>
<td>[1.7]</td>
<td>[4.9]</td>
</tr>
<tr>
<td></td>
<td>VAR(1)</td>
<td>87.9</td>
<td>6.50</td>
<td>-2.98</td>
<td>4.24</td>
<td>3.80</td>
<td>-2.57</td>
<td>-2.98</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[20.2]</td>
<td>[5.8]</td>
<td>[28.2]</td>
<td>[9.5]</td>
<td>[2.7]</td>
<td>[5.8]</td>
</tr>
<tr>
<td></td>
<td>VAR(2)</td>
<td>89.6</td>
<td>4.64</td>
<td>-2.79</td>
<td>4.19</td>
<td>2.68</td>
<td>-2.26</td>
<td>-2.79</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[22.5]</td>
<td>[5.2]</td>
<td>[27.1]</td>
<td>[9.6]</td>
<td>[1.9]</td>
<td>[5.2]</td>
</tr>
<tr>
<td></td>
<td>VAR(4)</td>
<td>89.9</td>
<td>4.72</td>
<td>-2.83</td>
<td>4.13</td>
<td>2.62</td>
<td>-2.30</td>
<td>-2.83</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>[20.3]</td>
<td>[4.9]</td>
<td>[26.6]</td>
<td>[8.5]</td>
<td>[1.6]</td>
<td>[4.9]</td>
</tr>
</tbody>
</table>

| Number of residuals | 179 | 179 | 179 | 179 | 179 | 179 |

**Table 2:** Tests for heteroskedasticity of skewness index and initial claims residuals

This table presents the results of a test for the heteroskedasticity of my estimated skewness index ($x_t$, top panels) as well as initial claims for unemployment, $\text{Claims}_t$. Test statistics are generated using a two-step procedure. In the first stage, I estimate an AR($p$) or VAR($p$) model using $x_t$ and $\text{Claims}_t$. I report the $R^2$ from this regression is in the third column. The left and right panels report Newey-West t-statistics (4 lags) and $R^2$ (in brackets) from regressions of absolute and squared residuals, respectively. The sample period is 1967:1 through 2012:3.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_0$</td>
<td>0.0065</td>
<td>Average idiosyncratic jump intensity</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>0.0026</td>
<td>Sensitivity of quarterly jump intensity to a one standard deviation change in $x_t$</td>
</tr>
<tr>
<td>$\mu_b$</td>
<td>-0.18</td>
<td>Average consumption decline given a disaster</td>
</tr>
<tr>
<td>$\sigma_b$</td>
<td>0.115</td>
<td>Standard deviation of disaster magnitude</td>
</tr>
<tr>
<td>$\rho_x$</td>
<td>0.8847</td>
<td>Persistence of $x_t$ process</td>
</tr>
<tr>
<td>$\rho_\sigma$</td>
<td>0.9446</td>
<td>Persistence of $\sigma_t^2$ process</td>
</tr>
<tr>
<td>$\mu_c$</td>
<td>0.02 / 4</td>
<td>Drift of consumption growth</td>
</tr>
<tr>
<td>$\mu_d$</td>
<td>0.0075 / 4</td>
<td>Drift of dividend growth</td>
</tr>
<tr>
<td>$\phi_c$</td>
<td>0.000366</td>
<td>Loading of expected consumption growth on $x_t$</td>
</tr>
<tr>
<td>$\phi_x$</td>
<td>0.025</td>
<td>Loading of expected dividend growth on $x_t$</td>
</tr>
<tr>
<td>$\varphi_c$</td>
<td>0.0125</td>
<td>Standard deviation of shock to $\Delta c_t$</td>
</tr>
<tr>
<td>$\varphi_d$</td>
<td>0.045</td>
<td>Standard deviation of independent shock to $\Delta d_t$</td>
</tr>
<tr>
<td>$\varphi_x$</td>
<td>$\sqrt{1 - \rho_x^2}$</td>
<td>Standard deviation of shock to $x_t$</td>
</tr>
<tr>
<td>$\varphi_\sigma$</td>
<td>0.1674</td>
<td>Standard deviation of shock to $\sigma_t^2$</td>
</tr>
<tr>
<td>$\pi_c$</td>
<td>2.5</td>
<td>Loading of dividend innovation on $\Delta c_t$ innovation</td>
</tr>
<tr>
<td>$\pi_x$</td>
<td>0.04</td>
<td>Loading of dividend innovation on $x_t$ innovation</td>
</tr>
<tr>
<td>$\pi_\sigma$</td>
<td>-0.0896</td>
<td>Loading of dividend innovation on $\sigma_t^2$ innovation</td>
</tr>
<tr>
<td>$\chi$</td>
<td>-0.66</td>
<td>Correlation of shocks to $x_t$ and $\sigma_t^2$</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>11</td>
<td>Relative risk aversion coefficient</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.9745</td>
<td>Rate of time preference (quarterly)</td>
</tr>
<tr>
<td>$\psi$</td>
<td>2</td>
<td>Intertemporal elasticity of substitution</td>
</tr>
</tbody>
</table>

Table 3: Summary of Parameters for the Quantitative Model

This table describes the parameters of the quantitative asset pricing model, along with the calibrated values. The time horizon of the model is quarterly. The additional free parameters, $\mu_\eta$ and $F_\eta$, are assumed without loss of generality to equal the expressions given in Lemma 1.
<table>
<thead>
<tr>
<th>Statistic</th>
<th>Baseline</th>
<th>BY</th>
<th>BKY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Persistence of $x_t$ (monthly)</td>
<td>0.96</td>
<td>0.979</td>
<td>0.975</td>
</tr>
<tr>
<td>Half-life of $x_t$ shocks</td>
<td>1.4 years</td>
<td>2.7 years</td>
<td>2.3 years</td>
</tr>
<tr>
<td>Persistence of $\sigma_t^2$ (monthly)</td>
<td>0.980</td>
<td>0.987</td>
<td>0.999</td>
</tr>
<tr>
<td>Half-life of $\sigma_t^2$ shock</td>
<td>2.9 years</td>
<td>4.4 years</td>
<td>57.7 years</td>
</tr>
<tr>
<td>Volatility of expected consumption growth (annualized)</td>
<td>0.15%</td>
<td>2.0%</td>
<td>1.5%</td>
</tr>
<tr>
<td>Volatility of expected dividend growth (annualized)</td>
<td>10.0%</td>
<td>6.1%</td>
<td>3.7%</td>
</tr>
<tr>
<td>Dividend growth predictability: $E_t \left[ \sum_{j=0}^{\infty} \rho^j [\Delta d_{t+1+j} - E(\Delta d_{t+1+j})</td>
<td>x_t = \sigma(x_t)] \right]$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0% annual discount rate</td>
<td>21.7%</td>
<td>24.6%</td>
<td>12.6%</td>
</tr>
<tr>
<td>5% annual discount rate</td>
<td>19.8%</td>
<td>20.6%</td>
<td>10.9%</td>
</tr>
<tr>
<td>10% annual discount rate</td>
<td>18.1%</td>
<td>17.6%</td>
<td>9.5%</td>
</tr>
</tbody>
</table>

**Table 4:** Comparison with benchmark long-run risk calibrations

This table compares several features of my calibrated model with comparable values implied by the parameters in Bansal and Yaron (2004, BY) and Bansal, Kiku, and Yaron (2012, BKY).
<table>
<thead>
<tr>
<th>Moment</th>
<th>Data Estimate</th>
<th>Model Median</th>
<th>2.5%</th>
<th>5%</th>
<th>25%</th>
<th>75%</th>
<th>95%</th>
<th>97.5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E[R_m - R_f]$</td>
<td>7.09</td>
<td>6.46</td>
<td>2.19</td>
<td>2.90</td>
<td>5.02</td>
<td>7.91</td>
<td>9.93</td>
<td>10.61</td>
</tr>
<tr>
<td>$\sigma(R_m)$</td>
<td>20.28</td>
<td>23.61</td>
<td>18.50</td>
<td>19.27</td>
<td>21.73</td>
<td>25.69</td>
<td>29.12</td>
<td>30.47</td>
</tr>
<tr>
<td>$E[R_f]$</td>
<td>0.57</td>
<td>0.46</td>
<td>-1.61</td>
<td>-1.22</td>
<td>-0.19</td>
<td>1.05</td>
<td>1.83</td>
<td>2.07</td>
</tr>
<tr>
<td>$\sigma(R_f)$</td>
<td>2.86</td>
<td>3.46</td>
<td>2.50</td>
<td>2.63</td>
<td>3.08</td>
<td>3.87</td>
<td>4.58</td>
<td>4.83</td>
</tr>
<tr>
<td>$E[\Delta c]$</td>
<td>1.93</td>
<td>2.00</td>
<td>1.43</td>
<td>1.53</td>
<td>1.81</td>
<td>2.19</td>
<td>2.46</td>
<td>2.54</td>
</tr>
<tr>
<td>$\sigma(\Delta c)$</td>
<td>2.16</td>
<td>2.04</td>
<td>1.59</td>
<td>1.66</td>
<td>1.87</td>
<td>2.21</td>
<td>2.47</td>
<td>2.57</td>
</tr>
<tr>
<td>$AC1(\Delta c)$</td>
<td>0.45</td>
<td>0.21</td>
<td>-0.01</td>
<td>0.02</td>
<td>0.13</td>
<td>0.28</td>
<td>0.39</td>
<td>0.42</td>
</tr>
<tr>
<td>$E[\Delta d]$</td>
<td>1.15</td>
<td>0.83</td>
<td>-5.64</td>
<td>-4.46</td>
<td>-1.27</td>
<td>2.77</td>
<td>5.33</td>
<td>6.15</td>
</tr>
<tr>
<td>$\sigma(\Delta d)$</td>
<td>11.05</td>
<td>13.78</td>
<td>10.33</td>
<td>10.83</td>
<td>12.50</td>
<td>15.20</td>
<td>17.68</td>
<td>18.61</td>
</tr>
<tr>
<td>$AC1(\Delta d)$</td>
<td>0.21</td>
<td>0.50</td>
<td>0.28</td>
<td>0.32</td>
<td>0.43</td>
<td>0.57</td>
<td>0.66</td>
<td>0.68</td>
</tr>
<tr>
<td>$Corr(\Delta c, \Delta d)$</td>
<td>0.55</td>
<td>0.41</td>
<td>0.16</td>
<td>0.21</td>
<td>0.33</td>
<td>0.49</td>
<td>0.58</td>
<td>0.61</td>
</tr>
</tbody>
</table>

Table 5: Bootstrapped distribution of model-implied moments

This table presents several moments of aggregate cash flows and asset prices, both from the data and the model. The data moments are reproduced from Bansal, Kiku, and Yaron (2012), who use real, annual data from 1930-2008. The remaining columns show the Monte Carlo distributions of 50,000 simulated paths of analogous quantities, which are simulated from the calibrated model and time-aggregated to an annual frequency. Each simulated path has the same length as the historical data.
<table>
<thead>
<tr>
<th>Moment</th>
<th>Data Estimate</th>
<th>Markets</th>
<th>Baseline</th>
<th>No Covariance</th>
<th>Less CF Vol</th>
<th>IID Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E[R_m - R_f]$</td>
<td>7.1</td>
<td>Inc</td>
<td>6.5</td>
<td>5.5</td>
<td>5.8</td>
<td>5.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RA</td>
<td>3.2</td>
<td>3.1</td>
<td>2.5</td>
<td>2.0</td>
</tr>
<tr>
<td>$\sigma(R_m)$</td>
<td>20.3</td>
<td>Inc</td>
<td>24.2</td>
<td>20.2</td>
<td>23.4</td>
<td>23.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RA</td>
<td>26.9</td>
<td>23.5</td>
<td>26.0</td>
<td>25.5</td>
</tr>
<tr>
<td>$E[R_f]$</td>
<td>0.6</td>
<td>Inc</td>
<td>0.4</td>
<td>0.4</td>
<td>0.6</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RA</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>$\sigma(R_f)$</td>
<td>2.9</td>
<td>Inc</td>
<td>3.7</td>
<td>3.7</td>
<td>3.5</td>
<td>3.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RA</td>
<td>0.3</td>
<td>0.3</td>
<td>0.1</td>
<td>0.0</td>
</tr>
<tr>
<td>$E[pd]$</td>
<td>3.4</td>
<td>Inc</td>
<td>3.4</td>
<td>3.4</td>
<td>3.5</td>
<td>3.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RA</td>
<td>4.5</td>
<td>4.1</td>
<td>5.1</td>
<td>6.2</td>
</tr>
<tr>
<td>$\sigma(pd)$</td>
<td>0.45</td>
<td>Inc</td>
<td>0.22</td>
<td>0.21</td>
<td>0.21</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RA</td>
<td>0.25</td>
<td>0.25</td>
<td>0.24</td>
<td>0.24</td>
</tr>
<tr>
<td>$AC1(pd)$</td>
<td>0.87</td>
<td>Inc</td>
<td>0.62</td>
<td>0.64</td>
<td>0.60</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RA</td>
<td>0.59</td>
<td>0.62</td>
<td>0.58</td>
<td>0.57</td>
</tr>
<tr>
<td>$E[\Delta c]$</td>
<td>1.9</td>
<td>Both</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>$\sigma(\Delta c)$</td>
<td>2.2</td>
<td>Both</td>
<td>2.1</td>
<td>2.1</td>
<td>2.1</td>
<td>2.1</td>
</tr>
<tr>
<td>$AC1(\Delta c)$</td>
<td>0.45</td>
<td>Both</td>
<td>0.23</td>
<td>0.23</td>
<td>0.23</td>
<td>0.23</td>
</tr>
<tr>
<td>$E[\Delta d]$</td>
<td>1.15</td>
<td>Both</td>
<td>0.68</td>
<td>0.75</td>
<td>0.70</td>
<td>0.80</td>
</tr>
<tr>
<td>$\sigma(\Delta d)$</td>
<td>11.1</td>
<td>Both</td>
<td>14.5</td>
<td>12.8</td>
<td>14.4</td>
<td>14.4</td>
</tr>
<tr>
<td>$AC1(\Delta d)$</td>
<td>0.21</td>
<td>Both</td>
<td>0.53</td>
<td>0.50</td>
<td>0.53</td>
<td>0.53</td>
</tr>
<tr>
<td>Corr($\Delta c, \Delta d$)</td>
<td>0.55</td>
<td>Both</td>
<td>0.40</td>
<td>0.44</td>
<td>0.41</td>
<td>0.36</td>
</tr>
</tbody>
</table>

**Table 6:** Data and model-implied moments for different specifications

This table presents several moments of aggregate cash flows and asset prices, from the data and different versions of the model. The data moments are reproduced from Bansal, Kiku, and Yaron (2012), who use real, annual data from 1930-2008. The remaining columns show the averages over a simulation of 1 million quarters of analogous quantities, which are simulated from the calibrated model and time-aggregated to an annual frequency. Please see the text for the parameter restrictions associated with the different specifications.
### Table 7: Correlations between labor market variables and predictor variables

This table reports univariate correlation coefficients between a number of monthly time series. Initial claims for unemployment insurance, divided by private sector employment, is my proxy for labor market uncertainty. Future employment cuts, is the negative of the logarithmic growth rate in private payroll employment over the next 3 months. The table also includes the Goyal and Welch (2008) predictors and combination forecasts which are constructed from these predictors. The first two combination forecasts are estimated using different sample periods. The last combination forecast uses predictors which are orthogonalized with respect to initial claims. Stars indicate statistical significance at the 1% level.
### Table 8: Predictive regressions for excess returns on Market and SMB portfolios

This table plots the $R^2$ values (in percentage points) from predictive regressions of cumulative returns on a number of univariate state variables. I consider the market excess return as well as the Fama and French (1993) SMB portfolio. I use overlapping monthly data for the regressions, and the sample period is 1967-2012, the period for which initial claims data are available. Newey-West t-statistics, with lag length equal to the forecast horizon minus 1, are in parentheses.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Market Excess Return</th>
<th></th>
<th></th>
<th>SMB Return</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3 mo</td>
<td>1 yr</td>
<td>2 yr</td>
<td>3 mo</td>
<td>1 yr</td>
<td>2 yr</td>
<td></td>
</tr>
<tr>
<td>Initial claims</td>
<td>2.40**</td>
<td>6.37**</td>
<td>4.96</td>
<td>4.29***</td>
<td>10.82***</td>
<td>9.47***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.42)</td>
<td>(2.44)</td>
<td>(1.58)</td>
<td>(3.63)</td>
<td>(3.90)</td>
<td>(2.58)</td>
<td></td>
</tr>
<tr>
<td>Equal-weighted forecast combinations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1928 - 2012</td>
<td>2.22**</td>
<td>8.96***</td>
<td>9.34**</td>
<td>3.25***</td>
<td>5.52*</td>
<td>2.69</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.07)</td>
<td>(2.71)</td>
<td>(2.15)</td>
<td>(2.61)</td>
<td>(1.91)</td>
<td>(0.92)</td>
<td></td>
</tr>
<tr>
<td>1967 - 2012</td>
<td>3.39**</td>
<td>14.88***</td>
<td>16.95***</td>
<td>3.84***</td>
<td>4.73**</td>
<td>1.50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.05)</td>
<td>(4.19)</td>
<td>(3.38)</td>
<td>(3.30)</td>
<td>(2.27)</td>
<td>(0.96)</td>
<td></td>
</tr>
<tr>
<td>1967 - 2012 (orth. predictors)</td>
<td>2.41</td>
<td>9.49***</td>
<td>10.70***</td>
<td>2.10**</td>
<td>1.62</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.64)</td>
<td>(2.82)</td>
<td>(3.57)</td>
<td>(2.53)</td>
<td>(1.51)</td>
<td>(0.32)</td>
<td></td>
</tr>
<tr>
<td>Univariate regressions with Goyal and Welch (2008) predictors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dy</td>
<td>0.97</td>
<td>3.12</td>
<td>4.29</td>
<td>1.16</td>
<td>1.71</td>
<td>0.48</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.44)</td>
<td>(1.47)</td>
<td>(1.31)</td>
<td>(1.44)</td>
<td>(1.06)</td>
<td>(0.33)</td>
<td></td>
</tr>
<tr>
<td>ep</td>
<td>0.32</td>
<td>0.79</td>
<td>0.54</td>
<td>0.36</td>
<td>0.50</td>
<td>0.31</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.48)</td>
<td>(-0.60)</td>
<td>(-0.42)</td>
<td>(-0.70)</td>
<td>(-0.48)</td>
<td>(-0.21)</td>
<td></td>
</tr>
<tr>
<td>bm</td>
<td>0.29</td>
<td>0.76</td>
<td>0.38</td>
<td>1.32*</td>
<td>4.23*</td>
<td>4.18</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.50)</td>
<td>(0.63)</td>
<td>(0.31)</td>
<td>(1.73)</td>
<td>(1.91)</td>
<td>(1.26)</td>
<td></td>
</tr>
<tr>
<td>svar</td>
<td>0.31</td>
<td>0.85*</td>
<td>1.72**</td>
<td>0.47</td>
<td>1.59**</td>
<td>2.10*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.44)</td>
<td>(1.75)</td>
<td>(2.26)</td>
<td>(1.58)</td>
<td>(2.06)</td>
<td>(1.84)</td>
<td></td>
</tr>
<tr>
<td>tbl</td>
<td>0.67</td>
<td>1.13</td>
<td>1.33</td>
<td>0.63</td>
<td>0.52</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.10)</td>
<td>(-0.84)</td>
<td>(-1.03)</td>
<td>(-1.35)</td>
<td>(-0.64)</td>
<td>(-0.19)</td>
<td></td>
</tr>
<tr>
<td>tms</td>
<td>1.76**</td>
<td>8.04***</td>
<td>13.11***</td>
<td>0.98</td>
<td>1.48</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.01)</td>
<td>(2.67)</td>
<td>(3.45)</td>
<td>(1.55)</td>
<td>(1.24)</td>
<td>(0.13)</td>
<td></td>
</tr>
<tr>
<td>dfy</td>
<td>1.29</td>
<td>3.82*</td>
<td>3.62</td>
<td>2.89***</td>
<td>5.22***</td>
<td>5.26**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.30)</td>
<td>(1.88)</td>
<td>(1.50)</td>
<td>(3.17)</td>
<td>(2.75)</td>
<td>(2.52)</td>
<td></td>
</tr>
<tr>
<td>lty</td>
<td>0.19</td>
<td>0.42</td>
<td>0.91</td>
<td>0.26</td>
<td>0.18</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.17)</td>
<td>(0.40)</td>
<td>(0.61)</td>
<td>(-0.63)</td>
<td>(-0.04)</td>
<td>(-0.17)</td>
<td></td>
</tr>
<tr>
<td>ntis</td>
<td>0.21</td>
<td>0.43</td>
<td>0.43</td>
<td>0.22</td>
<td>0.20</td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.20)</td>
<td>(-0.31)</td>
<td>(-0.34)</td>
<td>(-0.37)</td>
<td>(-0.13)</td>
<td>(-0.13)</td>
<td></td>
</tr>
<tr>
<td>infl</td>
<td>0.70</td>
<td>2.95**</td>
<td>1.57*</td>
<td>0.35</td>
<td>0.82</td>
<td>1.85</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.94)</td>
<td>(-2.09)</td>
<td>(-1.90)</td>
<td>(-0.88)</td>
<td>(1.26)</td>
<td>(1.42)</td>
<td></td>
</tr>
<tr>
<td>dfcfr</td>
<td>0.92</td>
<td>0.31</td>
<td>0.23</td>
<td>0.37</td>
<td>0.44*</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.46)</td>
<td>(0.88)</td>
<td>(0.55)</td>
<td>(1.21)</td>
<td>(1.78)</td>
<td>(0.86)</td>
<td></td>
</tr>
<tr>
<td>ltr</td>
<td>0.62</td>
<td>1.64***</td>
<td>0.78***</td>
<td>0.48</td>
<td>0.23</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.27)</td>
<td>(3.41)</td>
<td>(2.74)</td>
<td>(1.31)</td>
<td>(0.53)</td>
<td>(-0.27)</td>
<td></td>
</tr>
</tbody>
</table>
Appendix

A  Idiosyncratic Risk Process - Calibration and Estimation

A.1  Additional summary statistics for income growth rates

This section provides more detail about several key results from GOS about the nature of idiosyncratic labor income risk. In addition to providing nonparametric evidence of with time-varying, idiosyncratic tail risk, these statistics are inputs for my calibrated model of the labor income process in Section 2.3.

One obtains an intuitive measure of the asymmetry of a distribution by considering three conditional quantiles. A robust measure of skewness is Kelley’s skewness, which is defined as \[ \frac{Q_{90} - Q_{50}}{Q_{90} - Q_{10}} - \frac{Q_{50} - Q_{10}}{Q_{90} - Q_{10}}. \] The denominator is the distance between the 10\textsuperscript{th} and 90\textsuperscript{th} percentiles, a measure of the overall spread of the distribution. GOS show that, over longer horizons, the denominator is almost constant. The numerator splits \( Q_{90} - Q_{10} \) into two pieces. The first, \( Q_{90} - Q_{50} \), measures the width of the right tail, while the latter, \( Q_{50} - Q_{10} \), measures the width of the left tail. In most cases, increases in the former distance are “good”, indicating a higher likelihood of seeing large increases in wages. Increases in the latter distance indicate a higher exposure to large declines in wages.

Figure 7 shows the time series evolution of these spreads, for 1, 3, and 5 year trailing changes in wages. These statistics pool all observations in their sample, giving a snapshot of the entire cross-sectional distribution of wage changes across the U.S. population. As the economy moves from an expansion to a recession, the left tail of the distribution (\( Q_{50} - Q_{10} \), in Panel A) expands, indicating an increased likelihood of experiencing large decreases in wages, while the width of the right tail (\( Q_{90} - Q_{50} \), in Panel B) shrinks. There are more big losers in recessions and fewer big winners. Note that my use of trailing growth rates in the graphs means that the long horizon measures will tend to lag the recession bars.

Table 9 summarizes a number of GOS’s results on the distribution of 5-year wage changes, which control for cohort and life-cycle fixed effects and individuals’ previous earnings.\textsuperscript{59} For each year in

\[ \text{An alternative name for this measure, which is more popular in the finance literature, is “conditional asymmetry”. See, e.g., Ghysels et al. (2013).} \]

\[ \text{For example, a Kelly’s skewness of -20\% implies that the left tail makes up 60\% of the spread between the 10\textsuperscript{th} and 90\textsuperscript{th} percentiles, while the right tail contributes the remaining 40\% of the distance.} \]

\[ \text{A potential critique of Figure 7 is that the reported statistics pool the entire population of male earners together, which could overstate the asymmetry of the distribution of idiosyncratic shocks. Therefore, it is important to control for other observable characteristics as well, particularly lagged earnings. GOS control for lagged} \]

\[ \text{A.1} \]
### Panel A: Difference between 50th and 10th percentiles of trailing k year real income growth rate

<table>
<thead>
<tr>
<th>Year</th>
<th>Difference between quantiles</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980</td>
<td>0.3</td>
</tr>
<tr>
<td>1985</td>
<td>0.4</td>
</tr>
<tr>
<td>1990</td>
<td>0.5</td>
</tr>
<tr>
<td>1995</td>
<td>0.6</td>
</tr>
<tr>
<td>2000</td>
<td>0.7</td>
</tr>
<tr>
<td>2005</td>
<td>0.8</td>
</tr>
<tr>
<td>2010</td>
<td>0.9</td>
</tr>
</tbody>
</table>

### Panel B: Difference between 90th and 50th percentiles of trailing k year real income growth rate

<table>
<thead>
<tr>
<th>Year</th>
<th>Difference between quantiles</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980</td>
<td>0.3</td>
</tr>
<tr>
<td>1985</td>
<td>0.4</td>
</tr>
<tr>
<td>1990</td>
<td>0.5</td>
</tr>
<tr>
<td>1995</td>
<td>0.6</td>
</tr>
<tr>
<td>2000</td>
<td>0.7</td>
</tr>
<tr>
<td>2005</td>
<td>0.8</td>
</tr>
<tr>
<td>2010</td>
<td>0.9</td>
</tr>
</tbody>
</table>

**Figure 7:** Dynamic evolution of the cross-section of income growth rates over time

Panel A plots the evolution of the distance between the 50th and 10th percentiles, a measure of the width of the left tail, of the cross-sectional distribution of 1, 3, and 5-year trailing real income growth rates from the 10% sample of Social Security earnings records in GOS. Panel B reports the distance between the 90th and 50th percentiles, a measure of the width of the right tail. Data are from GOS Appendix Table A.13, which reports linearly detrended cross-sectional quantiles.
Table 9: Summary statistics for the cross sectional distribution of income growth rates

This table summarizes a number of statistics from the cross-section of 5-year log income growth rates, which are calculated from statistics reported by GOS using annual data from 1978-2011. I report the average of each statistic over the 91st through 95th percentiles of the 5-year average income distribution (see GOS for a detailed definition) and over time. The second column indicates the period over which the average value of the statistic is calculated, where “E”, “R”, and “R - E” denote expansions, recessions, and the difference between recessions and expansions, respectively.

their sample and for each of 100 different groups formed based on lagged wages, GOS calculate a number of quantiles of the cross-sectional distribution of income growth rates. They then average these statistics over expansion periods and recession periods, and compare the average levels of the different quantiles in expansions with those in recessions. In their classification, recession periods begin one year prior to the start of the recession and end several years after the recession has ended, in order to emphasize persistent changes in wages from recessions, as opposed to more temporary declines in income such as lost wages during unemployment spells.60 Expansions are 5-year periods which do not include a recession year.

The left panel of Table 9 reports the median, 10th percentile, 90th percentile, and Kelley’s skewness of five year income growth rates in expansions and recessions, respectively. The right panel reports the changes in quantile-based measures of scale—the inter-quartile range and the 90-10 percentile spread—in recessions versus expansions. The right panel also reports quantile-based earnings nonparametrically, placing each individual into one of 100 bins based upon his earnings over the previous 5 years, though similar results obtain when also controlling for age. I refer the reader to GOS for further details on this procedure.

measures of the width of the left and right tails of the cross-sectional distribution, respectively. Recall that increases in the width of the left tail indicate higher risk exposures.

Several features of Table 9 are particularly striking. First, high earners face a substantial degree of idiosyncratic labor income risk, even in expansions. The average 90-10 spread is 115 log percentage points. Second, the entire distribution shifts to the left in bad times; all of the quantiles are strictly lower in recessions relative to expansions. This shift is not specific to the three quantiles in the left panel. All of the reported quantiles are lower in recessions.

Third, the change in the 10th and 90th percentiles is larger than the change in the median, meaning that width of the left tail expands in recessions, while the right tail shrinks (as was the case in Figure 7). This cyclical asymmetry is reflected by the change in Kelley’s skewness, which decreases by 14 percentage points. In contrast, both measures of the overall spread of the distribution changes very little over the cycle. GOS demonstrate that a similar result holds for second moments, particularly at longer horizons.

Finally, and perhaps most importantly, the tails of the idiosyncratic wage growth distribution, as measured by extreme quantiles, are much more responsive to the cycle than the center of the distribution. Over 5 year periods which include a recession, the median change in wages is 3.8 log percentage points lower relative to expansions. Scale measures barely change at all. However, the extreme quantiles of income growth rates are highly cyclical. The 50-10 spread increases by 7.7 log points, indicating a higher risk of large wage declines, while the 50-25 and 75-50 spreads moves much less (3 and -1.7 log points, respectively). Turning to the right tail, where more statistics are available, I find that the 90-50 spread shrinks by a magnitude comparable to the 50-10 spread, while the more extreme tail quantiles contract by considerably larger amounts. The 95-50 and 99-50 spreads shrink by 15.5 and 26.7 log points, respectively.

The results in Table 9 suggest that, for those individuals who receive idiosyncratic shocks from the center of the distribution, the business cycle has a relatively mild impact on their labor income. However, for those who experience larger shocks, the cycle has a substantial quantitative impact. Ex post, aggregate shocks appear to be disproportionately borne by a small fraction of the population. Section 2.3 replicates these features with a simple model where labor income is exposed to infrequent but very large shocks whose distribution is state-dependent.

Table 9 reported a number of statistics for the cross-sectional distribution of income growth rates, which were averaged over the 91st through 95th percentiles of the earnings distribution. Table 10 shows that the same results hold for different segments of the earnings distribution. It reports the same statistics, averaged over different percentiles of the distribution. These ranges are indicated by different columns, where [96,100] in the first column indicates that we
### Table 10: Summary statistics for the cross sectional distribution of income growth rates

This table summarizes a number of statistics from the cross-section of 5-year log income growth rates, which are calculated from statistics reported by GOS using annual data from 1978-2011. Columns indicate averages of the statistic over different percentiles of the 5-year average income distribution (see GOS for a detailed definition), where 1 and 100 indicate the lowest and highest 1% of earners, respectively. The second column indicates the period over which the average value of the statistic is calculated, where “E”, “R”, and “R - E” denote expansions, recessions, and the difference between recessions and expansions, respectively.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Period</th>
<th>Average over percentiles of 5-year average income distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>[96,100]</td>
</tr>
<tr>
<td>Median</td>
<td>E</td>
<td>2.71</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>-4.46</td>
</tr>
<tr>
<td>10th Percentile</td>
<td>E</td>
<td>-76.48</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>-95.34</td>
</tr>
<tr>
<td>90th Percentile</td>
<td>E</td>
<td>76.63</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>53.51</td>
</tr>
<tr>
<td></td>
<td>R - E</td>
<td>-20.47</td>
</tr>
<tr>
<td>Kelley’s Skewness</td>
<td>E</td>
<td>-2.22</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>-22.30</td>
</tr>
</tbody>
</table>

Scale Measures

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>[96,100]</th>
<th>[91,100]</th>
<th>[91,95]</th>
<th>[76,100]</th>
<th>[76,95]</th>
<th>[76,90]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inter-Quartile Range</td>
<td>R - E</td>
<td>0.43</td>
<td>0.86</td>
<td>1.30</td>
<td>0.78</td>
<td>0.87</td>
<td>0.73</td>
</tr>
<tr>
<td>90-10 Percentile Spread</td>
<td>R - E</td>
<td>-4.26</td>
<td>-2.44</td>
<td>-0.62</td>
<td>-1.22</td>
<td>-0.46</td>
<td>-0.40</td>
</tr>
</tbody>
</table>

Left Tail Width Measures

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>[96,100]</th>
<th>[91,100]</th>
<th>[91,95]</th>
<th>[76,100]</th>
<th>[76,95]</th>
<th>[76,90]</th>
</tr>
</thead>
<tbody>
<tr>
<td>50-25 Percentile Spread</td>
<td>R - E</td>
<td>4.08</td>
<td>3.55</td>
<td>3.01</td>
<td>2.67</td>
<td>2.32</td>
<td>2.09</td>
</tr>
<tr>
<td>50-10 Percentile Spread</td>
<td>R - E</td>
<td>10.85</td>
<td>9.28</td>
<td>7.70</td>
<td>7.72</td>
<td>6.94</td>
<td>6.69</td>
</tr>
</tbody>
</table>

Right Tail Width Measures

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>[96,100]</th>
<th>[91,100]</th>
<th>[91,95]</th>
<th>[76,100]</th>
<th>[76,95]</th>
<th>[76,90]</th>
</tr>
</thead>
<tbody>
<tr>
<td>75-50 Percentile Spread</td>
<td>R - E</td>
<td>-3.66</td>
<td>-2.69</td>
<td>-1.72</td>
<td>-1.89</td>
<td>-1.45</td>
<td>-1.36</td>
</tr>
<tr>
<td>90-50 Percentile Spread</td>
<td>R - E</td>
<td>-12.97</td>
<td>-10.54</td>
<td>-8.12</td>
<td>-8.61</td>
<td>-7.52</td>
<td>-7.32</td>
</tr>
</tbody>
</table>

are describing the risk faced by the top 5% of earners. Thus, row one, column three of Table 9 reports the median changes in log income, averaged over the top 5 percentiles of the 5-year average income distribution and over 5 expansion periods.

Regardless of the specific group (column of Table 10) considered, the results are consistent with the discussion in the main text. The overall level of idiosyncratic risk is extremely high, with the level of the 90-10 spread exceeding 100 log percentage points for all groups. The entire distribution shifts to the left in recessions. Scale measures are relatively insensitive to the business cycle, while the extreme tails move much more strongly. Finally, including the highest
group of earners increases the overall degree of risk substantially.

### A.2 Skewness Index Estimation

#### A.2.1 Proof of Proposition 1

This section derives expressions for the moments of time-aggregated wages from my quarterly model in (1-2). For notational simplicity, I suppress \( i \) subscripts here. A first-order Taylor expansion yields that, for \( k \geq 4 \),

\[
w_{A,t} - w_{A,t-k} \approx \frac{1}{4} \Delta w_t + \frac{1}{2} \Delta w_{t-1} + \frac{3}{4} \Delta w_{t-2} + \sum_{j=3}^{k-1} \Delta w_{t-j} + \frac{3}{4} \Delta w_{t-k} + \frac{1}{2} \Delta w_{t-k-1} + \frac{1}{4} \Delta w_{t-k-2}.
\]

\[
\begin{align*}
&= \beta \cdot k + \frac{1}{4} \eta_t + \frac{1}{2} \eta_{t-1} + \frac{3}{4} \eta_{t-2} + \sum_{j=3}^{k-1} \eta_{t-j} + \frac{3}{4} \eta_{t-k} + \frac{1}{2} \eta_{t-k-1} + \frac{1}{4} \eta_{t-k-2} \\
&+ \frac{1}{4} \rho(L)(\eta_t + \eta_{t-1} + \eta_{t-2} + \eta_{t-3}) - \frac{1}{4} \rho(L)(\eta_{t-k} + \eta_{t-k-1} + \eta_{t-k-2} + \eta_{t-k-3}) \\
&+ \frac{1}{4} (\epsilon_t + \epsilon_{t-1} + \epsilon_{t-2} + \epsilon_{t-3}) - \frac{1}{4} (\epsilon_{t-k} + \epsilon_{t-k-1} + \epsilon_{t-k-2} + \epsilon_{t-k-3}) \\
\end{align*}
\]

\[
\equiv \beta \cdot k + \theta_k(L; \rho)\eta_t + \epsilon_{A,t} - \epsilon_{A,t-k},
\]  

(26)

where \( \theta_k(\cdot) \) is a polynomial in the lag operator whose second argument is the vector of coefficients for \( \rho(L) \). Next, I link the third central moments of time-aggregated wages with moments from the quarterly model. Since \( \eta_t \) is independent of \( \eta_{t-j} \) given the path of the aggregate state, then

\[
M_3[y_{A,t} - y_{A,t-k}] = k^3 M_3[\beta] + \sum_{j=0}^{\infty} [\theta_{k,j}(\rho)]^3 M_3[\eta_{t-j}] + M_3[\epsilon_{A,t} - \epsilon_{A,t-k}],
\]

\[
\equiv \phi_k(L; \rho) M_3[\eta_t] + k^3 M_3[\beta] + M_3[\epsilon_{A,t} - \epsilon_{A,t-k}],
\]  

(27)

where \( M_3(\cdot) \) denotes the third central moment, conditional on aggregate information.\(^{61}\) If we further assume that \( M_3(\eta_t) = a + b' y_t \), where \( y_t \) is a vector of observable state variables, then

\[
M_3[w_{A,t} - w_{A,t-k}] = c_k + b' \phi_k(L; \rho) z_t,
\]  

(28)

---

\(^{61}\)This follows because, given two independent random variables \( x \) and \( z \) with \( \mu_x \equiv E[x] \) and \( \mu_z \equiv E[z] \),

\[
\begin{align*}
E[(x + z - \mu_x - \mu_z)^3] &= E[(x - \mu_x)^3] + 3E[(x - \mu_x)^2(z - \mu_z)] + 3E[(x - \mu_x)(z - \mu_z)^2] + E[(z - \mu_z)^3] \\
&= E[(x - \mu_x)^3] + E[(z - \mu_z)^3] \equiv M_3[x] + M_3[z],
\end{align*}
\]

where we use independence to replace terms such as \( E[(x - \mu_x)^2(z - \mu_z)] \) with \( E[(x - \mu_x)^2]E[(z - \mu_z)] = 0. \)
where $\phi_k(L; \rho)$ is a known lag polynomial and $c_k \equiv \phi_k(1; \rho)a + k^3 M_3[\beta] + M_3[\epsilon_{A,t} + \epsilon_{A,t-k}]$, which is constant given our assumption that the third moments of $\beta_i$ and $\epsilon_t$ are state independent.

A.2.2 Skewness Indices - Parametric Approach

Motivation for Explanatory Variables

I consider four macroeconomic time series for inclusion in the vector $y_t$. The first variable, $\Delta emp_t$, is the quarterly change in the logarithm of private payroll employment. In section 2.3, I found that the tails of the income growth distribution are much more sensitive to the cycle relative to the center. If tail events are related to transitions between jobs, one would expect to see more large positive shocks and fewer large negative shocks when firms are hiring, generating a positive relation between $\Delta emp_t$ and cross-sectional skewness.

The second variable, $\Delta y_t$, is the quarterly change in real compensation to private sector employees, which is essentially the first moment of the cross-sectional distribution of income growth rates. If changes in the first moment are driven by changes in the tails, one would expect to see a positive relation between $\Delta y_t$ and cross-sectional skewness. All nominal variables are converted to real variables using the personal consumption expenditures (PCE) deflator.

The third variable, $pw_{t-1}$, is the lagged ratio of corporate profits to wages, detrended using a HP filter. This variable captures a potential timing mismatch between shocks received by firms and those received by workers. Relative to profits, the response of wages to aggregate shocks is more sluggish, generating cyclical variation in overall profitability. If profits and wages are cointegrated, $pw_{t-1}$ can be interpreted as an error-correction term. Thus, when profits are high relative to wages, it is likely that firm recently experienced a series of favorable shocks. Future wages are likely to be higher and firms are more likely to be hiring than firing, causing the right tail of the income growth distribution to expand and the left tail to contract.

Additional motivation for $\Delta y_t$ and $pw_{t-1}$ comes from Berk et al. (2010). They derive the optimal contract between a risk averse worker and a risk-neutral firm when the productivity of the match varies over time, extending Harris and Holmstrom (1982) to a setting where firms have a financial incentive to issue debt. Under the optimal contract, firms partially insure workers against productivity shocks. In normal times, wages rise less than 1 for 1 in response to positive shocks and stay constant in response to negative shocks. This insurance breaks down when

---

62 I filter the series to eliminate very low frequency movements in this ratio, which could be related to changes in the composition of the private sector relative to the economy as a whole over time. As such, I use a smoothing parameter of 12,800, 8 times higher than the standard quarterly choice of 1600. Similar results obtain if the series is detrended by using a 10-year backward-looking moving average.
firms encounter financial distress, dissolving completely if the firm goes bankrupt. Workers
whose contracts are terminated experience sudden, large declines in wages—i.e. idiosyncratic
“disaster risk” arises as an equilibrium outcome of the model.63

Berk et al. (2010) is a partial equilibrium model, lacking any sources of aggregate risk. However,
if one takes the structure of their optimal contract as given and applies it to a world where
aggregate productivity is time-varying, the implications for the cross-sectional distribution of
income growth rates are relatively clear. When productivity increases, firms’ average profit
margins increase, making it easier for firms to insure workers against bad shocks in the future.
When profitability increases, average wages also increase. Conversely, firms have lower risk-
bearing capacity when overall profitability is low and wages are falling, increasing the risk that
workers experience large negative shocks and making the income growth rate distribution more
negatively skewed.64

The last variable, ∆ct, is the change in the logarithm of real aggregate consumption (nondurables
plus services). The intuition for aggregate consumption is essentially identical to that for ∆yt.
One would expect the distribution of income growth rates to be more negatively skewed when
household consumption is falling. However, compared with ∆yt, there is more scope for a timing
mismatch between ∆ct and cross-sectional skewness. For example, households with a strong
precautionary savings motive could cut consumption today in response to bad news about the
distribution of future labor income growth, causing ∆ct to lead the cross-sectional moments.

Results

As a precursor to my regressions, Figure 8 summarizes the univariate forecasting performance
of the employment and compensation growth, perhaps two of the most natural candidates for
zt. It plots the time series of 1-year and 5-year third central moments from GOS, and weighted
moving averages of these first two variables, φk(L; 0)Δemp and φk(L; 0)Δyt, respectively. For
purposes of generating these graphs, I calculate the moving averages assuming that ρ(L) = 0,
which assumes that transitory shocks are completely state-independent. As we discuss in greater
detail below, changing ρ(·) primarily impacts the level of φk(L; 0)zt rather than its time series
variation. Similar results obtain with other choices of ρ(·).

At both horizons, the time-aggregated employment and income growth measures track the cross-
sectional moments quite closely. The latter works slightly better at the 5-year horizon, while

63Berk et al. (2010) write: “employees’ wages at the moment of termination will typically be substantially
greater than their competitive market wages. As a result, these entrenched employees face substantial costs
resulting from a bankruptcy filing.”
64Giving workers occasional opportunities to switch firms, as in on-the-job search models, could potentially
generate procyclical variation in the likelihood of experiencing large positive shocks as well.
Figure 8: Co-movement of aggregate variables with third central moment of idiosyncratic income growth rates

Panel A plots the co-movement of 5-year idiosyncratic third central moments from GOS with weighted moving averages of logarithmic employment growth and real compensation growth. Panel B repeats the analysis for a 1 year measures. Series are standardized to have mean zero and unit variance. The weights are the lag polynomials $\phi_{20}(L; 0)$ and $\phi_{4}(L; 0)$ for 5 year and 1 year changes, respectively, which are defined in equation (27).

both measures perform equally well at the 1-year horizon. The $R^2$’s from univariate regressions of 5-year moments on employment and income growth are 61% and 72%, respectively. For 1-year measures, these $R^2$’s are 68% and 67%, respectively. At these frequencies, the two moving averages are fairly highly correlated. This is perhaps unsurprising, because changes in the size of the workforce likely generate the lion’s share of variation in aggregate wages. In the data, the asymmetry of the idiosyncratic labor income growth distribution is tightly linked with the extensive margin.

Table 11 estimates the vector $b$ in (4) by regressing the time-aggregated skewness measures on several aggregate variables. Given the sample size, I limit attention to univariate and bivariate

---

65 See Table 11, Panel A.
specifications. Panel A sets $\rho(L) = 0$, while Panel B allows for a restricted MA(1) structure: $\rho(L) = \rho \cdot [1 + L]$.\footnote{Similar results obtain with different lag lengths.} In this latter specification, the partial derivative of $y_t$ with respect to $\eta_{it}$ on is $[1 + \rho]$ in quarters $t$ and $t + 1$, and 1 in later periods; thus, the temporary effect reinforces (dampens if $\rho < 0$) the permanent effect by an additional $\rho\%$. All estimates are obtained by minimizing the sum of squared residuals, which is an OLS regression when $\rho$ is held fixed, and nonlinear least squares otherwise.\footnote{I calculate standard which are robust to the presence of heteroskedasticity and autocorrelation. I use a Newey-West estimator for the long-run variance with 4 lags.}

Each panel includes the coefficients from three different estimations. In the left columns, I report coefficients from pooled GMM regressions which include both 1 and 5-year third central moments as dependent variables. Next, we reestimate the model using data from each horizon separately. The center columns use 5-year measures only, while the right columns use 1-year measures only. When estimating these univariate regressions in Panel B, I fix the value of $\rho$ at its estimated value from the bivariate model.\footnote{Thus, the associated standard errors are best interpreted as conditional on $\hat{\rho}$.}

Qualitatively, the picture is essentially the same across specifications. In models 1-4, each of the four proxies always has the expected (positive) sign and is highly statistically significant. When I allow $\rho \neq 0$ in Panel B, our estimates are generally positive, suggesting that permanent shocks have additional transitory effects. In the bivariate models 5 and 6, both variables always enter positively and are generally statistically significant. A combination of contemporaneous income or employment growth with a proxy for future labor market conditions $pw_{t-1}$ matches the skewness measures quite well. Our estimates in Model 7, where $\Delta y_t$ and $\Delta c_t$ are both generally significant but enter with opposite signs, are somewhat less intuitive. However, the time series of quarterly skewness measures from this model track those from the other, more intuitive models relatively closely.

Panel A of Figure 9 plots our estimates of quarterly conditional third central moments, $\hat{Y}_t z_t$, from the pooled GMM estimates of models 5-7 from Panel B of Table 11. The picture from the corresponding models in Panel A are essentially identical. Model 5, which includes employment growth and the profit-wage ratio, appears to capture a common, low-frequency component around which the more volatile estimates from Models 6 and 7 fluctuate. While all three time series are highly cyclical, peaking in expansions and bottoming out in recessions, these idiosyncratic risk measures exhibit substantial time series variation, even in periods without recessions. Moreover, a quarterly NBER recession indicator has almost no explanatory power.

With the exception of $\Delta c_t$, all of my proxies are capable of capturing the variation in the 5-year measures quite well. Models 1-3 and 5-7 generate $R^2$’s in excess of 60% at a 5-year horizon. The
Table 11: Regressions of third central moment of income growth on aggregate variables

This table presents the results from estimating equation (4) for different choices of $z_t$ by least squares. The dependent variable is the time series of third central moments from the cross section of income growth rates from GOS. Panel A restricts $\rho(L) = 0$, while Panel B estimates $\rho(L) = \rho \cdot [1 + L]$. The “pooled GMM” column combines information from 1 and 5 year moments, while the next two columns reestimate the models using data on 5 year and 1 year measures only, conditioning on $\hat{\rho}$ from the pooled specification. Newey-West standard errors, calculated with 4 lags, are in parentheses.

### Panel A: Specifications with $\rho(L) = 0$

<table>
<thead>
<tr>
<th>Model</th>
<th>Variable</th>
<th>Pooled GMM</th>
<th>5 year only</th>
<th>1 year only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\hat{\rho}$</td>
<td>$R^2$, $R^2_{2}$</td>
<td>$\hat{\rho}$</td>
</tr>
<tr>
<td>1</td>
<td>$\Delta emp_t$</td>
<td>1.7959***</td>
<td>0.340</td>
<td>1.5640***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.415)</td>
<td>(0.424)</td>
<td>(0.865)</td>
</tr>
<tr>
<td>2</td>
<td>$\Delta y_t$</td>
<td>1.1325***</td>
<td>0.708</td>
<td>0.9944***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.143)</td>
<td>(0.152)</td>
<td>(0.434)</td>
</tr>
<tr>
<td>3</td>
<td>$pw_{t-1}$</td>
<td>0.0438***</td>
<td>0.610</td>
<td>0.0402***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.006)</td>
<td>(0.007)</td>
<td>(0.865)</td>
</tr>
<tr>
<td>4</td>
<td>$\Delta c_t$</td>
<td>1.5459***</td>
<td>0.233</td>
<td>1.3163***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.383)</td>
<td>(0.388)</td>
<td>(1.691)</td>
</tr>
<tr>
<td>5</td>
<td>$\Delta emp_t$</td>
<td>1.2245***</td>
<td>0.766</td>
<td>0.9908***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.236)</td>
<td>(0.214)</td>
<td>(0.991)</td>
</tr>
<tr>
<td>6</td>
<td>$\Delta y_t$</td>
<td>0.8456***</td>
<td>0.808</td>
<td>0.6895***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.129)</td>
<td>(0.093)</td>
<td>(0.666)</td>
</tr>
<tr>
<td>7</td>
<td>$pw_{t-1}$</td>
<td>0.0202***</td>
<td>0.391</td>
<td>0.210***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.007)</td>
<td>(0.016)</td>
<td>(0.196)</td>
</tr>
</tbody>
</table>

### Panel B: Specifications with $\rho(L) = \rho \cdot [1 + L]$

<table>
<thead>
<tr>
<th>Model</th>
<th>Variable</th>
<th>Pooled GMM</th>
<th>5 year only</th>
<th>1 year only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\hat{\rho}$</td>
<td>$R^2$, $R^2_{2}$</td>
<td>$\hat{\rho}$</td>
</tr>
<tr>
<td>1</td>
<td>$\Delta emp_t$</td>
<td>1.5676***</td>
<td>0.5763</td>
<td>1.3323***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.473)</td>
<td>(0.371)</td>
<td>(0.455)</td>
</tr>
<tr>
<td>2</td>
<td>$\Delta y_t$</td>
<td>0.9565***</td>
<td>0.6002**</td>
<td>0.8390***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.160)</td>
<td>(0.237)</td>
<td>(0.196)</td>
</tr>
<tr>
<td>3</td>
<td>$pw_{t-1}$</td>
<td>0.0433***</td>
<td>-1.3590***</td>
<td>0.0393***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.007)</td>
<td>(0.285)</td>
<td>(0.016)</td>
</tr>
<tr>
<td>4</td>
<td>$\Delta c_t$</td>
<td>1.2088***</td>
<td>0.9957*</td>
<td>1.0426***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.376)</td>
<td>(0.448)</td>
<td>(0.101)</td>
</tr>
<tr>
<td>5</td>
<td>$\Delta emp_t$</td>
<td>1.1851***</td>
<td>0.4630</td>
<td>0.9456***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.261)</td>
<td>(0.313)</td>
<td>(0.167)</td>
</tr>
<tr>
<td>6</td>
<td>$\Delta y_t$</td>
<td>0.9522***</td>
<td>0.5240**</td>
<td>0.6568***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.125)</td>
<td>(0.243)</td>
<td>(0.326)</td>
</tr>
<tr>
<td>7</td>
<td>$pw_{t-1}$</td>
<td>0.0154*</td>
<td>0.0158***</td>
<td>0.0224*</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.008)</td>
<td>(0.005)</td>
<td>(0.012)</td>
</tr>
</tbody>
</table>

**Panel A** restricts $\rho(L) = 0$, while **Panel B** estimates $\rho(L) = \rho \cdot [1 + L]$. The “pooled GMM” column combines information from 1 and 5 year moments, while the next two columns reestimate the models using data on 5 year and 1 year measures only, conditioning on $\hat{\rho}$ from the pooled specification. Newey-West standard errors, calculated with 4 lags, are in parentheses.
Panel A: Quarterly third central moment of permanent shocks – all models, demeaned

<table>
<thead>
<tr>
<th>Year</th>
<th>Fitted third central moment, demeaned</th>
</tr>
</thead>
<tbody>
<tr>
<td>1960</td>
<td>-0.06</td>
</tr>
<tr>
<td>1970</td>
<td>-0.04</td>
</tr>
<tr>
<td>1980</td>
<td>-0.02</td>
</tr>
<tr>
<td>1990</td>
<td>0</td>
</tr>
<tr>
<td>2000</td>
<td>0.02</td>
</tr>
<tr>
<td>2010</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Panel B: 1 year lag polynomials

\[ \Phi_4(L; \rho) = \begin{cases} 
0 & \text{if } \rho = 0 \\
0.463 & \text{if } \rho = 0.463 \\
0 & \text{if } \rho = 0, \text{scaled} 
\end{cases} \]

Panel C: 5 year lag polynomials

\[ \Phi_4(L; \rho) = \begin{cases} 
0 & \text{if } \rho = 0 \\
\hat{\rho}(1 + L) & \text{if } \rho \text{ estimated} \\
0 & \text{if } \rho = 0, \text{scaled} 
\end{cases} \]

**Figure 9:** Key features from estimated regression models

Panel A plots pooled GMM estimates of quarterly conditional third central moments from the estimated specifications in Panel B of Table 11, i.e. \( \hat{b}'z_t \). A dashed vertical line indicates the beginning of the sample period used to estimate the skewness measures. Panel B plots the coefficients of the lag polynomial \( \Phi_4(L; \rho) \) in the regression equation (4) for 1 year skewness measures. The first series imposes \( \rho(L) = 0 \), while the second corresponds with the estimated \( \rho(L) = \hat{\rho}(1 + L) \) from Model 5. The third line rescales the first line so that the sum of the weights is the same as that from the estimated specification. Panel C repeats the analysis in Panel B for 5 year measures.

The inferior performance of Model 4 is somewhat unsurprising in light of my discussion above about a potential timing mismatch between \( \Delta c_t \) and idiosyncratic labor market shocks. Moreover, the 5-year \( R^2 \)'s are similar between the pooled GMM and univariate specifications in the left and middle columns, respectively. At a 1-year horizon, \( R^2 \)'s are also in excess of 60% for Models 1-2 and 5-7 in the univariate specifications in the left and right columns. However, the differences between the pooled GMM and univariate specifications are larger. In the pooled GMM specifications, these \( R^2 \)'s are between 10 and 20 percentage points lower in Panel B and substantially lower in Panel A.
If the data are generated according to equation (5), the slope coefficients from a regression of $M_3[y_{A,t} - y_{A,t-k}]$ on $\phi_k(L; \rho)$ is the same for all horizons $k$. My pooled GMM estimations impose this restriction. Therefore, I can check the validity of this assumption by comparing the estimated coefficients and $R^2$’s in the middle and right columns with those from the pooled estimation. Comparing the the middle and right columns, the coefficients estimated using the 1-year skewness measures only are generally much larger in magnitude than the pooled estimates. These differences are particularly stark in Panel A, where unrestricted 1-year specifications outperform the pooled GMM estimates by a wide margin.

These discrepancies, though still present, shrink substantially once I allow $\rho \neq 0$ (Panel B). When comparing the $R^2$’s from the pooled GMM specifications in Panels A and B, the first order effect of allowing $\rho \neq 0$ is an improvement in the fit for 1-year skewness measures. Panels B and C of Figure 9 offer an explanation for such a result. Panel B plots the coefficients of the lag polynomial $\phi_4(L; \rho)$ in the regression equation (4) for 1 year skewness measures. First, I plot the weights when $\rho(L) = 0$. Second, I plot the weights implied by $\rho(L) = \hat{\rho}(1 + L)$ from model 5. The third line rescales the first so that the sum of the weights matches the second, making it easier to compare the shapes of the two fitted polynomials.

At a 1-year horizon, both lag polynomials are tent-shaped, giving the highest weight to the third lag (the shock received in first quarter of the end year). The biggest difference is that the model with $\rho > 0$ has a much higher peak. The sum of weights is about 75% larger relative to the specification with $\rho = 0$. Second, the weighting function with $\rho > 0$ is asymmetric, overweighting more recent lags. Relative to the change in the sum of the weights, the change in the shape induced by $\rho \neq 0$—the difference between the solid gray and dashed blue lines—is less substantial. As such, the primary effect of allowing $\rho > 0$ is to increase the variance of $\phi_k(L; \rho)z_t$ by a factor of around 3, shrinking the 1-year regression coefficients towards zero.

Figure 9, Panel C shows the corresponding weights in the lag polynomials for 5 year measures. Once again, the specification with $\rho > 0$ puts higher weight on recent lags. However, as I am summing over a much larger number of lags, the overall effect is quite minor. The sum of the weights is much less sensitive to changes in $\rho$, and the weighting functions have essentially identical shapes after the 5th lag. Accordingly, changes in $\rho$ will have much larger effects on the 1-year moving averages relative to the 5-year moving averages.

Looking at Panel B of Table 11, there remains room for improvement. The coefficients from the unrestricted 1-year models in the right columns are still larger and the $R^2$’s are somewhat lower than their counterparts from the pooled GMM specifications. There appear to be additional dimensions of transitory risk which are not captured by my relatively simplistic model for $\rho(L)$. For example, models 5-6 in the right column tend to place higher weights on $\Delta y_t$ and $\Delta emp_t$. 

A.13
relative to the other columns, so contemporaneous labor market factors might have a closer connection with transitory risk than the forward-looking pw_{t-1}. The dynamics of the quarterly skewness measures are relatively insensitive to my specification of ρ(L). However, I am more concerned with the model’s performance at longer frequencies, which is quite strong.

In additional unreported results, I replicate the volatility tests from Table 2 for each of the conditional skewness indices. For most of the models, all three of the measures from section 2.4 proxy for uncertainty about future idiosyncratic risk. For some of these alternative specifications, I find that the cross-sectional volatility measure works particularly well.

A.2.3 Implementation Details - 3PRF Approach

Variables used to calculate the skewness index

Table 12 lists the variables which I use to construct my skewness index. Following Wu and Xia (2014), 97 of the variables are obtained from Global Insight, which is the subset of 120 series from Bernanke et al. (2005) which are available through the present. I augment these time series with 12 variables from the literature on return predictability, which are taken from Ivo Welch. The second column provides the Global Insight or Goyal and Welch (2008) mnemonic for each series, and the final column indicates which transformation, if any, I perform to the raw time series data.

Methodology

Here, I provide a number of additional details about the estimation procedure which generates my quarterly skewness index. Equation (4) shows that I can write a panel of time-aggregated cross-sectional moments (at different horizons k, measured at different times t) in “regression” form. I estimate b so as to minimize the sum of squared residuals in (4).

Equation (4) is linear in parameters conditional on the moving average coefficient, ρ. Further, the slope coefficients are identical across horizons; only the intercepts are horizon-specific. Therefore, for low-dimensional y_t, an OLS regression of the cross-sectional moments on φ_k(L, ρ)y_t and horizon-specific constants will minimize the sum of squared residuals for each value of ρ. Then, I choose the value of ρ which minimizes the total sum of squared residuals. Very little changes when I move from a low-dimensional to a high-dimensional y_t. However, rather than run an OLS regression, I use the 3PRF to estimate b for each value of ρ.69

69 I use the version of the 3PRF with a single factor and automatic proxy selection. See Kelly and Pruitt (2014) for further details. I am grateful to Bryan Kelly and Seth Pruitt for making available their code for the 3PRF.
<table>
<thead>
<tr>
<th>Mnemonic</th>
<th>Description</th>
<th>Transform</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPS11.M</td>
<td>INDUSTRIAL PRODUCTION INDEX - PRODUCTS, TOTAL</td>
<td>Δ log</td>
</tr>
<tr>
<td>IPS299.M</td>
<td>INDUSTRIAL PRODUCTION INDEX - FINAL PRODUCTS</td>
<td>Δ log</td>
</tr>
<tr>
<td>IPS12.M</td>
<td>INDUSTRIAL PRODUCTION INDEX - CONSUMER GOODS</td>
<td>Δ log</td>
</tr>
<tr>
<td>IPS13.M</td>
<td>INDUSTRIAL PRODUCTION INDEX - DURABLE CONSUMER GOODS</td>
<td>Δ log</td>
</tr>
<tr>
<td>IPS18.M</td>
<td>INDUSTRIAL PRODUCTION INDEX - NONDURABLE CONSUMER GOODS</td>
<td>Δ log</td>
</tr>
<tr>
<td>IPS25.M</td>
<td>INDUSTRIAL PRODUCTION INDEX - BUSINESS EQUIPMENT</td>
<td>Δ log</td>
</tr>
<tr>
<td>IPS32.M</td>
<td>INDUSTRIAL PRODUCTION INDEX - MATERIALS</td>
<td>Δ log</td>
</tr>
<tr>
<td>IPS34.M</td>
<td>INDUSTRIAL PRODUCTION INDEX - DURABLE GOODS MATERIALS</td>
<td>Δ log</td>
</tr>
<tr>
<td>IPS38.M</td>
<td>INDUSTRIAL PRODUCTION INDEX - NONDURABLE GOODS MATERIALS</td>
<td>Δ log</td>
</tr>
<tr>
<td>IPS43.M</td>
<td>INDUSTRIAL PRODUCTION INDEX - MANUFACTURING (SIC)</td>
<td>Δ log</td>
</tr>
<tr>
<td>IPS311.M</td>
<td>INDUSTRIAL PRODUCTION INDEX - OIL &amp; GAS WELL DRILLING &amp; MANUFACTURED HOMES</td>
<td>Δ log</td>
</tr>
<tr>
<td>IPS307.M</td>
<td>INDUSTRIAL PRODUCTION INDEX - RESIDENTIAL UTILITIES</td>
<td>Δ log</td>
</tr>
<tr>
<td>IPS10.M</td>
<td>INDUSTRIAL PRODUCTION INDEX - TOTAL INDEX</td>
<td>Δ log</td>
</tr>
<tr>
<td>UTL11.M</td>
<td>CAPACITY UTILIZATION - MANUFACTURING (SIC)</td>
<td>Δ log</td>
</tr>
<tr>
<td>PI001.M</td>
<td>PERSONAL INCOME, BIL$, SA</td>
<td>Δ log</td>
</tr>
<tr>
<td>A6M051.M</td>
<td>PERS INCOME LESS TRSF PMT (AR BIL, CHAIN 2009 $), SA-US</td>
<td>Δ log</td>
</tr>
<tr>
<td>LHEM.M</td>
<td>CIVILIAN LABOR FORCE: EMPLOYED, TOTAL (THOUS.,SA)</td>
<td>Δ log</td>
</tr>
<tr>
<td>LHNAG.M</td>
<td>CIVILIAN LABOR FORCE: EMPLOYED, NONAGRIC.INDUSTRIES (THOUS.,SA)</td>
<td>Δ log</td>
</tr>
<tr>
<td>LHUR.M</td>
<td>UNEMPLOYMENT RATE: ALL WORKERS, 16 YEARS and OVER (%.SA)</td>
<td>Δ log</td>
</tr>
<tr>
<td>LHUR680.M</td>
<td>UNEMPLOY.BY DURATION: AVERAGE(MEAN)DURATION IN WEEKS (SA)</td>
<td>Δ log</td>
</tr>
<tr>
<td>LHU5.M</td>
<td>UNEMPLOY.BY DURATION: PERSONS UNEMPL.LESS THAN 5 WKS (THOUS.,SA)</td>
<td>Δ log</td>
</tr>
<tr>
<td>LHU14.M</td>
<td>UNEMPLOY.BY DURATION: PERSONS UNEMPL.5 TO 14 WKS (THOUS.,SA)</td>
<td>Δ log</td>
</tr>
<tr>
<td>LHU15.M</td>
<td>UNEMPLOY.BY DURATION: PERSONS UNEMPL.15 WKS + (THOUS.,SA)</td>
<td>Δ log</td>
</tr>
<tr>
<td>CES001.M</td>
<td>EMPLOYEES, NONFARM - TOTAL NONFARM</td>
<td>Δ log</td>
</tr>
<tr>
<td>CES002.M</td>
<td>EMPLOYEES, NONFARM - TOTAL PRIVATE</td>
<td>Δ log</td>
</tr>
<tr>
<td>CES003.M</td>
<td>EMPLOYEES, NONFARM - GOODS-PRODUCING</td>
<td>Δ log</td>
</tr>
<tr>
<td>CES006.M</td>
<td>EMPLOYEES, NONFARM - MINING</td>
<td>Δ log</td>
</tr>
<tr>
<td>CES011.M</td>
<td>EMPLOYEES, NONFARM - CONSTRUCTION</td>
<td>Δ log</td>
</tr>
<tr>
<td>CES015.M</td>
<td>EMPLOYEES, NONFARM - MFG</td>
<td>Δ log</td>
</tr>
<tr>
<td>CES017.M</td>
<td>EMPLOYEES, NONFARM - DURABLE GOODS</td>
<td>Δ log</td>
</tr>
<tr>
<td>CES033.M</td>
<td>EMPLOYEES, NONFARM - NONDURABLE GOODS</td>
<td>Δ log</td>
</tr>
<tr>
<td>CES046.M</td>
<td>EMPLOYEES, NONFARM - SERVICE-PROVIDING</td>
<td>Δ log</td>
</tr>
<tr>
<td>CES048.M</td>
<td>EMPLOYEES, NONFARM - TRANSPORT, TRANSPORT, UTILITIES</td>
<td>Δ log</td>
</tr>
<tr>
<td>CES049.M</td>
<td>EMPLOYEES, NONFARM - WHOLESALE TRADE</td>
<td>Δ log</td>
</tr>
<tr>
<td>CES053.M</td>
<td>EMPLOYEES, NONFARM - RETAIL TRADE</td>
<td>Δ log</td>
</tr>
<tr>
<td>CES140.M</td>
<td>EMPLOYEES, NONFARM - GOVERNMENT</td>
<td>Δ log</td>
</tr>
<tr>
<td>CES154.M</td>
<td>AVG WKLY HOURS, PROD WRKRS, NONFARM - MFG</td>
<td>Δ log</td>
</tr>
<tr>
<td>CES155.M</td>
<td>AVG WKLY OVERTIME HOURS, PROD WRKRS, NONFARM - MFG</td>
<td>Δ log</td>
</tr>
<tr>
<td>PMEMP.M</td>
<td>NAPM EMPLOYMENT INDEX (PERCENT)</td>
<td>Δ log</td>
</tr>
</tbody>
</table>

Table 12: List of Variables Included in Calculation of Skewness Index (cont.)
<table>
<thead>
<tr>
<th>Mneumonic</th>
<th>Description</th>
<th>Transform</th>
</tr>
</thead>
<tbody>
<tr>
<td>PI031.M</td>
<td>PERSONAL CONSUMPTION EXPENDITURES, BIL$, SAAR</td>
<td>$\Delta \log$</td>
</tr>
<tr>
<td>PI032.M</td>
<td>PERSONAL CONSUMPTION EXPENDITURES - DURABLE GOODS, BIL$, SAAR</td>
<td>$\Delta \log$</td>
</tr>
<tr>
<td>PI033.M</td>
<td>PERSONAL CONSUMPTION EXPENDITURES - NONDURABLE GOODS, BIL$, SAAR</td>
<td>$\Delta \log$</td>
</tr>
<tr>
<td>PI034.M</td>
<td>PERSONAL CONSUMPTION EXPENDITURES - SERVICES, BIL$, SAAR</td>
<td>$\Delta \log$</td>
</tr>
<tr>
<td>HSFR.M</td>
<td>HOUSING STARTS: TOTAL FARM&amp;NONFARM (THOUS., SA)</td>
<td>log</td>
</tr>
<tr>
<td>HSNE.M</td>
<td>HOUSING STARTS: NORTHEAST (THOUS., U.S.A.)</td>
<td>log</td>
</tr>
<tr>
<td>HSMW.M</td>
<td>HOUSING STARTS: MIDWEST (THOUS., U.S.A.)</td>
<td>log</td>
</tr>
<tr>
<td>HSSOU.M</td>
<td>HOUSING STARTS: SOUTH (THOUS., U.S.A.)</td>
<td>log</td>
</tr>
<tr>
<td>HSWST.M</td>
<td>HOUSING STARTS: WEST (THOUS., U.S.A.)</td>
<td>log</td>
</tr>
<tr>
<td>HS6BR.M</td>
<td>HOUSING AUTHORIZED: TOTAL NEW PRIV HOUSING UNITS (THOUS., NSA)</td>
<td>log</td>
</tr>
<tr>
<td>HMOB.M</td>
<td>MOBILE HOMES: MANUFACTURERS’ SHIPMENTS (THOUS., OF UNITS, SAAR)</td>
<td>log</td>
</tr>
<tr>
<td>PMNV.M</td>
<td>NAPM INVENTORIES INDEX (PERCENT)</td>
<td></td>
</tr>
<tr>
<td>PMNO.M</td>
<td>NAPM NEW ORDERS INDEX (PERCENT)</td>
<td></td>
</tr>
<tr>
<td>PMDEL.M</td>
<td>NAPM VENDOR DELIVERIES INDEX (PERCENT)</td>
<td></td>
</tr>
<tr>
<td>MOCMQ.M</td>
<td>NEW ORDERS (NET) - CONSUMER GOODS and MATERIALS, 1996 $ (BCI)</td>
<td>$\Delta \log$</td>
</tr>
<tr>
<td>MSONDQ.M</td>
<td>NEW ORDERS, NONDEFENSE CAPITAL GOODS, IN 1996 $ (BCI)</td>
<td>$\Delta \log$</td>
</tr>
<tr>
<td>FSPCOM.M</td>
<td>S&amp;P’S COMMON STOCK PRICE INDEX: COMPOSITE (1941-43=10)</td>
<td>$\Delta \log$</td>
</tr>
<tr>
<td>FSPIN.M</td>
<td>S&amp;P’S COMMON STOCK PRICE INDEX: INDUSTRIALS (1941-43=10)</td>
<td>$\Delta \log$</td>
</tr>
<tr>
<td>DY (GW)</td>
<td>DIVIDEND YIELD</td>
<td></td>
</tr>
<tr>
<td>EP (GW)</td>
<td>EARNINGS-PRICE RATIO</td>
<td></td>
</tr>
<tr>
<td>BM (GW)</td>
<td>BOOK-TO-MARKET RATIO</td>
<td></td>
</tr>
<tr>
<td>SVAR (GW)</td>
<td>STOCK MARKET REALIZED VARIANCE</td>
<td></td>
</tr>
<tr>
<td>TBL (GW)</td>
<td>3 MONTH T-BILL RATE</td>
<td></td>
</tr>
<tr>
<td>TMS (GW)</td>
<td>TERM SPREAD</td>
<td></td>
</tr>
<tr>
<td>DFY (GW)</td>
<td>DEFAULT YIELD: BAA - AAA SPREAD</td>
<td></td>
</tr>
<tr>
<td>LTY (GW)</td>
<td>LONG-TERM YIELD</td>
<td></td>
</tr>
<tr>
<td>NTIS (GW)</td>
<td>NET ISSUANCE</td>
<td></td>
</tr>
<tr>
<td>INFL (GW)</td>
<td>INFLATION</td>
<td></td>
</tr>
<tr>
<td>DFR (GW)</td>
<td>CORPORATE - GOVT BOND RETURN</td>
<td></td>
</tr>
<tr>
<td>LTR (GW)</td>
<td>LONG-TERM BOND RETURN</td>
<td></td>
</tr>
<tr>
<td>EXRUK.M</td>
<td>FOREIGN EXCHANGE RATE: UNITED KINGDOM (CENTS PER POUND)</td>
<td>$\Delta \log$</td>
</tr>
<tr>
<td>EXRCAN.M</td>
<td>FOREIGN EXCHANGE RATE: CANADA (CANADIAN $ PER U.S.$)</td>
<td>$\Delta \log$</td>
</tr>
</tbody>
</table>

**Table 12:** List of Variables Included in Calculation of Skewness Index (cont.)
### INTEREST RATES

<table>
<thead>
<tr>
<th>Mneumonic</th>
<th>Description</th>
<th>Transform</th>
</tr>
</thead>
<tbody>
<tr>
<td>75 FYFF.M</td>
<td>INTEREST RATE: FEDERAL FUNDS (EFFECTIVE) (% PER ANNUM, NSA)</td>
<td></td>
</tr>
<tr>
<td>76 FYGM3.M</td>
<td>INTEREST RATE: U.S. TREASURY BILLS, SEC, MKT, 3-MO. (% PER ANNUM, NSA)</td>
<td></td>
</tr>
<tr>
<td>77 FYGM6.M</td>
<td>INTEREST RATE: U.S. TREASURY BILLS, SEC, MKT, 6-MO. (% PER ANNUM, NSA)</td>
<td></td>
</tr>
<tr>
<td>78 FYGT1.M</td>
<td>INTEREST RATE: U.S. TREASURY CONST MATURITIES, 1-YR. (% PER ANNUM, NSA)</td>
<td></td>
</tr>
<tr>
<td>79 FYGT5.M</td>
<td>INTEREST RATE: U.S. TREASURY CONST MATURITIES, 5-YR. (% PER ANNUM, NSA)</td>
<td></td>
</tr>
<tr>
<td>80 FYGT10.M</td>
<td>INTEREST RATE: U.S. TREASURY CONST MATURITIES, 10-YR. (% PER ANNUM, NSA)</td>
<td></td>
</tr>
<tr>
<td>83 FYGT1.M-FYFF.M</td>
<td>SPREAD: FYGT1.M-FYFF.M</td>
<td></td>
</tr>
<tr>
<td>84 FYGT5.M-FYFF.M</td>
<td>SPREAD: FYGT5.M-FYFF.M</td>
<td></td>
</tr>
</tbody>
</table>

### MONEY AND CREDIT QUANTITY AGGREGATES

<table>
<thead>
<tr>
<th>Mneumonic</th>
<th>Description</th>
<th>Transform</th>
</tr>
</thead>
<tbody>
<tr>
<td>86 ALCIBL00.M</td>
<td>COML&amp;IND LOANS OUTST IN 2009 $, SA-U</td>
<td>$log</td>
</tr>
<tr>
<td>87 CCINRV.M</td>
<td>CONSUMER CREDIT OUTSTANDING - NONREVOLVING (G19)</td>
<td>$log</td>
</tr>
<tr>
<td>88 FM1.M</td>
<td>MONEY STOCK: M1 (CURR, TRAV. CKS, DEM DEP, OTHER CKABLE DEP) (BIL$, SA)</td>
<td>$log</td>
</tr>
<tr>
<td>89 FM2.M</td>
<td>MONEY STOCK: M2 (M1 + O'NITE RPS, EURO$, G/P&amp;B/D MMMFS&amp;SAV &amp; SM TIME DEP (BIL$, SA)</td>
<td>$log</td>
</tr>
<tr>
<td>90 MBASE.M</td>
<td>REVISED MONETARY BASE-ADJUSTED-(FED RESERVE BANK-SAINT LOUIS), SA-US</td>
<td>$log</td>
</tr>
<tr>
<td>91 MNY2.M</td>
<td>M2 - MONEY SUPPLY - M1 + SAVINGS DEPOSITS, SMALL TIME DEPOSITS, &amp; MMMFS [H6], SA-US</td>
<td>$log</td>
</tr>
</tbody>
</table>

### PRICE INDICES

<table>
<thead>
<tr>
<th>Mneumonic</th>
<th>Description</th>
<th>Transform</th>
</tr>
</thead>
<tbody>
<tr>
<td>92 PMCP.M</td>
<td>APM COMMODITY PRICES INDEX (PERCENT)</td>
<td>$log</td>
</tr>
<tr>
<td>93 PWFS.A.M</td>
<td>PRODUCER PRICE INDEX; FINISHED GOODS (82=100, SA)</td>
<td>$log</td>
</tr>
<tr>
<td>94 PWFCSA.M</td>
<td>PRODUCER PRICE INDEX; FINISHED CONSUMER GOODS (82=100, SA)</td>
<td>$log</td>
</tr>
<tr>
<td>95 PWIMS.A.M</td>
<td>PRODUCER PRICE INDEX; INTERMEDIATE MATERIALS &amp; COMPONENTS (82=100, SA)</td>
<td>$log</td>
</tr>
<tr>
<td>96 PWCMSA.M</td>
<td>PRODUCER PRICE INDEX; CRUDE MATERIALS (82=100, SA)</td>
<td>$log</td>
</tr>
<tr>
<td>97 PUNEW.M</td>
<td>CPI-U: ALL ITEMS (82-84=100, SA)</td>
<td>$log</td>
</tr>
<tr>
<td>98 PU83.M</td>
<td>CPI-U: APPAREL &amp; UPKEEP (82-84=100, SA)</td>
<td>$log</td>
</tr>
<tr>
<td>99 PU84.M</td>
<td>CPI-U: TRANSPORTATION (82-84=100, SA)</td>
<td>$log</td>
</tr>
<tr>
<td>100 PU85.M</td>
<td>CPI-U: MEDICAL CARE (82-84=100, SA)</td>
<td>$log</td>
</tr>
<tr>
<td>101 PUC.M</td>
<td>CPI-U: COMMODITIES (82-84=100, SA)</td>
<td>$log</td>
</tr>
<tr>
<td>102 PUCD.M</td>
<td>CPI-U: DURABLES (82-84=100, SA)</td>
<td>$log</td>
</tr>
<tr>
<td>103 PUS.M</td>
<td>CPI-U: SERVICES (82-84=100, SA)</td>
<td>$log</td>
</tr>
<tr>
<td>104 PUXF.M</td>
<td>CPI-U: ALL ITEMS LESS FOOD (82-84=100, SA)</td>
<td>$log</td>
</tr>
<tr>
<td>105 PUXHS.M</td>
<td>CPI-U: ALL ITEMS LESS SHELTER (82-84=100, SA)</td>
<td>$log</td>
</tr>
<tr>
<td>106 PUXM.M</td>
<td>CPI-U: ALL ITEMS LESS MEDICAL CARE (82-84=100, SA)</td>
<td>$log</td>
</tr>
</tbody>
</table>

### AVERAGE HOURLY EARNINGS

<table>
<thead>
<tr>
<th>Mneumonic</th>
<th>Description</th>
<th>Transform</th>
</tr>
</thead>
<tbody>
<tr>
<td>107 CES277.M</td>
<td>AVG HRLY EARNINGS, PROD WRKRS, NONFARM - CONSTRUCTION</td>
<td>$log</td>
</tr>
<tr>
<td>108 CES278.M</td>
<td>AVG HRLY EARNINGS, PROD WRKRS, NONFARM - MFG</td>
<td>$log</td>
</tr>
</tbody>
</table>

### CONSUMER EXPECTATIONS

<table>
<thead>
<tr>
<th>Mneumonic</th>
<th>Description</th>
<th>Transform</th>
</tr>
</thead>
<tbody>
<tr>
<td>109 U0M083.M</td>
<td>BUSINESS CYCLE INDICATORS, CONSUMER EXPECTATIONS, NS</td>
<td></td>
</tr>
</tbody>
</table>

Table 12: List of Variables Included in Calculation of Skewness Index (cont.)
The 3PRF method of Kelly and Pruitt (2014) is intended to solve a linear, univariate forecasting problem, whereas (4) includes horizon-specific constants. Given that the constants in (4) aren’t of particular interest, I “partial them out” before estimating the slope coefficients. This amounts to subtracting the time-series mean of the cross-sectional moments (the dependent variable) and \( \phi_k(L, \rho) y_t \) (the independent variables) for each of the horizons \( k \).\(^{70}\) The regression involving these transformed variables no longer has an intercept. Since none of the coefficients depend on \( k \), I can ignore the panel dimension entirely and treat my data as if I were running a simple univariate regression. I estimate \( b \) by applying the 3PRF to the transformed variables.

Finally, as mentioned in the main text, I allow for some flexibility with respect to lead-lag relationships between the observed macroeconomic time series and the cross-sectional moments. For example, a moving average of returns may outperform the contemporaneous stock return as a proxy for the conditional skewness of \( \eta_t \). I allow for this possibility in a fairly straightforward manner. For each of the 109 time series, I choose the moving average transformation (lag polynomial) which best fits the data–as measured by the sum of squared residuals in the univariate version of (4)–from a finite set of potential transformations. The coefficients on the lag polynomials for these potential transformations of the data are given in Table 13.\(^{71}\) The 3PRF then chooses the optimal linear combination of these transformed variables.

### A.3 Further details on parametric income process

#### A.3.1 Calibrated specification

I maintain my assumptions on the labor income process from section 2.2. For ease of notation, I will suppress \( i \) superscripts here. Equation (26), in Appendix A.2.1, shows that the growth rate of annual labor income at horizon \( k \) is the sum of three pieces: profile heterogeneity, \( (\beta \cdot k) \), a state-independent transitory shock \( (\epsilon_{A,t} - \epsilon_{A,t-k}) \), and a weighted moving average of permanent shocks \( (\theta_k(L; \rho) \eta_t) \). For the first two terms, I adopt the fairly standard assumptions that \( \beta \sim N(0, \sigma^2_b) \) and \( \epsilon_{A,t} \sim N(0, \sigma^2_{\epsilon}) \).

My primary interest is on the last term, a weighted moving average of permanent shocks, \( \eta_t \). I assume that \( \eta_t = (J_{g,t} - E_t[J_{g,t}]) + (J_{b,t} - E_t[J_{b,t}]) + N(0, \sigma^2_n) \). \( J_{g,t} \) and \( J_{b,t} \) are compound

---

\(^{70}\)Note that this adjustment will depend on \( \rho \).  
\(^{71}\)I allow this selection to vary depending on \( \rho \). The weights are constructed using the pdf of a \textit{Beta}(\( \alpha, \beta \)) random variable (a popular choice in the literature on MiDaS estimation), where the two parameters are chosen on a rectangular grid where \( \alpha \in \{0.1, 1, 2, 3, 4, 5\} \) and \( \beta \in \{1, 3, 5, 7, 9\} \). This produces a flexible variety of shapes, as is evident from inspection of the lab polynomials in Table 13.
Table 13: Potential moving average transformations of underlying time series performed prior to calculating skewness index

For each one of the 109 quarterly time series which is considered for inclusion in my skewness index, I allow for some flexibility with respect to lead-lag relationships between the observed macroeconomic time series and the cross-sectional moments. For each of the underlying 109 time series, I choose from a finite set of potential moving average transformations (lag polynomials) which best fits the data—as measured by the sum of squared residuals in the univariate version of (4). The table provides the coefficients on the lag polynomials for these potential transformations of the data. I consider transformations involving 4 quarters and 8 quarters of data, which are provided in the left and right columns, respectively. The selected transformation is permitted to vary depending on $\rho$. 

<table>
<thead>
<tr>
<th>Models w. 4 lags</th>
<th>Models w. 8 lags</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>100 0 0 0</td>
<td>0 1 2 3 4 5 6 7</td>
</tr>
<tr>
<td>100 100 0 0</td>
<td></td>
</tr>
<tr>
<td>0 100 0 0</td>
<td></td>
</tr>
<tr>
<td>100 100 100 0</td>
<td></td>
</tr>
<tr>
<td>0 0 100 0</td>
<td></td>
</tr>
<tr>
<td>0 0 0 100</td>
<td></td>
</tr>
<tr>
<td>95 5 0 0</td>
<td>81 15 3 1 0 0 0</td>
</tr>
<tr>
<td>91 9 1 0</td>
<td>75 18 5 1 0 0 0</td>
</tr>
<tr>
<td>84 14 2 0</td>
<td>67 21 8 3 1 0 0</td>
</tr>
<tr>
<td>71 21 7 1</td>
<td>55 22 11 6 3 2 1</td>
</tr>
<tr>
<td>46 24 17 13</td>
<td>34 18 13 10 8 7 6 5</td>
</tr>
<tr>
<td>91 9 0 0</td>
<td>68 23 7 2 0 0 0 0</td>
</tr>
<tr>
<td>84 15 1 0</td>
<td>59 26 10 3 1 0 0 0</td>
</tr>
<tr>
<td>72 23 5 0</td>
<td>47 27 15 7 3 1 0 0</td>
</tr>
<tr>
<td>53 30 13 3</td>
<td>31 24 18 12 8 4 2 0</td>
</tr>
<tr>
<td>25 25 25 25</td>
<td>13 13 13 13 13 13 13 13</td>
</tr>
<tr>
<td>70 28 2 0</td>
<td>27 37 24 10 3 0 0 0</td>
</tr>
<tr>
<td>54 38 8 0</td>
<td>17 31 27 16 7 2 0 0</td>
</tr>
<tr>
<td>35 44 19 2</td>
<td>9 21 26 22 14 6 2 0</td>
</tr>
<tr>
<td>15 35 35 15</td>
<td>3 10 16 20 20 16 10 3</td>
</tr>
<tr>
<td>3 13 30 53</td>
<td>0 2 4 8 12 18 24 31</td>
</tr>
<tr>
<td>34 55 11 0</td>
<td>4 23 34 25 10 2 0 0</td>
</tr>
<tr>
<td>19 55 24 1</td>
<td>2 14 28 29 19 7 1 0</td>
</tr>
<tr>
<td>8 42 42 8</td>
<td>1 6 17 26 26 17 6 1</td>
</tr>
<tr>
<td>2 19 44 35</td>
<td>0 2 6 14 22 26 21 9</td>
</tr>
<tr>
<td>0 5 23 72</td>
<td>0 0 1 3 7 1 5 27 47</td>
</tr>
</tbody>
</table>
Poisson random variables with time-varying intensities and exponential increments, defined as

\[
J_{g,t} = \sum_{j=1}^{N_{g,t}} [\mu_s + \text{Exponential}(\sigma_s) - \sigma_s], \quad N_{g,t} \sim \text{Poisson}(\lambda_{0g} + \lambda_1 x_t) \quad (29)
\]

\[
J_{b,t} = \sum_{j=1}^{N_{b,t}} [-\mu_s + \sigma_s - \text{Exponential}(\sigma_s)], \quad N_{b,t} \sim \text{Poisson}(\lambda_{0b} - \lambda_1 x_t), \quad (30)
\]

where \(x_t\) is a scalar, and \(J_{g,t} = 0\) and \(J_{b,t} = 0\) when \(N_{g,t} = 0\) and \(N_{b,t} = 0\), respectively. In the relevant region of the parameter space, the probability that \(N_{g,t}\) or \(N_{g,t}\) is larger than 1 is essentially zero, so one can interpret \(\lambda_{0g} + \lambda_1 x_t\) and \(\lambda_{0b} - \lambda_1 x_t\) as the quarterly probabilities of experiencing good and bad shocks, respectively. State dependence manifests itself via time variation in these probabilities. The conditional skewness of \(\eta_t\) is proportional to \(\lambda_1 x_t\). When \(\lambda_1 x_t\) is high, large positive shocks become more likely while large negative shocks become less likely, shifting probability mass from the left to the right tail.

The parameters \(\mu_s\) and \(\sigma_s\) are the mean and standard deviation of these large shocks (jump increments) in log labor income, respectively. In the interest of parsimony, I assume that the jump size distribution for good shocks equals that for the bad shocks multiplied by negative 1. When the sum of the Poisson intensities is constant, this restriction implies that the variance of \(\eta_t\) is constant. By construction, then, my estimates are consistent with the evidence on the lack of cyclical variation in second moments from GOS. By allowing \(\lambda_{0g}\) to differ from \(\lambda_{0b}\), this process can generate substantial unconditional skewness.

As discussed earlier, if the data are generated according to (29-30), the quarterly skewness index from Figure 1 is a consistent estimate of \(x_t\) up to a constant of proportionality. Therefore, when calibrating the idiosyncratic shock distribution, I set \(x_t\) equal to my skewness index, normalized it to have mean zero and variance one. After this normalization, \(\lambda_{0g}\) and \(\lambda_{0b}\) capture the unconditional probabilities of experiencing good and bad shocks, while \(\lambda_1\) captures the marginal effect of a 1 standard deviation increase in \(x_t\) on the probability of a good shock.

My calibration tries to match a number of statistics from Table 9, plus several additional moment conditions from statistics reported by GOS. I estimate the parameters so as to minimize a weighted sum of squared errors between a number of model-implied and data-implied moments.\(^{22}\) First, I try to match the average distance between the median and the 10\(^{th}\) percentile of 5-year income growth rates in expansions and recessions. I also target the average distance between the 90\(^{th}\) percentile and the median in expansions and recessions. Second, I target the changes

\(^{22}\)I also place some practical constraints on the parameters, which restrict the variance of jump shocks, \(\sigma_s\), and guarantee that the fitted poisson intensities are non-negative for most of the sample. If the fitted intensity is negative, I truncate \(\lambda_1 x_t\) so that the minimum intensity is zero the sum of the fitted intensities is \(\lambda_{0b} + \lambda_{0g}\).
from recessions versus expansions in the left tail and right tail width measures. Third, I target the average distance between the 90\textsuperscript{th} and 10\textsuperscript{th} percentiles of 1-year growth rates in expansions. Finally, I add information about the standard deviation, skewness, and kurtosis from GOS.

I place two additional restrictions on the model. Under my assumptions, the distribution of \( w_{A,t} - w_{A,t-k} \) can be decomposed into the sum of a gaussian component and a non-gaussian component. The variance of the gaussian component depends on \( \sigma_{\beta}^2, \sigma_{\epsilon}^2, \) and \( \sigma_{n}^2 \). Given that data are only available two different horizons (1-year and 5-year), I need an additional restriction to achieve identification. I choose to shut off profile heterogeneity by setting \( \sigma_{\beta}^2 = 0 \), which is relatively innocuous given my focus on state-dependent shocks. With respect to the state-dependent transitory shock, I fix \( \rho \) at its estimated value of 0.45.

I am interested in matching the time series behavior of cross-sectional quantiles of time-aggregated income growth rates, which cannot be expressed in closed form. However, conditional on the parameters governing \( \rho(L) \) and the income process, which I collect in a vector \( \beta \), I can calculate its characteristic function, \( \varphi_{k,t}(\omega; \beta) \equiv E_t\{\exp\{i\omega \cdot (w_{A,t} - w_{A,t-k})\}\} \) analytically. Using Lévy’s theorem, I recover the probability density function \( f_{k,t}(z; \beta) \) by taking the inverse Fourier transform of \( \varphi_{k,t}(\omega; \beta) \),

\[
  f_{k,t}(z; \beta) = \frac{1}{\pi} \cdot \text{Real} \left[ \int_0^\infty \varphi_{k,t}(\omega; \beta) e^{-i\omega z} d\omega \right],
\]

which involves a single numerical integration. I use the fractional fast Fourier transform to efficiently evaluate \( f_{k,t}(z; \beta) \) on a fine grid over the support of \( w_{A,t} - w_{A,t-k} \). By integrating \( f_{k,t}(z; \beta) \), I quickly and accurately recover the conditional cdf and quantile functions.\textsuperscript{73} Expressions for \( \varphi_{k,t}(z; \beta) \) and further details about the procedure are in Appendix A.3.2.

Table 14 presents estimates of the parameters governing the labor income process. \( \lambda_{0g} + \lambda_{0b} \) is about 2%, suggesting that the probability of receiving a large shock within a given year is about 8%. \( \lambda_{0b} \) is larger than \( \lambda_{0g} \), implying that large negative shocks are more likely to occur than large positive shocks. \( \lambda_1 \) is 0.26%, implying that, on an annualized basis, a 1 standard deviation increase in \( x_t \) shifts 1% of the probability mass from bad to good shocks.

In stark contrast with the jump shocks, the annualized standard deviation of the state-independent permanent, gaussian shock is only 3.7%, implying that permanent income is relatively safe when no jumps occur. The contribution from transitory shocks is more substantial. The calibrated value of \( \sigma_{\epsilon} \) is 13.5%, implying that the standard deviation of \( \epsilon_{A,t} - \epsilon_{A,t-k} \) is 19%.

\textsuperscript{73}The whole procedure takes about 2 milliseconds for each time period. I run some diagnostics using a variety of parametric densities and find that approximation errors associated with the estimated quantiles are on the order of \( 10^{-8} \).

A.21
This table presents the estimated parameters governing the distribution of idiosyncratic labor income shocks. Estimates are obtained by minimizing a weighted sum of squared errors between model-implied and data moments, both of which are displayed in Table 15. These moments are averages of quantiles or moments of the cross-sectional distribution of income from GOS.

Table 15 provides some goodness-of-fit statistics for the calibrated model. The left panel compares the moments implied by the calibrated model with their counterparts in the data. Generally speaking, the fit is quite close. By setting $\lambda_{0b} > \lambda_{0g}$, the model replicates the negative unconditional asymmetry which is observed in the data. The model comes pretty close to matching the cyclical variation in the 50-10 and 90-10 spreads, slightly overestimating variation in the former and underestimating the latter. It matches the level of the 90-10 spread at 1 and 5 year horizons almost perfectly. While the exact magnitudes of changes in the other quantiles do not match perfectly, particularly for the 99-50 percentile spread, the fit is reasonably close. The presence of rare, large shocks generates substantial cyclical variation in the tails of the distribution, while leaving the central quantiles essentially unchanged.

In addition to quantile-based measures from Table 9, I target several time series averages of cross-sectional moments from GOS. First, the model-implied standard deviation is relatively close, though slightly lower, to its value in the data.\textsuperscript{74} GOS provide a number of reasons to prefer quantile-based skewness measures such as those reported in Table 9. I also compare the calibrated (moment-based) skewness measures to their counterparts in the data. The fitted model has some trouble matching the level of unconditional skewness, though it does a better job of matching the observed change in skewness from expansions.\textsuperscript{75}

I also calculate the kurtosis of 5-year earnings growth. Guvenen et al. (2014b) emphasize the

\textsuperscript{74}It is worth noting that the fitted standard deviation is a bit further off at a 1-year horizon, suggesting that there is room for improvement in my specification of transitory risk.

\textsuperscript{75}Recall from the discussion above that the constant term, $a_4$, in (4) capturing unconditional skewness is more difficult to identify relative to the slope coefficient $b$. The same difficulty applies here as well.
Cross-sectional distribution of income growth rates

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Period</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>50-10 Percentile Spread</td>
<td>5 yr E</td>
<td>64.14</td>
<td>63.29</td>
</tr>
<tr>
<td></td>
<td>5 yr R</td>
<td>71.82</td>
<td>72.79</td>
</tr>
<tr>
<td></td>
<td>5 yr R-E</td>
<td>7.68</td>
<td>9.50</td>
</tr>
<tr>
<td>90-50 Percentile Spread</td>
<td>5 yr E</td>
<td>51.66</td>
<td>51.64</td>
</tr>
<tr>
<td></td>
<td>5 yr R</td>
<td>43.37</td>
<td>44.27</td>
</tr>
<tr>
<td></td>
<td>5 yr R-E</td>
<td>-8.30</td>
<td>-7.37</td>
</tr>
<tr>
<td>50-25 Percentile Spread</td>
<td>5 yr R-E</td>
<td>3.01</td>
<td>1.81</td>
</tr>
<tr>
<td>75-50 Percentile Spread</td>
<td>5 yr R-E</td>
<td>-1.72</td>
<td>-0.79</td>
</tr>
<tr>
<td>95-50 Percentile Spread</td>
<td>5 yr R-E</td>
<td>-15.47</td>
<td>-15.15</td>
</tr>
<tr>
<td>99-50 Percentile Spread</td>
<td>5 yr R-E</td>
<td>-26.71</td>
<td>-18.62</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>5 yr E</td>
<td>66.52</td>
<td>62.36</td>
</tr>
<tr>
<td>Skewness (Moment)</td>
<td>5 yr E</td>
<td>-1.18</td>
<td>-0.29</td>
</tr>
<tr>
<td></td>
<td>5 yr R</td>
<td>-1.80</td>
<td>-0.76</td>
</tr>
<tr>
<td></td>
<td>5 yr R-E</td>
<td>-0.62</td>
<td>-0.47</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>5 yr</td>
<td>≈ 10</td>
<td>10.53</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Horizon</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean - 50 Percentile</td>
<td>1 yr</td>
<td>0.790</td>
</tr>
<tr>
<td>50-10 Percentile Spread</td>
<td>1 yr</td>
<td>0.520</td>
</tr>
<tr>
<td>90-50 Percentile Spread</td>
<td>1 yr</td>
<td>0.659</td>
</tr>
<tr>
<td>Kelley skewness</td>
<td>1 yr</td>
<td>0.709</td>
</tr>
<tr>
<td>50-10 Percentile Spread</td>
<td>3 yr</td>
<td>0.801</td>
</tr>
<tr>
<td>90-50 Percentile Spread</td>
<td>3 yr</td>
<td>0.702</td>
</tr>
<tr>
<td>90-10 Percentile Spread</td>
<td>5 yr</td>
<td>0.671</td>
</tr>
<tr>
<td>90-10 Percentile Spread</td>
<td>3 yr</td>
<td>0.695</td>
</tr>
<tr>
<td>Kelley skewness</td>
<td>3 yr</td>
<td>0.667</td>
</tr>
<tr>
<td>90-10 Percentile Spread</td>
<td>5 yr</td>
<td>0.762</td>
</tr>
<tr>
<td>90-10 Percentile Spread</td>
<td>5 yr</td>
<td>0.789</td>
</tr>
</tbody>
</table>

Table 15: Goodness-of-fit statistics for calibrated income process

This table presents goodness-of-fit statistics for the calibrated model for idiosyncratic labor income risk. The left panel compares model-implied moments with their counterparts in the data. These moments are averages of quantiles or moments of the cross-sectional distribution of income growth rates for individuals in the 91st through 95th percentiles of the income distribution, most of which are in Table 9. The right panel calculates the $R^2$ from univariate regressions of time series of model-implied statistics on the same statistics from GOS for the cross-section of income growth rates for male earners in the U.S. population.

An extremely high degree of excess kurtosis observed in the SSA earnings data, which is at odds with the assumption of normally distributed shocks that is ubiquitous in the literature on calibrating earnings distributions. Their estimates, which are constructed using similar methods to GOS, suggest that the kurtosis of 5-year income growth rates for relatively high earners is about 10. My model with jump shocks generates an average kurtosis of 10.5, consistent with this evidence.

The right panel of Table 15 compares model-implied time series with their counterparts in the data. Recall that GOS report annual time series for the shape of the cross-sectional distribution of income growth rates, where the entire U.S. population of male earners is treated as a single cross-section. In contrast, my calibrated model targets the level of income risk faced by relatively high earners. However, I can still assess whether our fitted model qualitatively matches the time series dynamics of these cross-sectional moments. At 1, 3, and 5-year horizons, I report the $R^2$ from a univariate regression of the GOS data on model-implied statistics. I consider four different time series, two of which are plotted in Figure 7: the difference between the mean and the median, the 50-10 spread, the 90-10 spread, and Kelley’s skewness.
Overall, my model matches essentially all of these time series moments quite well. The $R^2$'s tend to be highest at a 3-year horizon, which is somewhat reassuring given that no data from that horizon were used to estimate the linear index, $x_t$. I have some trouble matching the transitory variation in the 50-10 spread in recessions, suggesting that I am omitting some important sources of transitory risk. The high $R^2$'s on the Kelley’s skewness measure suggest that we do a reasonable job of capturing the cyclical variation in the overall asymmetry of the distribution.

A.3.2 Quasi-analytic calculation of conditional quantiles

This section details the steps performed in order to calculate conditional quantiles of the calibrated earnings process. First, I derive the characteristic functions for the idiosyncratic component of the cross-sectional distribution of time-aggregated labor income, $\varphi_{k,t}(\omega; \beta)$, conditional on the skewness index ($x_t$) and parameters governing the earnings process ($\beta$). I begin from equation (26). By assumption, the first term involving profile heterogeneity is zero, and all shocks are independent conditional on $x_t$. Then, it follows that

$$\log \varphi_{k,t}(\omega; \beta) = \sum_{j=0}^{\infty} \left[ h(\theta_{k,j}(\rho) \cdot \omega, x_{t-j}) \right] - \omega^2 \sigma_e^2, \tag{32}$$

where $h(\omega, x_t) \equiv \log E_t[\exp(i\omega \eta_t)|x_t]$ and $\theta_{k,j}(\rho)$ is the $j^{th}$ term in the lag polynomial defined implicitly by equation (26). Given my distributional assumptions,

$$h(\omega, x_t) = (\lambda_0 + \lambda_1 x_t)[\exp(i\omega \cdot (\mu_s - \sigma_s) - \log(1 - i\omega \sigma_s)) - 1 - \mu_s i\omega] + (\lambda_0 - \lambda_1 x_t)[\exp(i\omega \cdot (\sigma_s - \mu_s) - \log(1 + i\omega \sigma_s)) - 1 + \mu_s i\omega] - \frac{1}{2} \omega^2 \sigma_n^2.$$

Thus, for a given $z$ and $\omega$, I can evaluate the integrand in (31) in closed form.

As is well known, one can approximate the integral in (31) numerically using the trapezoidal rule, a method which also has a number of desirable computational properties. The trapezoidal rule approximates the integral as a sum over an equally-spaced grid from $[0, \Omega]$. By using an equally-spaced grid, I can exploit the computational efficiency of the Fast Fourier Transform in order to solve for the density $f_{k,t}(z; \beta)$ over a fine grid, $[-Z, Z]$ which covers the support of the distribution of idiosyncratic shocks.\textsuperscript{76} Prior to executing the Fourier inversion algorithm, I also normalize the distribution of $w_{A,t} - w_{A,t-k}$ in order to have mean zero and variance 1. I use $2^{10} = 1024$ grid points, and I set $\Omega = 20$ and $Z$ equal to 9 standard deviations.\textsuperscript{77}

\textsuperscript{76}Following the method in Mastro (2013), I use the Fractional Fast Fourier Transform to break the (otherwise mechanical) link between $\Omega$ and $Z$ which is required in order to use the Fast Fourier Transform. This dramatically improves the accuracy of the algorithm.

\textsuperscript{77}This combination of tuning parameters produced highly accurate results for a variety of parametric densities.
The final step is to integrate the density in order to solve for the conditional CDF, and, in turn, the conditional quantiles. After some experimentation, I find that Simpson’s rule (which eliminates a first-order bias term which is associated with the trapezoid rule) produces extremely accurate results. Finally, I use quadratic interpolation of the conditional CDF in order to solve for the desired conditional quantiles. A number of simulation exercises confirmed that the approximation errors associated with the Fourier inversion, integration, and interpolation procedure are extremely small. Further details are available upon request.

B Theoretical Model Appendix

B.1 Full Structure of General Model

The model is a Lucas (1978) endowment economy with incomplete markets. Agents’ consumption stream derives from two types of assets (trees), each of which delivers an uncertain stream of future cash flows (fruit). Between periods, total fruit output grows stochastically and the growth of each tree is potentially subject to aggregate and idiosyncratic shocks. The first type of tree, human capital ($H_i^t$), is a claim on future labor income, which will equal consumption in equilibrium. Dividend income, other sources of income are included as well. In addition, agents have the option to purchase shares ($N_{kt}$) in $K$ other financial assets in zero net supply, paying dividends ($D_{kt}$).

In the model, the key distinction between the two types of assets is that labor income is subject to idiosyncratic risk, meaning that different investors will receive different returns over the same holding period because their trees will not all grow at the same rate. Defining the aggregate quantity $C_t = \int I_C i^t di$, the fruit production of the first type of tree grows at rate $C_t/C_{t-1} \times \exp(\eta_i^t)$, where $\eta_i^t$ is a shock which is independently and identically distributed across agents satisfying $E[\exp(\eta_i^t)] = 1$.\textsuperscript{78} Households are unable to buy or sell human capital, nor trade contingent claims on realizations of $\eta_i^t$.\textsuperscript{79} Dividend income is only subject to aggregate risk, so cash flows from trees of the second type grow at the same rate ($D_{kt}/D_{k,t-1}$). Finally, the total supply of each type of tree in the economy is fixed, so that, in equilibrium, aggregate consumption will equal total fruit production.

At time 0, each agent begins with an initial endowment $H_{i0}^t$. Thereafter, each agent chooses her

\textsuperscript{78}I will formalize my assumptions about $\eta_i^t$ later in the section.
\textsuperscript{79}Since the financial assets are in zero net supply, no-trade will be an equilibrium. Therefore, I could assume that households are able to buy and sell human capital without affecting any of the results below. The key friction is the inability to write contingent claims on $\eta_i^t$. 

A.25
consumption (the numeraire) and investment $N^i_t$ to maximize (5). All financial assets are in zero net supply, so market clearing will imply $N^i_{kt} = 0$ for all $i$, $k$, and $t$. These assumptions imply the following budget constraint

$$C^i_t + \sum_{k=1}^{K} P_{kt} N^i_{kt} = C_t \exp(\eta^i_t) H^i_{t-1} + \sum_{k=1}^{K} (P_{kt} + D_{kt}) N^i_{k,t-1}$$

subject to $\sum_{k=1}^{K} P_{kt} N^i_{kt} > -W$, where $P_{kt}$ is the price of the $k^{th}$ financial asset. The borrowing constraint, which will not bind in equilibrium, is simply present in order to rule out Ponzi schemes. Under my assumptions, $H^i_t = \exp(\eta^i_t) H^i_{t-1}$. This will imply that

$$C^i_t = H^i_t C_t \quad \Rightarrow \quad \Delta c^i_t = \Delta c_t + \eta^i_t.$$  

(34)

Assumption 1 gives my general model for aggregate dynamics.

**Assumption 1.** Aggregate variables evolve according to the stationary VAR model:

$$y_{t+1} = \mu_y + F_y y_t + G_{y,t} z_{y,t+1} + J_{y,t+1},$$

(35)

with $y_0$ given, where

(i) $z_{t+1}$ is i.i.d. $N(0,1)$,

(ii) $G_{y,t} G_{y,t}'$ is a symmetric, positive semi-definite matrix, a function of $y_t$,

(iii) $F_y$ has all of its eigenvalues inside the unit circle.

(iv) $J_{y,t+1}$ is a compound Poisson shock with mutually independent, i.i.d. increments and arrival intensity vector $\lambda_{y,t}$, a function of $y_t$, and

(v) $\Delta c_{t+1} = S_c y_{t+1}$ and $\Delta d_{k,t+1} = S_k y_{t+1}$ for $L \times 1$ vectors $S_c$ and $S_1, \ldots, S_K$.

I summarize the jump size distribution for the compound Poisson shocks with $\Psi_y(u)$, the $L \times 1$ vector-valued function whose $j^{th}$ element is the moment-generating function of the size distribution for the $j^{th}$ jump component. I need little structure on $\Psi_y(u)$ beyond the existence of such a function (and boundedness for certain values of $u$). $G_{y,t}$ need not have full rank. For example, one can impose cointegration restrictions on consumption and dividends or make dividends a levered claim on aggregate consumption via appropriate restrictions on $F_y$ and $G_{y,t}$.

\footnote{Without loss of generality, we normalize the total supply of human capital to equal 1.}
Assumption 2 allows the distribution of the idiosyncratic shock to depend on the realization of the aggregate state vector, $y_{t+1}$. This structure means that, ex post, aggregate shocks (e.g. consumption declines) need not be distributed equally across agents. Denote agent $i$’s private information by the filtration $\mathcal{F}^i_t$ and public information by $\mathcal{F}_t = \bigcap_i \mathcal{F}^i_t$.

**Assumption 2.** The following statements are true.

(i) $\eta_t = 1_M \tilde{\eta}^i_t$, and, conditional on $y_{t+1}$, $\tilde{\eta}^i_t$ is generated according to

$$
\tilde{\eta}^i_{t+1} = G_{\eta,t+1} z^i_{\eta,t+1} + J^i_{\eta,t+1} - \log E \left[ \exp \left( G_{\eta,t+1} z^i_{\eta,t+1} + J^i_{\eta,t+1} \right) \bigg| y_{t+1} \right],
$$

where $z^i_{\eta,t+1}$ is a vector of standard normal random variables that is i.i.d. across agents and over time, $G_{\eta,t+1}$ is a symmetric, positive semi-definite matrix. $J^i_{\eta,t+1}$ is a compound Poisson shock with mutually independent, i.i.d. increments (across agents and over time for a given agent) and arrival intensity vector $\lambda_{\eta,t+1}$.

(ii) $y_t \in \mathcal{F}_t$ and the joint distribution of $(y_{t+1}, \eta^i_{t+1})|\mathcal{F}^i_t$ is the same as the joint distribution of $(y_{t+1}, \eta^i_{t+1})|y_t$.

As above, I will describe the jump size distribution for the idiosyncratic shocks by $\Psi_\eta(u)$ be, an $M \times 1$ vector-valued function whose $j$th element is the moment-generating function of the size distribution for the $j$th jump component.

As discussed in the text, Assumption 2.i is the most important and restrictive. It implies that, conditional on public information, $\tilde{\eta}^i_t$ does not depend on any of its past realizations. Following Toda (2014b), Assumption 2.ii says that all agents have rational expectations and consider the same set of information when choosing their investments. It also says that $y_t$, which is common knowledge, is a sufficient statistic for describing aggregate and idiosyncratic dynamics.

Assumption 3 places general restrictions on the model which are necessary to ensure that, after performing the Campbell and Shiller (1988) approximation, the model generates valuation ratios which are exponential affine in the state vector $y_t$.

**Assumption 3.** The following statements are true.

(i) $G_{y,t} G_{y,t}^t = h_y + \sum_{j=1}^L H_{y,j} y_{j,t}$, where $h_y$ and $H_{y,1}, \ldots, H_{y,L}$ are $L \times L$ matrices.

(ii) $\lambda_{y,t} = l_0 + l_1 y_t$, where $l_0$ and $l_1$ are $L \times 1$ and $L \times L$ matrices.

(iii) The $(1,1)$ element of $G_{\eta,t} G_{\eta,t}^t$ equals $h_\eta + H_{\eta,1}^t y_t$, where $h_\eta$ and $H_{\eta,1}$ are a scalar and a $M \times 1$ vector, respectively. All other elements of $G_{\eta,t} G_{\eta,t}^t$ are zero.
(iv) \( \lambda_{n,t} = l_{n0} + l_{n1}y_t \), where \( l_{n0} \) and \( l_{n1} \) are \( M \times 1 \) and \( M \times L \) matrices, respectively.

Assumptions 3.i-ii are standard restrictions which ensure that the model’s solution falls into the affine class.\(^{81}\) In the absence of idiosyncratic shocks, my framework nests long-run risk-type representative agent models with Poisson jumps, such as DY.

Assumptions 3.iii-vi parameterize the idiosyncratic shocks. Assumption 3.iii allows for a normally-distributed “diffusion” shock, and the variance of this shock is allowed to be state-dependent.\(^{82}\) As such, I can easily allow for CCSV within our model. Given that GOS find little evidence of CCSV in their Social Security Administration dataset, my analysis will focus more on state-dependence in the “jump” shocks. However, the theoretical implications of and intuition for time-varying volatility of the Gaussian shocks are similar. Assumption 3.iv parallels Assumption 3.ii, allowing for state-dependence in the idiosyncratic jump intensities.

Note that equation (36) includes a location adjustment. Lemma 1 says that under the distributional assumptions above, this adjustment is affine.

**Lemma 1.** Let Assumptions 2 and 3.iii-iv hold. Then,

\[
\log E[\exp(G_{\eta,t+1}z_{\eta,t+1}^i + J_{\eta,t+1}^i) | y_{t+1}] = \mu_\eta + F_\eta y_{t+1},
\]

where \( \odot \) denotes element-by-element multiplication,

(i) \( \mu_\eta = -1/2h_\eta0e_1 - l_{\eta0}(\Psi_\eta(1_M) - 1_M), \) and

(ii) \( F_\eta = -1/2e_1 \odot h_\eta1 - l_{\eta1} \odot [(\Psi_\eta(1_M) - 1_M) \odot 1'_L]. \)

Therefore, I can fully decouple any assumptions about the time-variation in the distribution of idiosyncratic consumption risk from aggregate consumption while preserving linearity of the process. My solutions below replace the expectation with this affine form.

An equilibrium in this economy is a sequence of state-contingent prices \( \{P_{1,t}, \ldots, P_{K,t}\}_{t=0}^\infty \) and allocations \( \{C_{i,t}, i \in I\}_{t=0}^\infty \) which solves agents’ optimization problems and satisfies market clearing in the capital markets. I restrict attention to symmetric (no-trade) equilibria where all agents consume their endowments. Such an equilibrium can obtain since all agents have identical homothetic preferences and access to the same investments, making their first order conditions

\(^{81}\)The key property I exploit is that \( \log E[\exp(u'y_{t+1})] \) and \( \log E[\exp(u \cdot \eta_{t+1}^i)|y_{t+1}] \) are affine functions of \( y_t \) and \( y_{t+1} \), respectively. Therefore, my solution method extends to other families of conditional distributions that also have this property. For example, Bekaert and Engstrom (2013) show that the sum of gamma random variables with time-varying shape parameters satisfies this property.

\(^{82}\)Without loss of generality, I can concentrate all of the “diffusion risk” in the first element of \( \tilde{\eta}_t \), so \( M \), the dimension of \( \eta_t^i \), is solely determined by the number of independent sources of jump risk.
identical. Market clearing is trivially satisfied. Toda (2014b) establishes the existence and essential uniqueness of a symmetric equilibrium in a similar environment.

In the text, I provide quasi-analytical solutions to a log-linearized model with affine, jump-diffusion dynamics. Tractability maintains without these specific distributional assumptions, in the sense that the wealth distribution does not enter the state space when solving the model, provided that shocks have the permanent, proportional form in (6).

The general solution obtains by imposing that, in equilibrium, (8) and (9) must hold for all assets in the investment opportunity set. Plugging the consumption claim and financial asset returns into (9) yields

\begin{align*}
1 &= E_t \left[ \delta^\theta \left( \frac{C_{t+1}}{C_t} \right)^{1-\gamma} \left( \frac{WC_{t+1} + 1}{WC_t} \right)^\theta \exp((1-\gamma)\eta_{t+1}) \right] \\
1 &= E_t \left[ \delta^\theta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} \left( \frac{WC_{t+1} + 1}{WC_t} \right)^{-(1-\theta)} \exp(-\gamma \cdot \eta_{t+1}) \left( \frac{P_{k,t+1} + 1}{D_{k,t+1}} \right) \left( \frac{P_{k,t} + 1}{D_{k,t}} \right) \right],
\end{align*}

a system of nonlinear equations involving the two key valuation ratios, \( WC_t \) and \( P_{kt}/D_{kt} \). Since all of exogenous quantities in (38-39) are stationary by Assumptions 1-3, the model may be solved numerically by finding the value of \( WC_t \) that satisfies (38) for each value of \( y_{t+1} \) in the state space. Then, given the solution for \( WC_t \), one can use (39) to solve for the equilibrium price-dividend ratios for the financial assets.

**B.2 The Term Structure of Risk Premia**

In this section, I derive expressions for the term structure of risk premia. At least three types of claims are of potential interest:

- Dividend strips / “Zero-coupon” equity: \( D_{k,t+h} \), a single dividend payment from one of the risky assets,
- Non-defaultable bonds: a real or nominal risk-free payment at time \( t + h \), or
- Consumption strips: \( C_{t+h}^i \), an individual agent’s consumption at time \( t + h \).

Prices for the first two types of assets are (mostly) observable and, as such, supply additional dimensions with which to test the model. While the prices for individual consumption strips are unobservable due to market incompleteness, they help to reveal information about the nature of discounting over different time horizons.
Within my framework, we can use identical methods to price zero-coupon bond and equity claims by judiciously parameterizing the selector vectors for the financial assets. For example, I can price an asset that delivers a risk-free, constant real payoff by assuming that its selector matrix is zero. The associated “dividend” prices are real bond prices, up to an irrelevant constant of proportionality. Nominal, default-free bonds are also easy to price. If I assume that the log of the inflation rate \( \pi_t \) equals \( S'_y y_{t+1} \), then the real log change in the value of a fixed coupon payment is \(-\pi_t\). By assuming that the “dividend” of one of the risky assets grows at rate \(-S'_y y_{t+1}\), the prices of its “dividends” are proportional to nominal bond prices.

Let \( P^h_{k,t} \) be the price of a zero-coupon equity claim, an asset which pays \( D_{k,t+h} \) at time \( t+h \). Trivially, no arbitrage requires that \( P^0_{k,t} = D_{k,t} \). Then \( R^h_{k,t+1} \equiv P^{h-1}_{k,t+1}/P^h_{k,t} \) is the holding period return from \( t \) to \( t+1 \) for an investor who purchased an \( h \)-period zero-coupon equity claim at time \( t \). No arbitrage also implies \( P_{k,t} = \sum_{h=1}^{\infty} P^h_{k,t} \), so

\[
R_{k,t+1} = \frac{P_{k,t+1} + D_{k,t+1}}{P_{k,t}} = \sum_{h=0}^{\infty} \frac{P^h_{k,t+1}}{P^h_{k,t}} = \sum_{h=0}^{\infty} \frac{P^{h+1}_{k,t+1}}{P^{h}_{k,t}} = \sum_{h=1}^{\infty} \frac{P^h_{k,t}}{P^h_{k,t}} \cdot R^h_{k,t+1}, \quad (40)
\]

meaning that \( R_{k,t+1} \), the return of the claim on the entire dividend stream, is a weighted average of the claims on the individual zero-coupon equity claims. It follows that the risk premium for asset \( k \) is a weighted average of the risk premia for its associated zero-coupon equity claims.

Proposition 5 says that the valuation ratios for the zero-coupon consumption and dividend claims are affine functions of \( y_t \).

**Proposition 5.** Let Assumptions 1-3 hold. The log-linearized model satisfies

\[
(i) \quad \log(P^h_{c,t}/C_t) \equiv w^h_{c,t} = A^h_0 + y'_t A^h,
\]

\[
(ii) \quad \log(P^h_{k,t}/D_{k,t}) \equiv p^h_{k,t} = A^h_{0,k} + y'_t A^h_k, \text{ for } k = 1, \ldots, K.
\]

for all \( t \) and \( h \geq 0 \), where \( A^h_0, A^h_{0,1}, \ldots, A^h_{0,K} \) are scalars and \( A^h, A^h_1, \ldots, A^h_K \in \mathbb{R}^K \).

An immediate implication, in light of the discussion above, is that real and nominal bond yields are also affine functions of \( y_t \). Expressions such as \( \log E_t[D_{k,t+h}/D_{k,t}] \) are affine as well. Therefore, I can study how the term structures of real bond yields, expected dividend growth rates, and risk premia evolve over time.
B.3 Additional Discussion of Quantitative Model Calibration

I set $\lambda_1$, the sensitivity of the disaster probability to a change in $x_t$, exactly equal to its calibrated value from Table 14. Given my emphasis only on the state-dependent component of idiosyncratic risk, I set the unconditional disaster probability equal to $2.5 \times \lambda_1$, which implies that the probability of the fitted intensity going negative is very small, though still nonzero.\(^{83}\)

$\mu_c$ and $\varphi_c$ are set to generate a mean and volatility of aggregate consumption growth which is roughly in line with the data. Note that, given that I have chosen the idiosyncratic risk parameters to correspond with the income risk faced by relatively high earners, it is not immediately obvious that “aggregate consumption” in the model should be tied to NIPA consumption data. For instance, Malloy et al. (2009) find evidence that the average consumption growth of stockholders is more highly correlated with returns relative to non stockholders. Parker and Vissing-Jorgensen (2010), Guvenen et al. (2014c), and Guvenen et al. (2014a) provide additional evidence of above-average cyclicality of high earners. Nonetheless, as predictability of average consumption growth is not the main focus of the paper, I maintain the NIPA benchmark for ease of comparison with the literature.

I set the drift of aggregate dividend growth ($\mu_d$) equal to 75 basis points per year, somewhat lower than the mean of aggregate consumption growth. I assume that a 1 standard deviation increase in $x_t$ increases expected dividend growth by 2.5% in the following quarter. AGiven the degree of persistence in $x_t$, dividends are fairly predictable at short horizons, but are fairly difficult to predict at longer horizons. I set the loading of the dividend innovation on the consumption growth innovation, $\pi_c$, equal to 2.5, which is identical to the value in Bansal et al. (2012). The correlation between consumption growth and dividend growth innovations is 52%. My choice of $\varphi_d$ implies that the quarterly volatility of the dividend growth innovation is just shy of 6%.

The parameters $\pi_x$ and $\pi_{\sigma}$ imply that dividend innovations are moderately correlated with $x_t$ and $\sigma_t^2$ shocks; pairwise correlations are 31% and -25%, respectively.

For the preference parameters, I set $\gamma = 11$ and $\psi = 2$, which fall in the standard range of choices in the literature with Epstein-Zin preferences. Given that $\psi > 1$, agents have a preference for the early resolution of uncertainty, implying that they are willing to pay a premium for assets whose returns hedge against bad news about the state vector (low $x_t$ or high $\sigma_t^2$). The discount factor $\delta$ is chosen to roughly match the real risk free rate. It is worth noting that matching observed risk-free rates the presence of idiosyncratic risk necessitates the use of discount factors which are substantially lower relative to standard choices in representative agent models.

\(^{83}\)My calibration embeds a negative correlation between $x_t$ and $\sigma_t^2$ shocks, implying that the unconditional distribution of $x_t$ is negatively skewed (consistent with the data). Since the disaster probability decreases in $x_t$, the likelihood that the fitted intensity goes negative is quite small.
B.4 Proofs of Propositions

B.4.1 Proof of Lemma 1 (Aggregation Restrictions)

Using the independence on $z_{\eta,t+1}^i$ and $J_{\eta,t+1}^i$, it follows that

$$E[\exp(\eta_{t+1}^i)|y_{t+1}, y_t] = \exp(1'_M \mu_\eta + 1'_M F_\eta y_t + 1/2(h_\eta + h_\eta')y_t)E[\exp(1'_M J_{\eta,t+1}^i)|y_{t+1}, y_t]$$

$$= \exp[1'_M \mu_\eta + 1'_M F_\eta y_t + 1/2(h_\eta + h_\eta')y_t]$$

$$\times \exp(1'_M (l_{\eta 0}(\Psi_\eta(1_M) - 1_M) + l_{\eta 1} \odot [(\Psi_\eta(1_M) - 1_M) \odot 1'_L]y_t)), $$

where I used the moment-generating function of the normal distribution and a compound Poisson process to go from the first to the second line. To establish the claim, the log of this expression has to equal zero for all values of $y_t$. Substituting in the given expressions for $\mu_\eta$ and $F_\eta$ yields zero, so the restriction holds.

B.4.2 Proof of Proposition 2 (Wealth-Consumption Ratios)

I will begin by solving for the wealth-consumption ratio coefficients, then proceed to solve for the price-dividend ratios. My solution method closely follows Eraker and Shaliastovich (2008) and Drechsler and Yaron (2011), who solve representative agent models with jump-diffusion shocks in continuous time and discrete time, respectively. Additional details and discussion are available in Eraker and Shaliastovich (2008) and Drechsler and Yaron (2011).

Before working with the Euler Equations, I introduce two lemmas, which provide analytical expressions for expectations of linear functions of the state vector, $\eta_{t+1}^i$ and $y_{t+1}$, respectively.

**Lemma 2.** Let Assumptions 2 and 3.iii-iv hold. Then,

$$E[\exp(u \cdot \eta_{t+1}^i)|y_{t+1}] = \exp[\beta_0(u) + \beta'(u)y_{t+1}],$$

where $\beta_0(u): \mathbb{R} \to \mathbb{R}$ and $\beta(u): \mathbb{R} \to \mathbb{R}^L$ for $u \in \mathbb{R}$ are given by

(i) $\beta_0(u) = \mu'_\eta 1_M u + 1/2u^2h_\eta + l'_\eta(\Psi_\eta(u1_M) - 1_M)$ and

(ii) $\beta(u) = F'_\eta 1_M u + 1/2u^2H_\eta + l'_\eta(\Psi_\eta(u1_M) - 1_M)$.
Proof. By definition, \( \eta_{t+1}^i = 1_M \eta_{t+1}^i \). By the conditional independence of \( z_{n,t+1}^i \) and \( J_{n,t+1}^i \),

\[
\log E[\exp(u \cdot \eta_{t+1}^i) | y_{t+1}] = u_1' \mu_y + F_y y_{t+1} + \log E[\exp(u_1' G_{n,t+1} z_{n,t+1}^i) | y_{t+1}]
\]

\[
\log E[\exp(u_1' G_{n,t+1} z_{n,t+1}^i) | y_{t+1}] = \frac{1}{2} u_1^2 y_{t+1} G_{n,t+1} G'_{n,t+1} 1_M = \frac{1}{2} u_1^2 [h_{n,0} + H'_{n,1} y_{t+1}]
\]

\[
\log E[\exp(u_1' J_{n,t+1}^i) | y_{t+1}] = \lambda_{n,t+1} [\Psi_y(u_1) - 1_M] = [l_{y0} + l_{y1} y_{t+1}] [\Psi_y(u_1) - 1_M],
\]

where I evaluate expectations using the moment generating functions of the multivariate normal and compound Poisson distributions then plug in the assumed affine functional form from Assumptions 3.iii-iv. Collecting the constants and terms multiplying \( y_{t+1} \) yields the desired result.

\[\Box\]

**Lemma 3.** Let Assumptions 1 and 3 hold. Then,

\[E_t[\exp(u' y_{t+1})] = \exp[f(u) + g(u)' y_t],\]  

where \( f(u) : \mathbb{R}^L \rightarrow \mathbb{R} \) and \( g(u) : \mathbb{R}^L \rightarrow \mathbb{R}^L \) for \( u \in \mathbb{R}^L \) are given by

(i) \( f(u) = \mu_y u + \frac{1}{2} u' h_y u + l_{y0}' [\Psi_y(u) - 1_L] \)

(ii) \( g(u) = F_y u + \frac{1}{2} [u' H_{yi} u]_{i=1,\ldots,L} + l_{y1}' [\Psi_y(u) - 1_L] \)

and \([u' H_{yi} u]_{i=1,\ldots,L}\) is the \( L \times 1 \) vector whose \( i^{th} \) component equals \( u' H_{yi} u \).

Proof. The proof is similar to that from the previous proposition. I start by using the conditional independence of \( z_{y,t+1} \) and \( J_{y,t+1} \) to write

\[
\log E_t[\exp(u' y_{t+1})] = u' [\mu_y + F_y y_t] + \log E_t[\exp(u' G_{y,t+1} z_{y,t+1})] + \log E_t[\exp(u' J_{y,t+1})]
\]

\[
\log E_t[\exp(u' G_{y,t+1} z_{y,t+1})] = \frac{1}{2} u' G_{y,t} G'_{y,t} u = \frac{1}{2} u' h_y u + \frac{1}{2} \sum_{j=1}^L u' H_{y,j} u \cdot y_{j,t}
\]

\[
\log E_t[\exp(u' J_{y,t+1})] = \lambda_{y,t} [\Psi_y(u) - 1_L] = [l_{y0} + l_{y1} y_t] [\Psi_y(u) - 1_L].
\]

Again, I use moment generating functions to evaluate expectations, plug in the functional forms from Assumptions 3.i-ii, then collect terms. See also Drechsler and Yaron (2011) section A.1. \[\Box\]

I will assume that the wealth-consumption ratio \( wc_t = A_0 + A'_y y_t \). By Assumptions 1-2, I can write consumption growth in vector notation as \( \Delta c_t^i = S'_c y_t + \eta_t^i \). Combining (12) with my
assumption, the log-linearized Euler equation for the consumption claim is

\[
1 = E_t \left[ \exp \left\{ \theta \log \delta + \theta \kappa_c + \theta (\rho_c - 1) A_0 + (1 - \gamma) (S'_{t} y_{t+1} + \eta_{t+1}^i) + \theta (\rho_c A' y_{t+1} - A' y_{t}) \right\} \right] \\
= \exp(\theta \log \delta + \theta \kappa_c + \theta (\rho_c - 1) A_0 - \theta A' y_{t}) E_t \left[ \exp \left\{ [(1 - \gamma) S'_c + \theta \rho_c A'] y_{t+1} + (1 - \gamma) \eta_{t+1}^i \right\} \right] \\
= \exp(\theta \log \delta + \theta \kappa_c + \theta (\rho_c - 1) A_0 - \theta A' y_{t}) \\
\times E_t \left[ \exp \left\{ [(1 - \gamma) S'_c + \theta \rho_c A'] y_{t+1} + \log E_t [(1 - \gamma) \eta_{t+1}^i | y_{t+1}] \right\} \right] \\
= \exp(\theta \log \delta + \theta \kappa_c + \theta (\rho_c - 1) A_0 - \theta A' y_{t}) \\
\times E_t \left[ \exp \left\{ [(1 - \gamma) S'_c + \theta \rho_c A'] y_{t+1} + \beta_0 (1 - \gamma) + \beta (1 - \gamma) y_{t+1} \right\} \right],
\]

where the second and third equalities use the law of iterated expectations and the last line follows from Lemma 2. Using Lemma 3 to evaluate the expectation yields

\[
\theta (\log \delta + \kappa_c + (\rho_c - 1) A_0) - \theta A' y_{t} = -f((1 - \gamma) S'_c + \theta \rho_c A + \beta(1 - \gamma)) + \beta_0 (1 - \gamma) - g((1 - \gamma) S'_c + \theta \rho_c A + \beta(1 - \gamma)) y_{t}. \tag{47}
\]

Since the Euler equation holds for each \( y_t \) in the state space, the solution must satisfy

\[
f((1 - \gamma) S'_c + \theta \rho_c A + \beta(1 - \gamma)) + \beta_0 (1 - \gamma) = -\theta (\log \delta + \kappa_c + (\rho_c - 1) A_0) \tag{43}
\]

\[
g((1 - \gamma) S'_c + \theta \rho_c A + \beta(1 - \gamma)) = A \theta, \tag{44}
\]

an \((L+1)\)-dimensional system of equations in \( A \) and \( A_0 \). This system does not have an analytical solution in general; however, it is relatively straightforward to solve the system numerically.

In addition to the primitive parameters governing preferences and cash flows, the system (43-44) also depends on the log-linearization constants \( \kappa_c \) and \( \rho_c \). Following Drechsler and Yaron (2011) and Eraker and Shaliastovich (2008), I choose the linearization point to equal the unconditional mean of the wealth-consumption ratio. In particular, I choose \( \overline{wc} \) so that

\[
\log (\rho_c) - \log (1 - \rho_c) = \overline{wc} = E(wc_t) = A_0 + A'E(y_t), \tag{45}
\]

which, when combined with the definition of \( \kappa_c \), implies that (see DY equation A.2.2)

\[
\kappa_c + (\rho_c - 1) A_0 = -\log \rho_c - (\rho_c - 1) A'E(y_t). \tag{46}
\]

I can then substitute (46) into (43), yielding

\[
f((1 - \gamma) S'_c + \theta \rho_c A + \beta(1 - \gamma)) + \beta_0 (1 - \gamma) = -\theta (\log \delta - \log \rho_c - (\rho_c - 1) A'E(y_t)), \tag{47}
\]

leaving (44) and (47), an exactly identified system of equations in \( A \) and \( \rho_c \). Then, given these
solutions, I can use the expressions above to derive $A_0$, $\kappa_c$, and $\kappa$.

I will assume that the price-dividend ratio for asset $k$, $pd_{k,t} = A_{0,k} + A_k'y_t$. By Assumption 1, I can write dividend growth as $\Delta d_{k,t} = S_k'y_t$. Since the dividend claims are financial assets, I can price them using the projected pricing kernel in (13). Note that (13) becomes

$$m_{t+1} = \kappa - \gamma \Delta c_{t+1} - (1 - \theta)(\rho_c w_{c,t+1} - wc_t) + \beta (-\gamma)'y_{t+1},$$

(48)

where $\kappa \equiv \theta \log \delta - (1 - \theta)\kappa_c + \beta_0(-\gamma)$. Plugging in the projected kernel, the log-linearized Euler equation for the $k^{th}$ dividend claim is

$$1 = \exp[\kappa - (1 - \theta)(\rho_c - 1)A_0 + \kappa_k + (\rho_k - 1)A_{0,k} + (1 - \theta)A_k'y_t - A_k'y_t] \times E_t \left[ \exp \left\{ [-A_k' + S_k' + \rho_k A_k']y_{t+1} \right\} \right].$$

As before, using Lemma 3 to evaluate the expectation and taking logs yields the $(L + 1)$-dimensional system of equations

$$f(-\Lambda + S_k + \rho_k A_k) = -[\kappa - (1 - \theta)(\rho_c - 1)A_0 + \kappa_k + (\rho_k - 1)A_{0,k}]$$

(49)

$$g(-\Lambda + S_k + \rho_k A_k) = A_k - (1 - \theta)A.$$

(50)

Once again, I choose the linearization constants in order to linearize around the unconditional mean log price-dividend ratio. In order to obtain a more accurate solution, I allow the linearization constants $\kappa_k$ and $\rho_k$ to differ across assets. This amounts to replacing equation (49) with

$$f(-\Lambda + S_k + \rho_k A_k) = -[\kappa - (1 - \theta)(\rho_c - 1)A_0 - \log \rho_k - (\rho_k - 1)A_k'E(y_t)].$$

(51)

### B.4.3 Proof of Proposition 3 (Risk Premia)

Given my solution for the price-dividend ratio, the log-linearized market return is

$$r_{k,t+1} = \kappa_k + (\rho_k - 1)A_{0,k} + (S_k' + \rho_k A_k')y_{t+1} - A_k'y_t \equiv \kappa_k + (\rho_k - 1)A_{0,k} + B_k'y_{t+1} - A_k'y_t.$$
Following Drechsler and Yaron (2011), I decompose the projected pricing kernel and the return on a risky asset into jump and Gaussian components

\[
m_{t+1} = \kappa - (1 - \theta)(\rho c - 1)A_0 + (1 - \theta)A' y_{t+1} - \Lambda'(\mu y + F y y_{t+1} + G y, z_{y, t+1}) + -\Lambda' J_{y, t+1}
\]

\[
r_{k, t+1} = \kappa + (\rho c - 1)A_0k - A' k y_{t+1} + B' k(\mu y + F y y_{t+1} + G y, z_{y, t+1}) + B' J_{y, t+1}
\]

Drechsler and Yaron (2011) show that the risk premium may be decomposed as

\[
\log(E_t[R_{k, t+1}]) - r_{f, t+1} = -\text{cov}(m_{t+1}, r^g_{k, t+1}) + \log(E_t[\exp(r^J_{k, t+1})]) + \log(E_t[\exp(m_{t+1})]) - \log(E_t[\exp(r^J_{k, t+1} + m_{t+1})])
\]

\[
= B' G y, t G^t y, t A + \lambda^I y, t \Psi y(B_k - \Lambda) - \Psi y(-\Lambda)
\]

\[
\equiv B' G y, t G^t y, t A + \lambda^I y, t \Omega_k.
\]

See Drechsler and Yaron (2011), section A.4, for further details.

### B.4.4 Proof of Proposition 4 (ICAPM)

In this section, I derive an ICAPM representation. Many of the steps of the derivation follow Campbell et al. (2014), so I highlight the incremental effects of adding incomplete markets. From Proposition 2, the log of the average return on the consumption claim is an affine function of the state vector, \(y_t\). Following Campbell (1993), I substitute out \(\Delta c_{t+1}\) using the identity

\[
\Delta c_{t+1} \approx r_{c, t+1} - \kappa_c - \rho_c w_{c t+1} + w_{ct}.
\]

Plugging (53) into the log-linearized pricing kernel (12), one obtains

\[
m_{t, t+1} = \text{const.} - \gamma(r_{c, t+1} + \eta_{t+1}) + \frac{\theta}{\psi}(\rho_c w_{c t+1} - w_{ct}).
\]

Plugging (54) into the Euler equation for \(r_{i, t+1}\) and projecting out \(\eta_{i, t+1}\) yields

\[
1 = E_t \left[ \exp \left( \text{const.} + (1 - \gamma)(r_{c, t+1} + \nu^*_{t+1}) + \frac{\theta}{\psi}(\rho_c w_{c t+1} - w_{ct}) \right) \right],
\]

where \(\nu^*_{t+1} \equiv \frac{1}{1 - \gamma} \log E_{t+1}[\exp(1 - \gamma)\eta_{t+1}y_{t+1}]\). My distributional assumptions imply

\[
wc_t = \text{const.} + (\psi - 1)[E_t r_{c, t+1} + E_t \nu^*_{t+1}] + \rho_c E_t w_{c t+1} + \frac{1}{2} \psi^2 \theta_t,
\]
where \( \vartheta_t \) is a Jensen’s inequality term, defined as

\[
\vartheta_t \equiv 2 \log E_t \left[ \exp \left( (1 - \gamma)(r_{c,t+1} - E_t[r_{c,t+1}] + \nu^*_{t+1} - E_t[\nu^*_{t+1}]) + \frac{\theta}{\psi} \rho_c (wc_{t+1} - E_t[wc_{t+1}]) \right) \right].
\] (57)

In the absence of jump risk, \( \vartheta_t \) equals \( \text{Var}_t[m^i_{t+1} + r^i_{c,t+1}] \), i.e. the risk-neutral variance of the consumption claim. When jumps are present, there is an analogous term capturing Gaussian volatility and jump risk.

Iterating forward on (56) and assuming that \( \lim_{j \to \infty} \rho^j wc_{t+1} = 0 \), one obtains

\[
wc_t = \text{const.} + E_t \sum_{j=0}^{\infty} \rho^j [\psi - 1](r_{c,t+1+j} + \nu^*_{t+1+j}) + \frac{\psi}{2\gamma} \vartheta_{t+j}] \quad (58)
\]

\[
\rho_c [wc_{t+1} - E_t wc_{t+1}] = [E_{t+1} - E_t] \sum_{j=1}^{\infty} \rho^j [(\psi - 1)(r_{t+1+j} + \nu^*_{t+1+j}) + \frac{\psi}{2\gamma} \vartheta_{t+1}] \quad (59)
\]

\[
\equiv (\psi - 1)[N_{DR,t+1} + N_{FIR,t+1}] + \frac{\psi}{2\gamma} N_{UNC,t+1}, \quad (60)
\]

where discount rate news \( N_{DR,t+1} \) are also defined using the decomposition

\[
r_{c,t+1} - E_t r_{c,t+1} = [E_{t+1} - E_t] \sum_{j=0}^{\infty} \rho^j [\Delta c_{t+1+j} - \rho \cdot r_{c,t+2+j}] \equiv N_{CF,t+1} - N_{DR,t+1}. \quad (61)
\]

The key difference with respect to the representative agent model is the second term, \( N_{FIR,t+1} \).

The subscript \( FIR \) is shorthand for future idiosyncratic risk, which captures news about the higher moments of idiosyncratic shocks.\(^{84}\) Equation (60) says that, when the EIS is greater than 1, the wealth-consumption ratio is higher when agents get good news about the distribution of idiosyncratic risk, as summarized by the cross-sectional certainty equivalent \( \nu^*_{t+1} \). The third term \( N_{UNC,t+1} \) captures news about uncertainty, i.e. the higher moments of future aggregate shocks. Plugging (60-61) into the projected pricing kernel (13), (16) obtains.

In empirical implementations of the ICAPM, it is standard to assume that the VAR is configured such that \( r_{c,t+1} = S_r y_{t+1} \). Then, using a similar method from above, I can also solve for the wealth-consumption ratios, and, in turn, an affine expression for \( \vartheta_t \). The solution must satisfy

\[
f \left( (1 - \gamma)S_r + \frac{\theta}{\psi} \rho_c A + \beta(1 - \gamma) \right) + \beta_0(1 - \gamma) = -\theta \log \delta + \frac{\theta}{\psi} \log \rho_c + (\rho_c - 1)A'E(y_t) \quad (62)
\]

\[
g \left( (1 - \gamma)S_r + \frac{\theta}{\psi} \rho_c A + \beta(1 - \gamma) \right) = \frac{\theta}{\psi} A, \quad (63)
\]

\(^{84}\)This term is related but not identical to the indirect effect discussed in the previous section. It captures the intuition that agents may be willing to pay a premium to hedge against increases in idiosyncratic risk in future periods. However, idiosyncratic risk also affects the average return on consumption, so the appropriate definition of discount rate news may be different.
an \((L + 1)\)-dimensional system of equations in \(A\) and \(\rho_c\). Given \(A\) and \(\rho_c\), the expression for \(\vartheta_t\) in (57) is a known, affine function of \(y_t\).

B.4.5 Proof of Proposition 5 (Term Structure)

I will begin by establishing the result for the dividend claim, then proceed to the consumption claim. The proof is by induction. First, I establish that \(pd^0_{k,t} = A^0_{0,k} + (A^h_0)'y_t\). By our no arbitrage restriction that \(D_{k,t} = P^0_{k,t}\), implying that \(pd^0_{k,t} = 0\) and thus that \(A^0_{0,k} = 0\) and \(A^0_k = 0\). Next, I must show that \(pd^h_{k,t} = A^h_{0,k} + (A^h_k)'y_t\) implies \(pd^h_{k,t+1} = A^h_{0,k} + (A^h_k)'y_t\). Combining the Euler equation with the dividend strip return yields

\[
pd^{h+1}_{k,t} = \log E_t[\exp\{m_{t+1} + \Delta d_{k,t+1} + pd^h_{k,t+1}\}]
\]

\[
A^{h+1}_{0,k} + (A^h_k)'y_t = \log E_t[\exp\{m_{t+1} + A^h_{0,k} + (S_k + A^h_k)'y_{t+1}\}]
\]

\[
= \log E_t[\exp\{m_0 + (1 - \theta)A'y_t + A^h_{0,k} + (S_k + A^h_k - \Lambda)'y_{t+1}\}]
\]

\[
= m_0 + A^h_{0,k} + f(S_k + A^h_k - \Lambda) + g(S_k + A^h_k - \Lambda) + (1 - \theta)A'y_t
\]

where \(m_0 \equiv \kappa - (1 - \theta)(\rho_c - 1)A_0\). Matching coefficients yields the recursions

\[
A^{h+1}_{0,k} = m_0 + A^h_{0,k} + f(S_k + A^h_k - \Lambda)
\]

\[
A^{h+1}_k = g(S_k + A^h_k - \Lambda) + (1 - \theta)A,
\]

which establishes the claim. One obtains coefficients for the real risk-free asset, by setting \(S_k = 0\) in (64-65). Analogously, the coefficients for expected real dividend growth are obtained by setting \(m_0 = 0\) and \(\Lambda = 0\). Next, I turn to the consumption claim. The only substantive difference is that, since the return on the consumption claim depends on \(\eta_{t+1}^c\), I cannot use the projected version of the pricing kernel. Instead, I work with Euler equation directly to evaluate expectations. All other steps in the proof are the same. Again, no arbitrage requires that \(A^0_0 = 0\) and \(A^0 = 0\). Then, I show that \(wc^h_t = A^h_0 + (A^h)'y_t\) implies \(wc^{h+1}_t = A^{h+1}_0 + (A^{h+1})'y_t\):

\[
w^{h+1}_c = \log E_t[\exp\{m^i_{t+1} + \Delta c_{t+1} + \eta^i_{t+1} + wc^h_{t+1}\}]
\]

\[
A^{h+1}_0 + (A^h)'y_t = \theta \log \delta - (1 - \theta)(\rho_c - 1)A_0 + A^h_0 + (1 - \theta)A'y_t
\]

\[
+ \log E_t[\exp\{[(1 - \gamma)S_c - (1 - \theta)\rho_c A + A^h]'y_{t+1} + (1 - \gamma)\eta^i_{t+1}\}]
\]

\[
= \bar{m}_0 + A^h_0 + (1 - \theta)A'y_t
\]

\[
+ \log E_t[\exp\{[(1 - \gamma)S_c - (1 - \theta)\rho_c A + \beta(1 - \gamma) + A^h]'y_{t+1}\}]
\]

\[
= \bar{m}_0 + A^h_0 + f(S_c + A^h - \tilde{\Lambda}) + g(S_c + A^h - \tilde{\Lambda}) + (1 - \theta)A'y_t
\]

A.38
where \( \tilde{m}_0 \equiv \theta \log \delta - (1 - \theta)(\kappa_c + (\rho_c - 1)A_0) + \beta_0(1 - \gamma) \) and \( \tilde{\Lambda} \equiv \gamma S_c + (1 - \theta)\rho_c A + \beta(1 - \gamma) \). The third equality uses the law of iterated expectations and Lemma 2, and the fourth equality follows from Lemma 3. Matching coefficients yields the recursion

\[
A_{h+1}^0 = \tilde{m}_0 + A_{h,k}^0 + f(S_k + A_k^h - \tilde{\Lambda}) \quad (66)
\]

\[
A_{h+1}^h = g(S_k + A_k^h - \tilde{\Lambda}) + (1 - \theta)A, \quad (67)
\]

so the only substantive difference between (64-65) and (66-67) comes from the definitions of \( \tilde{m}_0 \) and \( \tilde{\Lambda} \). Further note that \( \tilde{m}_0 = m_0 - \beta_0(-\gamma) + \beta_0(1 - \gamma) \) and \( \tilde{\Lambda} = \Lambda - \beta(-\gamma) + \beta(1 - \gamma) \), so the difference between the recursions comes entirely from the projection terms.

**B.4.6 Risk-free rate**

From the Euler equation, the one-period risk-free rate satisfies

\[
rf_{t+1} = -\log \left( E_t[\exp(m_{t+1})] \right)
\]

\[
= -[\kappa - (1 - \theta)(\rho_c - 1)A_0 + (1 - \theta)A'y_t + \log \left( E_t[\exp(-\Lambda y_{t+1})] \right)]
\]

\[
= rf_0 - [g(-\Lambda) + (1 - \theta)A'y_t],
\]

where \( rf_0 \equiv -[\kappa - (1 - \theta)(\rho_c - 1)A_0 + f(-\Lambda)] \). Terms involving \( \eta \) drop out from the expression because of the conditional independence of \( r_{k,t+1} \) and \( \eta_{k+1} \).