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

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

Very preliminary – All comments welcome!

This version: August 2014
First version: July 2014

Abstract

Discrete events, particularly transitions between jobs, are thought to be one of the largest sources of wage dispersion. These switches often induce changes in wages that can be quite large, highly persistent, and largely uninsurable. Labor market transitions can be idiosyncratic tail events, potentially having a disproportionate impact on welfare without affecting aggregate quantities. The nature of these events is highly cyclical; transitions are more likely to be favorable if they occur during expansions relative to recessions. As such, agents may require a premium for investing in assets which underperform when labor market event risk is high, a feature absent from leading asset pricing models. This paper explores the potential importance of this channel in affecting asset prices. We provide empirical evidence on the plausibility of event risk in explaining the shape of the idiosyncratic distribution of income growth rates as well as the relationship between event risk and aggregate variables. Next, we formalize its role within a general affine, jump-diffusion asset pricing framework with heterogeneous agents and incomplete markets, making our results immediately applicable to a wide class of existing models for aggregate dynamics. We propose a model with a single state variable–capturing the probability of large, idiosyncratic decreases in consumption—that quantitatively matches the level and dynamics of the equity premium, even with i.i.d. aggregate consumption growth. Consistent with the model’s predictions, initial claims for unemployment, suitably normalized, is a highly robust predictor of returns, outperforming a number of conventional predictors, including the dividend yield.

*Department of Economics, UCSD, 9500 Gilman Drive, La Jolla CA 92039, 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 Brendan Beare, Jonathan Berk, Jules Van Binsbergen, John Campbell, Steven Davis, Martin Eichenbaum, Marjorie Flavin, James Hamilton, Simon Gilchrist, Fatih Guvenen, Bryan Kelly, Ralph Kojien, Arthur Korteweg, David Lagakos, Hanno Lustig, Serdar Ozkan, Chris Parsons, David Romer, Martin Schneider, Frank Schorfheide, Rossen Valkanov, Annette Vissing-Jorgensen, Russ Wermers, Johannes Wieland, Motohiro Yogo, and Josef Zechner for helpful discussions and comments. All errors are my own.
1 Introduction

Tail events can play an important role in determining asset prices despite their relative infrequency. One potential resolution to the equity premium puzzle, first suggested by Rietz (1988) and Barro (2006), is to incorporate rare disasters—a small probability of an extremely large drop in aggregate consumption. This additional risk exposure reconciles the high expected excess return on stocks with the relatively small covariance between stocks and aggregate consumption. Extensions which allow for a time-varying disaster probability and/or magnitude can reproduce salient features of levels and dynamics of risk premia. However, a primary critique is that the parameters of these models governing the probability and magnitudes of disasters are challenging to estimate given the length of the time series of available data.

This paper considers an economic environment where tail events are cross-sectional, rather than aggregate phenomena. Labor income is risky, virtually uninsurable, and quantitatively important—comprising the largest single component of aggregate consumption. A growing body of empirical literature suggests that workers face a substantial amount of idiosyncratic labor income risk. Much of this literature stresses the importance of labor market transitions in explaining the large amount of variability in labor income. One strand focuses on job displacement risk—large, highly persistent, and uninsurable declines in income which are often linked to the extensive margin. These events resemble rare but idiosyncratic disasters, since they need not be accompanied by large declines in aggregate consumption or output. Also important are situations where a worker voluntarily switches jobs, presumably because she receives a better outside offer. Both types of “idiosyncratic tail events” can result in large, persistent, and uninsurable changes in income. Thus, they can have disproportionate impacts on welfare even though their realizations only hit a small fraction of households each period.

When markets are incomplete, investors are willing to pay a premium to hedge against states where labor market event risk is high. This paper makes a case for state-dependent, idiosyncratic

---

2 See e.g. Julliard and Ghosh (2012).
3 Standard moral hazard arguments suggest that insurance markets for idiosyncratic labor productivity shocks are unlikely to function well.
4 Losing a job is often associated with both temporary and permanent losses in income. Unemployment insurance can act as a hedge against the former, but not the latter, type of risk. As such, the job displacement literature focuses on the long-term earnings losses that persist even after a worker has found a new job. See Krebs (2007) for a survey of the job displacement literature and a quantitative examination of its role in calculating the welfare costs of business cycles.
5 For example, Guvenen, Ozkan, and Song (2013), report that from 2007-2009, average annual wages declined by 6.5%. Over the same period, the median worker experienced essentially no change in income (an increase of 0.1%), while ten percent of workers suffered a decline of 60 (log) percentage points or more.
tail risk as a key driver of the dynamics of risk premia at business cycle frequencies. Agents strongly dislike recessions because, as the economy contracts, the left tail of the cross-sectional distribution of consumption growth becomes fatter and the right tail becomes thinner. Job displacement events loom larger and lucrative outside offers dry up. Since stock prices fall and uncertainty rises in recessions, stocks are a bad hedge against this source of risk, which can lead to a large and countercyclical equity premium.

We propose and test a number of implications of a model where predictable changes in idiosyncratic tail risk induce predictable changes in risk premia. First, we provide empirical evidence of state-dependence in the conditional tails of the cross-sectional distribution of labor income growth rates. Unlike aggregate tail events, for which the data are necessarily limited, idiosyncratic tail events occur every period. Our analysis builds upon statistics from Guvenen, Ozkan, and Song (2014b, hereafter “GOS”) which are calculated from panel administrative earnings data. While the center of the earnings growth distribution is relatively insensitive to the business cycle, the tails are highly responsive. Our estimates also suggest that changes in aggregate wages are far from equally spread across agents; instead, they appear to be primarily driven by changes in the tails of the cross-sectional distribution.

We develop a method which allows us to estimate the higher frequency (quarterly) dynamics of idiosyncratic risk from cross-sectional moments measured at lower frequencies. This procedure yields a quarterly index capturing the level of idiosyncratic risk at a point in time. We find that three macroeconomic variables, employment growth, aggregate wage growth, and the ratio of corporate profits to wages, are very good proxies for the shape of the cross-sectional distribution of income growth rates. These findings are quite intuitive; large positive (negative) wage changes are more (less) likely when wages are rising, firms are hiring, and profitability is high. Our idiosyncratic risk index is highly persistent and cyclical, and it exhibits substantial time series variation, even in periods without recessions.

Second, we 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. We solve the model with arbitrary, jump diffusion dynamics (stochastic volatility and time-varying, compound Poisson jumps) for aggregate cash flows. Symmetrically, idiosyncratic shocks are allowed to have state-dependent Gaussian and jump components. These idiosyncratic jump components provide an analytically tractable way of capturing infrequent, large changes in consumption. In the model, agents are ex-ante identical and ex-post heterogeneous, facing the same distribution
of idiosyncratic shocks at each point in time, which preserves the analytical tractability of a representative agent economy.\footnote{This assumption was introduced in an asset pricing context by Constantinides and Duffie (1996) and is also used in Krebs (2007, 2003), Constantinides (2002), Angeletos (2007), and Toda (2014a, 2014b).}

To complement to our solution for a fully-specified endowment-based model, we also derive an intertemporal capital asset pricing (ICAPM) representation of the stochastic discount factor from our incomplete markets economy, extending a recent contribution by Campbell et al. (2012). This representation reveals that, in addition to several risk factors which also appear in representative 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. The ICAPM model reveals that, when the EIS > 1 and idiosyncratic risk is fairly persistent, news about future idiosyncratic risk likely carries a substantially higher weight.

While we emphasize labor income throughout, our 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. In our 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.\footnote{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 our model, uncertainty has an additional, direct effect on preferences when it is liked with the distribution of idiosyncratic shocks. Herskovic et al. (2014) make a similar argument.}

Third, we test one of the key implications of our incomplete markets model, namely that the equity premium is high when labor market uncertainty is high. We demonstrate that initial claims for unemployment, an observable proxy for labor market uncertainty, is a powerful, highly robust predictor of broad market returns. Over the 1967-2012 sample—the period for which initial claims data are available—initial claims outperforms 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. Moreover, initial claims is an even stronger predictor of the excess return on the Fama and French (1993) small-minus-big portfolio.

Finally, we illustrate the quantitative importance of idiosyncratic tail risk in affecting the dynam-
ics of risk premia within a stylized model. In order to make as stark of a comparison as possible, we deliberately shut down almost all sources of aggregate risk. This model has a single state variable, capturing the probability experiencing a “personal disaster”, and i.i.d., homoskedastic aggregate consumption growth. Despite its simplicity, the model matches a number of key asset pricing moments quite well; the equity premium is high and time-varying, and stock returns are excessively volatile.

**Related Literature.** This paper lies at the intersection of literatures in finance, macroeconomics, and labor economics. Mankiw (1986) first suggested uninsurable risk as an early potential resolution to the equity premium puzzle. In representative agent models, aggregate and individual agents’ consumption move in lockstep, so a 1% decline in consumption is equally shared across agents. Mankiw’s (1986) model demonstrates that welfare and asset pricing implications may be different if such a decline is concentrated, ex post, among a small fraction of agents. Our theoretical model embeds such a concentration mechanism within a dynamic environment.

Constantinides and Duffie (1996) propose a tractable asset pricing model with uninsurable idiosyncratic risk, where the the variance of permanent, idiosyncratic shocks is state dependent. Given arbitrary aggregate consumption and return processes, they construct an idiosyncratic shock process which prices assets properly, leading them to conclude that the incomplete markets model places no testable restrictions on the joint behavior of aggregate consumption and returns. Storesletten et al. (2004) provide evidence that the volatility of persistent shocks to individuals’ wages is more volatile in contractions as compared with expansions, a phenomenon often referred to as countercyclical cross-sectional volatility (“CCSV”).

GOS, using a larger panel of income records from the Social Security Administration, find little evidence of state-dependent volatility and argue that it is the *skewness* of the persistent income shock distribution which varies over the business cycle. Their results are consistent with findings of Davis and Von Wachter (2011) on the earnings losses for workers who lost jobs in mass layoff events. Whereas Davis and Von Wachter (2011) emphasize earnings losses for relatively low skilled workers, GOS’ results demonstrate that essentially all workers are exposed to cyclical variation in skewness. Our analysis builds heavily on calculations in GOS, and we discuss their

---

8Beginning with Storesletten et al. (2004), studies which use income data from the PSID tend to find evidence of CCSV. See, e.g. Huggett and Kaplan (2013) for a recent example. This finding is not without controversy. For example, Krebs (2007) argues that this result goes away if one allows for a time trend in idiosyncratic volatility.

9Davis and Von Wachter (2011) and Krebs (2007) provide an excellent discussion of earlier empirical results on earnings losses associated with unemployment. Mckay and Papp (2012) present evidence from the PSID that the variance of idiosyncratic income shocks increases when the unemployment rate is high. They can generate such a result using a search-and-matching model with on-the-job search.
findings in greater detail in the next section of the paper. While we emphasize time-varying skewness throughout, our general theoretical model allows for CCSV, and the intuition for how CCSV affects risk premia is essentially identical to that discussed here.

Constantinides and Duffie (1996) and several related studies assume that agents have identical CRRA preferences, so any amplification coming from incomplete markets must arise from a contemporaneous correlation between the higher moments of the idiosyncratic shock distribution and returns.10 They also tend to be static, seeking to explain the level of the equity premium. Cochrane (2005, Ch. 21.2) surveys the early literature on asset pricing with incomplete markets and argues that the time variation in the variance of idiosyncratic shocks would have to be implausibly large in order to meaningfully affect the equity premium with moderate levels of risk aversion. Such a correlation is particularly unlikely to arise in the data given that stock returns tend to lead the business cycle, whereas labor markets tend to lag. In contrast with returns, labor market indicators tend to be highly persistent.

While the implications of the static incomplete markets model have been relatively well-explored, comparatively little work has explored dynamic implications. Toda (2014b) extends the incomplete markets model to a setting with recursive preferences and Markov dynamics.11 When agents have Epstein and Zin (1989) preferences and a preference for the early resolution of uncertainty, they are willing to pay a premium to hold assets whose returns hedge against bad news about the distribution of consumption growth in future periods.

Bansal and Yaron (2004) demonstrated that, with these preferences, one can generate a large equity premium if expected (aggregate) consumption growth has a highly persistent, mean-reverting component and stochastic volatility. This central idea has led to a large and rapidly growing family of long run risk models.12 Virtually all of these models assume that markets are complete; therefore, the only source of priced risk is aggregate consumption. Direct estimation can be quite challenging, as the key (highly persistent) state variables often must be filtered from aggregate consumption data, which exhibit little autocorrelation.13 A similar mechanism

---

10 See, e.g., Cogley (2002), Krebs (2003), Krebs (2007), and Storesletten et al. (2007)

11 Toda (2014b) establishes the existence and uniqueness of a competitive equilibrium when agents have general recursive preferences and access to a menu of linear investment technologies. These investments are subject to both aggregate and uninsurable, state-dependent idiosyncratic risk. Toda (2014a) shows how to embed this mechanism within a production setting. See also De Santis (2005).

12 See, e.g. Bansal et al. (2012) for a discussion of multiple extensions of the long-run risk model.

13 Bansal and Yaron (2004) write: “Shephard and Harvey (1990) show that in finite samples, it is very difficult to distinguish between a purely i.i.d. process and one which incorporates a small persistent component. While it is hard to distinguish econometrically between the two alternative processes, the asset pricing implications are very different.” For example, the calibration proposed in Bansal et al. (2012) includes a stochastic volatility process with a monthly autocorrelation of 0.999.
operates in models which generate large risk premia with persistent variation in the probability or severity of experiencing a macroeconomic disaster.

Our theoretical model also features long-run risks, and the intuition behind it is quite similar. Moreover, our assumptions on aggregate cash flow dynamics are sufficiently general to encompass the vast majority of models considered to-date. We allow for additional dimensions of priced risk which do not affect the distribution of aggregate consumption. If, as the data suggest, extensive margin measures are good proxies for the higher moments of the distribution of aggregate shocks, then these state variables need not be inferred from the data. Quantitatively speaking, these measures are highly persistent and highly correlated with valuation ratios. Beeler and Campbell (2012) argue that the long-run risk model makes counterfactual predictions about the predicability of aggregate consumption by the price-dividend ratio. We demonstrate that the price-dividend ratio has substantial predictive content for our idiosyncratic risk index.

While we work within a reduced-form framework, theoretical motivations for cyclical variation in labor market event risk are plentiful. Berk et al. (2010) and Lagakos and Ordoñez (2011) provide examples where the optimal contract between workers and firms involves partial insurance; wages optimally move less than one-for-one in response to productivity shocks. In Berk et al. (2010), financial distress can cause this insurance to break down, disappearing completely if the contract is terminated. Ex post, losses are highly concentrated among workers who switch jobs. If all firms are more likely to be distressed in some states than others, the risk of large losses is time-varying. Moreover, firms’ provision of partial insurance will imply that output declines by more than wages in these bad states, causing firm profits to fall precisely when labor market event risk is high.

In the search-and-matching literature, which is too voluminous to survey here, virtually all of the variation in workers’ labor income occurs when workers switch between jobs. Thus, any changes in aggregate quantities from are fully concentrated among those who switch jobs. We also present evidence of pro cyclical variation in the likelihood of receiving large, positive shocks. On-the-job search models emphasize workers’ real option to increase future wages via future outside offers. The exercise of these options is more lucrative in good times, fattening the

\[ \text{Autocorrelation of initial claims for unemployment} = 0.986.\]

\[ \text{Berk and Walden (2012) present a model with two types of agents in which labor market contracts enable one type of agents to perfectly share idiosyncratic risks, concentrating all aggregate risk with the second group.}\]

\[ \text{The intuition from Shleifer and Vishny (1992) likely applies to the market for skilled labor (i.e. high-earners, who are more likely to participate in financial markets). When a worker with substantial industry-specific knowledge switches jobs, she will be most productive if she stays within the same industry. However, if the switch occurs because her firm encounters financial distress, other firms in the same industry (who are best positioned to productively use this knowledge) may also be constrained, lowering her wages at the next job.}\]
right tail of the income growth distribution. Thus, our reduced-form specification provides an approximate, but tractable way of thinking about potential asset pricing implications of search frictions in incomplete markets.

Kuehn et al. (2013) and Hall (2014) both discuss interactions between labor market search and asset prices. Kuehn et al. (2013) propose a production-based asset pricing model with search frictions and show that these frictions can endogenously generate rare disasters in the aggregate, which helps to generate a large, time-varying equity premium. Unlike our framework, they assume that a representative agent can pool income from employed and unemployed workers. Hall (2014) makes the broader argument that rises in discount rates should be associated with increases in unemployment, because they affect firms’ incentives for creating new jobs. Both papers provide evidence that the equity premium is forecastable using observable proxies of market tightness, the key state variable in search-and-matching models.

Two recent working papers, contemporaneous with our own, also study asset pricing implications of time-varying, state-dependent risk. Both provide empirical analyses in support of the mechanism and calibrate theoretical models, both of which are special cases of our general 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 construct a stylized model where aggregate consumption and dividend growth are i.i.d but the higher moments of household consumption growth are persistent and show that it is capable of matching key asset pricing moments. We provide a detailed discussion comparison between our quantitative model and theirs in Section 7. They present qualitative evidence that household skewness measures are priced in the cross-section, though the estimated risk prices are statistically insignificant.

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 show that that this factor is correlated with measures of household income risk. Their leading measure is the change in the dispersion (measured as the difference between the 90th and 10th percentiles) of year-on-year income growth rates from GOS. Finally, they show that they can replicate these cross-sectional patterns in a model where the volatility of idiosyncratic consumption growth shocks is persistent and countercyclical.

---

17 Our results suggest, therefore, that potential feedback from search fractions could be substantially increased when households are unable to insure against job loss.

18 These empirical findings compliment ours, 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. Also, survey-based consumption measures are more susceptible to measurement errors relative to administrative earnings data.
A key difference between these two papers and ours is the way in which idiosyncratic risk is modeled. We allow idiosyncratic shocks to be generated via an affine jump diffusion process. Herskovic et al. (2014) emphasize changes in second moments of idiosyncratic shocks; therefore, their analysis is focused on the implications of persistent variation in CCSV. In the income process of Constantinides and Ghosh (2014), the only source of the asymmetry is a small adjustment due to Jensen’s inequality. Their idiosyncratic risk specification generates very little skewness, and persistent variation in even moments is the primary source of time-varying risk premia in their model. Our process can generate arbitrarily high levels of skewness.

Moreover, our general solution clarifies the theoretical mechanisms at play. We show that the asset pricing implications of time-varying idiosyncratic risk can be summarized by the time series behavior of 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. When the idiosyncratic risk distribution distribution is driven by a single state variable, we can infer the time series dynamics of the certainty equivalent directly from the cross-sectional moments of the data. This makes it possible to test qualitative predictions of the model without needing to impose a specific functional form on the DGP of idiosyncratic shocks.

Relative to these papers, we also allow for a much richer specification of aggregate dynamics, which enables us to study the interactions between aggregate and idiosyncratic risk factors. To clarify the mechanics associated with time-varying idiosyncratic risk, we downplay these dynamics in our quantitative model. However, such an extension is likely to be useful going forward given the tight link between the first moment and the tails of the cross-sectional income growth distribution.

The remainder of the paper is organized as follows. We begin with some motivating evidence on state-dependent tail risk and its evolution over time. Section 2 presents some nonparametric results, while Section 3 describes the evolution of idiosyncratic risk over time. Sections 4 and 5 present and solves our general theoretical asset pricing model. Section 6 introduces our proxy for labor market uncertainty and tests the key implication of our model for return predictability. Section 7 presents our stylized, quantitative model, and Section 8 concludes.

## 2 Motivating Evidence

In this section, we briefly review several key implications of statistics calculated by GOS, whose calculations are used as the basis for our analysis. GOS obtain 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 (163 million total observations). They calculate a number of statistics for the cross-section of income growth rates, and demonstrate how these distributions evolve over time. Their data cover the period from 1978-2011.

GOS define \( \bar{y}_{i\tau} \) as the real wage of individual \( i \) in year \( \tau \). They then calculate \( y_{i\tau} = \bar{y}_{i\tau} - g(x_{i\tau}) - \lambda_{\tau} \), where \( g(x_{i\tau}) \) is a life-cycle component. GOS report statistics for the cross-section of standardized income growth rates \( y_{i\tau} - y_{i\tau-h} \) for various values of \( h \). The key, highly robust result from their analysis is that, in their dataset, the variance of idiosyncratic shocks is almost acyclical, while there is quantitatively important cyclical variation in skewness/asymmetry of the distribution. This finding stands in stark contrast to earlier work by Storesletten et al. (2004), among others, which suggested that the variance of idiosyncratic shocks is countercyclical.

One obtains an intuitive measure of the asymmetry of a distribution by considering three conditional quantiles. GOS report the evolution over time of the 10\(^{th} \), 50\(^{th} \), and 90\(^{th} \) percentiles of the cross-section of income growth rates. A robust measure of skewness is Kelley’s skewness, which is defined as \[
\frac{(Q_{90} - Q_{50}) - (Q_{50} - Q_{10})}{Q_{90} - Q_{10}}.
\] \(^{19}\) The denominator is the distance between the 10\(^{th} \) and 90\(^{th} \) percentiles, a measure of the overall spread of the distribution. GOS show that, over longer horizons, this distance is almost constant. The numerator takes this distance and splits it into two pieces, so the overall measure is between -1 and 1. The first, \( Q_{90} - Q_{50} \), is a measure of 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.\(^ {20}\)

Figure 1 shows the time series evolution of these spreads, for 1, 3, and 5 year trailing changes in wages, respectively, i.e. \( y_{i,h\tau} - y_{i,\tau-h} \) for \( h = 1, 3, 5 \). 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. The graph demonstrates that, as the economy moves from an expansion to a recession, the left tail of the distribution (\( Q_{50} - Q_{10} \), plotted in Panel A) expands, indicating an increased likelihood of experiencing large decreases in wages, while the width of the right tail (\( Q_{90} - Q_{50} \), plotted in Panel B) shrinks. There are more big losers in recessions and fewer big winners. Note that our use of trailing growth rates in the graphs means that the long horizon measures will tend to lag the recession bars.

\(^{19}\) An alternative name for this measure, which is more popular in the finance literature, is “conditional asymmetry”. See, e.g., Ghysels et al. (2013).

\(^{20}\) For example, a Kelly’s skewness of -20% implies that the left tail makes up 60% of the spread between the 10\(^{th} \) and 90\(^{th} \) percentiles, while the right tail contributes the remaining 40% of the distance.
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 1: 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 $y_{i,h\tau} - y_{i,\tau}$ for $h = 1, 3, 5$ 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. $\tau$ is on the horizontal axis. Data are from GOS Appendix Table A.13, which reports linearly detrended cross-sectional quantiles.

Table 1 summarizes a number of GOS’s results on the distribution of 5-year wage changes, which control for individuals’ previous earnings.\footnote{A potential critique of Figure 1 is that the reported statistics pool the entire population of male earners together, which could overstate the asymmetry of the distribution of idiosyncratic shocks. Consider a simple example where the only impact of the business cycle on wages is through the mean (a pure location shift). If there is heterogeneity in the cyclicality of average earnings across groups, pooling individuals will generate asymmetry in the cross-sectional distribution. Therefore, it is important to control for other observable characteristics as well, particularly lagged earnings. GOS control for lagged 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...} For each year in 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 years and recession years, 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.\footnote{GOS define income growth rates starting in 1979, 1989, 1999, and 2006 as those which include recessions. Expansions average over income growth rates starting in 1983, 1984, 1993, 1994, and 2002.} Expansions are 5-year periods which do not include a recession year.

Given our interest in asset pricing, we focus on the evolution of idiosyncratic risk faced by relatively high earners over time—i.e. those who are likely to participate in financial markets. We summarize this information by averaging these statistics over different ranges of the income distribution. These ranges are indicated by different columns, where [96,100] in the first column indicates that we are describing the risk faced by the top 5% of earners. Thus, row one, column three of Table 1 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.

The top section of Table 1 reports the median, 10\textsuperscript{th} percentile, 90\textsuperscript{th} percentile, and Kelley’s skewness of five year income growth rates in expansions and recessions, respectively. The middle section reports the changes in quantile-based measures of scale—the inter-quartile range and the 90 - 10 percentile spread—in recessions versus expansions. The last two sections report quantile-based 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.

Regardless of the specific group (column of Table 1) considered, a consistent picture emerges. First, high earners face a substantial degree of idiosyncratic labor income risk, even in expansions. The average 90-10 spread is in excess of 100 log percentage points for all of the groups. Second, the entire distribution shifts to the left; all of the quantiles are strictly lower in recessions relative to expansions. This shift is not specific to the three quantiles in the top panel. All of the quantiles considered in our analysis are lower in recessions.

Third, the change in the 10\textsuperscript{th} and 90\textsuperscript{th} percentiles is larger than the change in the median, meaning that width of the left tail expands in recessions, while the right tail shrinks. This cyclical asymmetry is reflected by the change in Kelley’s skewness, which decreases by 13-20 percentage points (depending on the group) with the largest change occurring in the group of controlling for age. We refer the reader to GOS for further details on this procedure.
Table 1: 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 1: Summary statistics for the cross sectional distribution of income growth rates

<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></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>2.71</td>
<td>2.12</td>
</tr>
<tr>
<td>R</td>
<td>-4.46</td>
<td>-3.44</td>
</tr>
<tr>
<td>10th Percentile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>-76.48</td>
<td>-69.55</td>
</tr>
<tr>
<td>R</td>
<td>-95.34</td>
<td>-84.79</td>
</tr>
<tr>
<td>90th Percentile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>76.63</td>
<td>64.91</td>
</tr>
<tr>
<td>R</td>
<td>53.51</td>
<td>47.23</td>
</tr>
<tr>
<td>R - E</td>
<td>-20.47</td>
<td>-16.20</td>
</tr>
<tr>
<td>Kelley’s Skewness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>-3.22</td>
<td>-7.03</td>
</tr>
<tr>
<td>R</td>
<td>-22.30</td>
<td>-23.52</td>
</tr>
<tr>
<td>Scale Measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inter-Quartile Range</td>
<td>R - E</td>
<td>0.43</td>
</tr>
<tr>
<td>90-10 Percentile Spread</td>
<td>R - E</td>
<td>-4.26</td>
</tr>
<tr>
<td>Left Tail Width Measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50-25 Percentile Spread</td>
<td>R - E</td>
<td>4.08</td>
</tr>
<tr>
<td>50-10 Percentile Spread</td>
<td>R - E</td>
<td>10.85</td>
</tr>
<tr>
<td>Right Tail Width Measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td>75-50 Percentile Spread</td>
<td>R - E</td>
<td>-3.66</td>
</tr>
<tr>
<td>90-50 Percentile Spread</td>
<td>R - E</td>
<td>-12.97</td>
</tr>
</tbody>
</table>

highest earners. In contrast, both measures of the overall spread of the distribution changes very little over the cycle. GOS demonstrate that the 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. Consider the [91,95] column of Table 1. 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, we 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 1 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 3.2 replicates these features with a simple model where labor income is exposed to infrequent but very large shocks whose distribution is state-dependent.

3 The Evolution of Idiosyncratic Risk Over Time

3.1 Identifying proxies for conditional skewness

GOS emphasize changes in skewness in recession years relative to expansion years. 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. In this section, we bridge this gap by identifying several macroeconomic variables which are good proxies for the skewness of the idiosyncratic wage distribution. These proxies are available at higher frequencies and are available over a longer time period, which improves the power of asset pricing tests.

3.1.1 Statistical framework

Our reduced-form model for labor income is a version of the canonical permanent income life-cycle model. Let $y_{it}$ and $x_{it}$ be an individual’s log labor income (after subtracting common
shocks and a deterministic life-cycle component) and age in period $t$, respectively. We assume

$$y_{it} = \alpha_i + \beta_i x_{it} + \xi_{it} + \rho(L) \cdot \eta_{it} + \epsilon_{it}$$  \quad (1)

where $\eta_{it}$ is a shock that is independently distributed over time conditional on the aggregate state, which we assume can be characterized by a finite-dimensional random vector, $z_t$. $\alpha_i$ and $\beta_i$ allow for heterogeneity in income levels and growth rates and are randomly drawn from a time-invariant bivariate distribution with mean zero and finite third moments. $\xi_{it}$ is the permanent component to wages. As we discuss below, our theoretical model requires that $\xi_{it}$ be a random walk process. Alternatively, one could allow for $\xi_{it}$ to follow a persistent, stationary process such as an AR(1). We impose this restriction throughout, noting that empirical estimates of the AR(1) parameter in other studies are generally close to 1.\textsuperscript{23}

We also allow for a transitory component in labor income. The first term captures its conditional mean, which can depend on current and past $\eta_{it}$ via $\rho(L)$, allowing permanent shocks to have additional temporary effects on measured income. For example, large negative realizations of $\eta_t$ could be accompanied by unemployment spells, leading to temporary interruptions in the flow of labor income.\textsuperscript{24} The second term, $\epsilon_{it}$, 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 $\epsilon_{it}$, we maintain this assumption for parsimony.\textsuperscript{25}

Asset pricing tests tend to be conducted using high frequency (e.g. monthly or quarterly) data, but wage data are available on an annual basis. We use macroeconomic time series which are sampled at a quarterly frequency in our analysis, then we derive the implications for the higher moments of idiosyncratic wage changes, where wages are measured annually. To do so, we make use of a simple log-linear approximation for time-aggregated wages. For notational simplicity, we suppress $i$ subscripts here.

Define $Y_{A,t} = \sum_{j=0}^{3} \exp(y_{t-j})$ and $y_{A,t} \equiv \log Y_{A,t}$, so that $Y_{A,t}$ is a four quarter moving average of labor income, measured at the end of quarter $t$. $Y_{A,t}$ is observed when the index $t$ corresponds

\textsuperscript{23}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.

\textsuperscript{24}Signing bonuses could generate similar effects for large positive shocks.

\textsuperscript{25}Our aggregation result goes through if, given the aggregate state, the third central moment of $\epsilon_{it}$ is constant.
with the fourth quarter. A first-order Taylor expansion yields that, for \( k \geq 4 \),

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

\[
= \beta \cdot k + \frac{1}{4} \eta_t + \frac{1}{2} \eta_{t-1} + \frac{3}{4} \eta_{t-2} + \sum_{j=3}^{k-1} \eta_{t-j} + \frac{3}{4} \eta_{t-k} + \frac{1}{2} \eta_{t-k-1} + \frac{1}{4} \eta_{t-k-2}
\]

\[
+ \frac{1}{4} \rho(L)(\eta_t + \eta_{t-1} + \eta_{t-2} + \eta_{t-3}) - \frac{1}{4} \rho(L)(\eta_{t-k} + \eta_{t-k-1} + \eta_{t-k-2} + \eta_{t-k-3})
\]

\[
+ \frac{1}{4}(\epsilon_{t-k} + \epsilon_{t-k-1} + \epsilon_{t-k-2} + \epsilon_{t-k-3})
\]

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

where \( \theta_k(\cdot) \) is a polynomial in the lag operator whose second argument is the vector of coefficients for \( \rho(L) \). This approximation replaces an arithmetic mean with a geometric mean, a solution proposed in a related context by Mariano and Murasawa (2003).\(^{26}\) Henceforth, we ignore approximation errors, treating (3) as the “true model” for time-aggregated labor income measures. Simulation results in the Appendix demonstrate that these errors are negligible for the parametric model in Section 3.2, particularly over longer horizons.

Next, we 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

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

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

where \( M_3(\cdot) \) denotes the third central moment, conditional on aggregate information.\(^{27}\) If we further assume that \( M_3(\eta_t) = a + b' z_t \), where \( z_t \) is a vector of observable state variables, then

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

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

---

\(^{26}\)See also Camacho and Perez-Quiros (2010) for a similar approach.

\(^{27}\)This follows because, given two independent random variables \( x \) and \( y \) with \( \mu_x \equiv E[x] \) and \( \mu_y \equiv E[y] \),

\[
E[(x + y - \mu_x - \mu_y)^3] = E[(x - \mu_x)^3] + 3E[(x - \mu_x)^2(y - \mu_y)] + 3E[(x - \mu_x)(y - \mu_y)^2] + E[(y - \mu_y)^3]
\]

\[
= E[(x - \mu_x)^3] + E[(y - \mu_y)^3] \equiv M_3[x] + M_3[y],
\]

where we use independence to replace terms such as \( E[(x - \mu_x)^2(y - \mu_y)] \) with \( E[(x - \mu_x)^2] E[(y - \mu_y)] = 0 \).
is constant given our assumption that the third moments of $\beta_i$ and $\epsilon_t$ are state independent.

Equation (5) says that, when the third central moment of the permanent shock $\eta_t$ is affine in an observable vector $z_t$, under our assumptions on the income process, we can recover $b$ semi-parametrically with a regression. Given a time series of time-aggregated third-central moments and a model for $\rho(L)$, $b$ is the vector of slope coefficients from a regression of $M_3[y_{A,t} - y_{A,t-k}]$ on a constant and $\phi_k(L; \rho)z_t$. We can pool the information from time-aggregated moments which are measured over different horizons, improving efficiency by imposing the cross-equation restriction that $b$ is the same at all horizons. Relative to $b$, the unconditional level of skewness $a$ is harder to identify, because it interacts with the third central moments of profile heterogeneity and transitory shocks to determine the constant $c_k$.

Is it plausible to assume $M_3(\eta_t) = a + b'z_t$? A sufficient condition is that the cumulant-generating function (the log of the moment-generating function) of $\eta_t$ is linear in $z_t$. Most distributions used in theoretical asset pricing models satisfy this condition, since linear cumulant-generating functions often lead to exponential affine solutions for prices, facilitating analytical tractability.28 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. Assuming independence, linear combinations of these processes also satisfy this property. For example, if $\eta_t$ has a compound Poisson distribution with arrival intensity $\lambda_t = \lambda_0 + \lambda_1'z_t$, our regression recovers $\lambda_1$ up to a constant of proportionality. The next section discusses a number of key properties of this distribution, which provides an analytically tractable way to represent infrequent, large shocks (“jump risk”).

3.1.2 Explanatory variables

We consider four macroeconomic time series for inclusion in the vector $z_t$. The first variable, $\Delta emp_t$, is the quarterly change in the logarithm of private payroll employment. In the previous section, we 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, we 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

---

28 Otherwise, one could still motivate our specification using standard linear projection arguments.
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.\(^29\) 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, $pw_{t-1}$ is 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 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.\(^30\)

Berk et al. (2010) is a partial equilibrium model, lacking any sources of aggregate risk. However, if we take the structure of their optimal contract as given and apply 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.\(^31\)

\(^29\) We 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, we 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.

\(^30\) Berk 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.”

\(^31\) Giving 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 2: 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 (4).

The last variable, $\Delta c_t$, is the change in the logarithm of real aggregate consumption (nondurables plus services). The intuition for aggregate consumption is essentially identical to that for $\Delta y_t$. We would expect the distribution of income growth rates to be more negatively skewed when household consumption is falling. However, compared with $\Delta y_t$, there is more scope for a timing mismatch between $\Delta c_t$ 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 $\Delta c_t$ to lead the cross-sectional moments.
3.1.3 Results

As a precursor to our regressions, Figure 2 summarizes the univariate forecasting performance of the employment and compensation growth, perhaps two of the most natural candidates for $z_t$. 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, $\phi_k(L; 0)\Delta emp_t$ and $\phi_k(L; 0)\Delta y_t$, respectively. For purposes of generating these graphs, we calculate the moving averages assuming that $\rho(L) = 0$, which assumes that transitory shocks are completely state-independent. As we discuss in greater detail below, changing $\rho(\cdot)$ primarily impacts the level of $\phi_k(L; 0)z_t$ rather than its time series variation. Similar results obtain with other choices of $\rho(\cdot)$.

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 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 2 estimates the vector $b$ in (5) by regressing the time-aggregated skewness measures on several aggregate variables. Given the sample size, we limit attention to univariate and bivariate specifications. Panel A sets $\rho(L) = 0$, while Panel B allows for a restricted MA(1) structure: $\rho(L) = \rho \cdot [1 + L]$. In this latter specification, the partial derivative of $y_{it}$ 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.

Each panel includes the coefficients from three different estimations. In the left columns, we 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. 

---

32 See Table 2, Panel A.
33 Similar results obtain with different lag lengths.
34 We calculate standard which are robust to the presence of heteroskedasticity and autocorrelation. We use a Newey-West estimator for the long-run variance with 4 lags.
### Table 2: Regressions of third central moment of income growth on aggregate variables

This table presents the results from estimating equation (5) 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.

<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>Coefficient</td>
<td>( \rho )</td>
<td>( R^2 )</td>
</tr>
<tr>
<td>1</td>
<td>( \Delta \text{emp}_t )</td>
<td>1.7959***</td>
<td>0</td>
<td>0.599</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.415)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>( \Delta y_t )</td>
<td>1.1325***</td>
<td>0</td>
<td>0.708</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.143)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>( p w_{t-1} )</td>
<td>0.0438***</td>
<td>0</td>
<td>0.610</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.006)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>( \Delta c_t )</td>
<td>1.5459***</td>
<td>0</td>
<td>0.233</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.383)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>( \Delta \text{emp}_t )</td>
<td>1.2245***</td>
<td>0</td>
<td>0.766</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.236)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>( \Delta y_t )</td>
<td>0.8456***</td>
<td>0</td>
<td>0.808</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.129)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>( \Delta c_t )</td>
<td>0.0258***</td>
<td>-1.3590***</td>
<td>0.614</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.007)</td>
<td>(0.285)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \Delta y_t )</td>
<td>1.8210***</td>
<td>0</td>
<td>0.853</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.123)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \Delta c_t )</td>
<td>-1.9281***</td>
<td>-1.8194***</td>
<td>-1.4021</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.230)</td>
<td>(0.249)</td>
<td></td>
</tr>
</tbody>
</table>

<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>Coefficient</td>
<td>( \rho )</td>
<td>( R^2 )</td>
</tr>
<tr>
<td>1</td>
<td>( \Delta \text{emp}_t )</td>
<td>1.5676***</td>
<td>0.5763</td>
<td>0.545</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.473)</td>
<td>(0.371)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>( \Delta y_t )</td>
<td>0.9565***</td>
<td>0.6092**</td>
<td>0.676</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.160)</td>
<td>(0.237)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>( p w_{t-1} )</td>
<td>0.0433***</td>
<td>-1.3590***</td>
<td>0.614</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.007)</td>
<td>(0.285)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>( \Delta c_t )</td>
<td>1.2088***</td>
<td>0.8957*</td>
<td>0.239</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.376)</td>
<td>(0.448)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>( \Delta \text{emp}_t )</td>
<td>1.1851***</td>
<td>0.4639</td>
<td>0.702</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.261)</td>
<td>(0.313)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>( \Delta y_t )</td>
<td>0.7952***</td>
<td>0.5240**</td>
<td>0.752</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.125)</td>
<td>(0.243)</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>( \Delta c_t )</td>
<td>1.5525***</td>
<td>0.4392**</td>
<td>0.810</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.140)</td>
<td>(0.217)</td>
<td></td>
</tr>
</tbody>
</table>

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.
measures only. When estimating these univariate regressions in Panel B, we fix the value of $\rho$ at its estimated value from the bivariate model.\textsuperscript{35}

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 we 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 3 plots our estimates of quarterly conditional third central moments, $\hat{b}'z_t$, from the pooled GMM estimates of models 5-7 from Panel B of Table 2. 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 our 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 inferior performance of Model 4 is somewhat unsurprising in light of our 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 middle 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 $k$. Our pooled GMM estimations im-

\textsuperscript{35} Thus, the associated standard errors are best interpreted as conditional on $\hat{\rho}$. 

22
### Figure 3: 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 2, 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 (5) 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.

Pose this restriction, while the univariate regressions in the middle and right columns allow us to test it. Therefore, we 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 we 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 3 offer an explanation for such a result. Panel B plots the coefficients of the lag polynomial $\phi_4(L; \rho)$ in the regression equation (5) for 1 year skewness measures. First, we plot the weights when $\rho(L) = 0$. Second, we 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 3, 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 we are 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 2, 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 our 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 $\Delta emp_t$ 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 our specification of $\rho(L)$. However, we are more concerned with the model’s performance at longer frequencies, which is quite strong.
3.2 Estimating a parametric model with labor market event risk

The analysis in the previous section was semi-parametric, allowing us to make statements about the time-series behavior of idiosyncratic skewness without any distributional assumptions. In order to close a quantitative asset pricing model, such assumptions will be necessary. This section adopts a more parametric approach. We fit a simple model with labor market event risk that simultaneously matches the cyclical variation in cross-sectional skewness from Section 2 and the time-series dynamics emphasized in Section 3.1 quite well.

We maintain our assumptions on the labor income process from the previous section. Then, from Equation (3), the growth rate of annual labor income 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, we adopt the fairly standard assumptions that \(\beta \sim N(0, \sigma_b^2)\) and \(\epsilon_{A,t} \sim N(0, \sigma_{\epsilon}^2)\).

Our primary interest is on the last term, a weighted moving average of permanent shocks, \(\eta_t\). We assume that \(\eta_t = (J_{g,t} - E_t[J_{g,t}]) + (J_{b,t} - E_t[J_{b,t}]) + N(0, \sigma_n^2)\). \(J_{g,t}\) and \(J_{b,t}\) are compound Poisson random variables with time-varying intensities and exponential increments, defined as

\[
J_{g,t} = \sum_{j=1}^{N_{g,t}} [\mu_s + \text{Exponential}(\sigma_s) - \sigma_s], \quad N_{g,t} \sim \text{Possion}(\lambda_{0g} + \lambda_1' z_t) \tag{6}
\]

\[
J_{b,t} = \sum_{j=1}^{N_{b,t}} [-\mu_s + \sigma_s - \text{Exponential}(\sigma_s)], \quad N_{b,t} \sim \text{Possion}(\lambda_{0b} - \lambda_1' z_t), \tag{7}
\]

where \(J_{g,t} = 0\) and \(J_{b,t} = 0\) when \(N_{g,t} = 0\) and \(N_{b,t} = 0\), respectively. The first component, \(J_{g,t}\), is a good shock, which captures infrequent, large upward adjustments in consumption—“climbing the ladder”. For example, these changes could come as a result of a promotion or the arrival of an attractive outside offer. The second component, \(J_{b,t}\), is a bad shock, capturing large downward adjustments—“falling off the ladder”—driven by events such as job loss. We also allow for a normally-distributed state-independent neutral shock which hits every period.

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 \(\lambda_{0g} + \lambda_1' z_t\) and \(\lambda_{0b} - \lambda_1' z_t\) 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 \(\lambda_1' z_t\). When \(\lambda_1' z_t\) 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, we assume that the jump size distribution for good shocks equals that for the bad shocks multiplied by negative 1. One can also show that, when the sum of the Poisson intensities is constant, this restriction implies that the even moments of $\eta_t$ are constant. By construction, our estimates are consistent with the evidence on the lack of cyclical variation in second moments from GOS. By allowing $\lambda_{0,g}$ to differ from $\lambda_{0,b}$, this process can generate substantial unconditional skewness.

Our objective is to quantify the labor income risk of a “representative” stockholder. We will try to match a number of statistics from the [91,95] column of Table 1. 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.\textsuperscript{36} We estimate the parameters so as to minimize the sum of squared errors between a number of model-implied and data-implied moments from Table 1. First, we try to match the average distance between the median and the 10\textsuperscript{th} percentile of 5-year income growth rates in expansions and recessions. We also target the average distance between the 90\textsuperscript{th} percentile and the median in expansions and recessions. Second, we target the changes from recessions versus expansions in the left tail and right tail width measures. Third, we target the average distance between the 90\textsuperscript{th} and 10\textsuperscript{th} percentiles of 1-year growth rates in expansions.

For $z_t$, we use the quarterly skewness estimate from the pooled GMM estimate of Model 5 in Panel B of Table 2. While this model did not have the best performance, as gauged by goodness-of-fit statistics, the resulting linear index is less noisy than those obtained from the other models. Figure 3 shows this variable does a good job of picking up the lower-frequency variation in the other measures. Given that poisson intensities must satisfy a nonnegativity constraint, this improves the ability of the model to fit the data while satisfying this constraint.\textsuperscript{37} We normalize it to have mean zero and variance 1 to aid in interpreting $\lambda_{0,g}$, $\lambda_{0,b}$, and $\lambda_1$. $\lambda_{0,g}$ and $\lambda_{0,b}$ capture the unconditional probabilities of experiencing good and bad shocks, while $\lambda_1$ captures the marginal effect of a 1 standard deviation increase in $z_t$ on the probability of a good shock.

We place two additional restrictions on the model. Under our assumptions, the distribution\textsuperscript{36}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 cyclical of the incomes of extremely high earners, particularly those in the top 1%. See also Guvenen, Kaplan, and Song (2014a). While the higher cyclicity 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.\textsuperscript{37}In the estimation, we allow for the fitted intensities to go slightly negative (we place a constraint bounding them below at -0.0025). When calculating conditional quantiles, we truncate $\lambda_1 z_t$ so that the minimum intensity is zero the sum of the fitted intensities is $\lambda_{0,b} + \lambda_{0,g}$
of \( y_{A,t} - y_{A,t-k} \) can be decomposed into the sum of a gaussian component and a non-gaussian component. The variance of the gaussian component depends on \( \sigma^2_\beta, \sigma^2_\epsilon, \) and \( \sigma^2_n \). Given that data are only available two different horizons (1-year and 5-year), we need an additional restriction to achieve identification. We choose to shut off profile heterogeneity by setting \( \sigma^2_\beta = 0 \), which is relatively innocuous given our focus on state-dependent shocks. Given the relative insensitivity of 5-year moments to \( \rho(L) \), we adopt the simplest specification of \( \rho(L) = 0 \).

We are 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 we collect in a vector \( \beta \), we can calculate its characteristic function, \( \varphi_{k,t}(\omega; \beta) \equiv E_t[\exp\{i\omega \cdot (y_{A,t} - y_{A,t-k})\}] \) analytically. Using Lévy’s (1925) theorem, we recover the probability density function \( f_{k,t}(x; \beta) \) by taking the inverse Fourier transform of \( \varphi_{k,t}(\omega; \beta) \),

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

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

Table 3 presents estimates of the parameters governing the labor income process. \( \lambda_0g + \lambda_0b \) is about 2.36%, suggesting that the probability of receiving a large shock within a given year is about 9.5%. \( \lambda_0b \) is larger than \( \lambda_0g \), implying that large negative shocks are more likely to occur than large positive shocks. \( \lambda_1 \) is 0.33%, implying that, on an annualized basis, a 1 standard deviation increase in \( z_t \) shifts 1.33% of the probability mass from bad to good shocks.

The magnitudes associated with the state-dependent shocks are quite large. Conditional on receiving a large shock, the average absolute change in log wages is 64.1% and the standard deviation is 38.4%. Incorporating Jensen’s inequality, this estimate implies that the average decline in wages from a large negative shock is about 45%! Positive shocks induce an average increase of 110%. In comparison, the annualized standard deviation of the state-independent permanent, gaussian shock is only 6%, implying that permanent income is relatively safe when no jumps occur. The contribution from transitory shocks is more substantial. Our estimate of \( \sigma_\epsilon \) is 13.5%, implying that the standard deviation of \( \epsilon_{A,t} - \epsilon_{A,t-k} \) is 19%.

---

\(^{38}\) The whole procedure takes about 2 milliseconds for each time period. We run some diagnostics using a variety of parametric densities and find that approximation errors associated with the estimated quantiles are on the order of \( 10^{-8} \).
### Table 3: GMM calibration estimates of idiosyncratic income process parameters

This table presents the estimated parameters governing the distribution of idiosyncratic labor income shocks. Estimates are obtained by minimizing the sum of squared errors between model-implied moments and their counterparts in the data, both of which are displayed in Table 4. These moments are averages of quantiles of the cross-sectional distribution of income growth rates for individuals in the 91st through 95th percentiles of the income distribution, taken from Table 1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_g$</td>
<td>0.73%</td>
<td>Average quarterly intensity of large positive shocks</td>
</tr>
<tr>
<td>$\lambda_b$</td>
<td>1.63%</td>
<td>Average quarterly intensity of large negative shocks</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>0.33%</td>
<td>Sensitivity of quarterly intensity of large shocks to a one standard deviation shock to business cycle factor</td>
</tr>
<tr>
<td>$\mu_s$</td>
<td>64.1%</td>
<td>Absolute value of average change in log wages given a large shock</td>
</tr>
<tr>
<td>$\sigma_s$</td>
<td>38.4%</td>
<td>Standard deviation of a large shock to wages</td>
</tr>
<tr>
<td>$\sigma_n$</td>
<td>3.01%</td>
<td>Standard deviation of quarterly state-independent permanent shock</td>
</tr>
<tr>
<td>$\sigma_c$</td>
<td>13.5%</td>
<td>Standard deviation of annual state-independent temporary shock</td>
</tr>
</tbody>
</table>

Figure 4, Panel A plots the evolution of the fitted probabilities of good and back shocks, respectively, over time. 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 eventually revert back to lower levels.

Figure 4, 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 state vector, $\{z_t\}_{t=0}^T$. To emphasize the changes in higher moments, we subtract the median from each quantile. At a 1-year horizon, the central quantiles barely move at all. Consistent with Table 1, the interquartile range is essentially unchanged, while the more extreme quantiles (2.5, 5, 95, 97.5) are highly state dependent. These extreme quantiles move up and down together, increasing in expansions and falling significantly in recessions.

Figure 5 plots several densities associated with the fitted model. Panel A plots the estimated jump size distributions, helping to illustrate the magnitudes associated with the large shocks. Given the bounded support of the exponential distribution, the minimum absolute change in wages is $\mu_s - \sigma_s$, which equals 25% in calibrated model, about four times the annualized standard deviation of the permanent gaussian shock.
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 5, Panel B characterizes the densities of the permanent component of year-on-year changes in wages in expansions and recessions, respectively. To generate the figures, we randomly sample with replacement from the observed values of \( z_t \) in expansion quarters and recession quarters, respectively, then plot the densities of year-on-year changes in the permanent component of wages (i.e. we strip out transitory shocks). We 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, \( \Delta y_t \).\(^{39} \) We use a log scale on the vertical axis

\(^{39}\)We choose the slope and intercept to exactly match the first two numbers in the [91,95] column of Table 1.
Panel A plots the densities of the jump size distributions for good and bad shocks. Panel B 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 5: Densities from fitted model

Panel B: Densities of jump size distributions

Panel B: Log densities of permanent component of annual wage growth (F12M / L12M)

in order to better show the changes in the tails.\textsuperscript{40} 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 different. In expansions, the density is relatively close to symmetric, left tail is slightly fatter than the left 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.

Table 4 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 is capable of replicating the negative unconditional asymmetry which is observed in the data. By allowing for rare, large shocks, we can easily generate large cyclical variation in the tails of the distribution, while

\textsuperscript{40}On this scale, the pdf of a normal distribution is a quadratic.
leaving the central quantiles essentially unchanged. We also match the level of the 90-10 spread in expansions almost perfectly. We do not perfectly match the cyclical variation in the 50-10 and 90-10 spreads; the fitted model slightly overestimates the variation in the former and underestimates the latter.

The right panel of Table 4 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, our calibrated model targets the level of income risk faced by relatively high earners. However, we 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, we report the $R^2$ from a univariate regression of the GOS data on model-implied statistics. We consider four different time series, two of which are plotted in Figure 1: the difference between the mean and the median, the 50-10 spread, the 90-10 spread, and Kelley’s skewness.

Overall, our model matches these time series moments quite well. The $R^2$ is always highest at a 3-year horizon, which is somewhat reassuring given that no data from that horizon were used.

### Table 4: Goodness-of-fit statistics for calibrated income process

This table presents the goodness-of-fit statistics for the estimated 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 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 1. 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.
to estimate the linear index, $z_t$. We have some trouble matching the transitory variation in the 50-10 spread in recessions, suggesting that we are 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.

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 equally across households; instead, state-dependent shocks appear to be disproportionately borne by those who receive very large positive or negative shocks. We are able to replicate these features with a simple model where labor income is subject to idiosyncratic “jump risk”. Moreover, our observable proxy for idiosyncratic risk is highly persistent. In the next section, we consider the asset pricing implications of a model which allows for all of these features.

4 Theoretical Framework

In this section, we embed an endowment-based asset pricing model with heterogeneous agents and incomplete markets within a general affine, jump-diffusion framework. Our setup most closely resembles the model in Toda (2014b), itself based upon the seminal contribution of Constantinides and Duffie (1996). We place more structure on the stochastic environment, similar to Drechsler and Yaron (2011, hereafter “DY”) and Eraker and Shaliastovich (2008), which leads to approximate analytical solutions.

Our 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. In addition, agents may 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 will be 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 $C_t \equiv \int C^i_t di$, 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 \( E[\exp(\eta^i_t)] = 1 \).\(^{41}\) Households are unable to buy or sell human capital, nor trade contingent claims on realizations of \( \eta^i_t \).\(^{42}\) 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.

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 over consumption, with identical recursive preferences following Epstein and Zin (1989) and Weil (1989):

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

where \( \psi \) governs the elasticity of intertemporal substitution (EIS) and \( \gamma \) is the coefficient of relative risk aversion.\(^{43}\) At time 0, each agent begins with an initial endowment \( H^i_0 \). Thereafter, each agent chooses her consumption (the numeraire) and investment \( N^i_{kt} \) in order to maximize (9). All financial assets are in zero net supply, so market clearing will imply \( N^i_{kt} = 0 \) for all \( i, k, \) and \( t \).\(^{44}\)

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 \( \sum_{k=1}^{K} P_{kt}N^i_{kt} > -W_t \), 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. In the section that follows, we will restrict attention to symmetric, no-trade equilibria.

Under our assumptions, \( H^i_t = \exp(\eta^i_t)H^i_{t-1} \). This will imply that

\[
C^i_t = H^i_t C_t \quad \Rightarrow \quad \Delta c^i_t = \Delta c_t + \eta^i_t,
\]

where \( c_t = \log(C_t) \). We denote levels with capital letters and logs with lower case letters, e.g.

\(^{41}\)We will formalize our assumptions about \( \eta^i_t \) later in the section.
\(^{42}\)Since the financial assets are in zero net supply, no-trade will be an equilibrium. Therefore, we could assume that households are able to buy and sell human capital without affecting any of the results below. The key assumption is the inability to write contingent claims on \( \eta^i_t \).
\(^{43}\)Krebs (2007) and Toda (2013) 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 \( \delta \).
\(^{44}\)Without loss of generality, we normalize the total supply of human capital to equal 1.
\(\Delta d_{k,t} = \log(D_{k,t})\). This is a special case of our permanent income model from section 3.1 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 endowments of \(H^0_i\).\(^{45}\) While the elimination of transitory risk is a substantial departure from the data, 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.\(^{46}\)

Assumption 1 gives our general model for aggregate dynamics.

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

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

with \(y_0\) given, where \(z_{t+1}\) is i.i.d. \(N(0,1)\), \(G_{y,t}G_{y,t}'\) is a symmetric, positive semi-definite matrix, \(F_y\) has all of its eigenvalues inside the unit circle. \(J_{y,t+1}\) is a compound Poisson shock with mutually independent, i.i.d. increments and arrival intensity vector \(\lambda_{y,t}\). Further, \(\Delta c_{t+1} = S_c' y_{t+1}\) and \(\Delta d_{k,t+1} = S_k' y_{t+1}\) for \(L \times 1\) vectors \(S_c\) and \(S_1, \ldots, S_K\).

We 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. We need little structure on \(\Psi_y(u)\) beyond the existence of such a function.

Assumption 1 allows for a very flexible array of dynamics, nesting a number of popular asset pricing models as special cases. For example, Assumption 1 nests the Bansal and Yaron (2004) long-run risk model, as well as a discrete-time version of the Wachter (2011) time-varying rare disaster model.\(^{47}\) The vector \(y_{t+1}\) can include lagged values of consumption, income, or dividend growth, in addition to other state variables of interest. It is also worth noting that \(G_{y,t}\) 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 \(G_{y,t}\).

Our next assumption places some structure on the idiosyncratic shocks, \(\eta_{t+1}^i\). Assumption 2

---

\(^{45}\) If agents can write contingent claims (e.g. there are no short sales constraints) on the aggregate state, they can effectively smooth out any deterministic differences in the growth rate of the endowment process.

\(^{46}\) See, e.g. Bansal, Kiku, and Yaron (2010) and Dew-Becker and Giglio (2013).

\(^{47}\) The disaster model obtains if we assume that both \(S_c\) and \(S_k\) have a common exposure to one of the Poisson jump components, whose probability varies over time.
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.

1. \( \eta^i_t = 1_M \tilde{\eta}^i_t \), and, conditional on \( y_{t+1} \), \( \tilde{\eta}^i_t \) is generated according to
\[
\tilde{\eta}^i_{t+1} = \mu_\eta + E_\eta y_{t+1} + G_{\eta,t+1} z^i_{\eta,t+1} + J^i_{\eta,t+1},
\]
where \( z^i_{\eta,t+1} \) is a vector of standard normal random variables that is i.i.d. across agents and over time, \( G_{\eta,t} \) 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 \( \lambda_{\eta,t+1} \).

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

3. \( E[\exp(\eta^i_{t+1}) | y_{t+1}, y_t] = 1 \) almost surely for all \( y_{t+1}, y_t \in \mathbb{R}_L \).

As above, we 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.

Of the three conditions, 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. It guarantees that agents are always ex-ante identical in the model and thus that we need not consider the wealth distribution in order to study asset prices. This assumption implies that all individuals always face the same level of human capital risk. In the data, one might imagine that the distribution of idiosyncratic shocks hitting individuals who are currently unemployed could be quite different from that faced by those who currently have jobs. Assumption 2.i rules out this sort of individual-specific state dependence. The model allows us to think about unemployment-type shocks, but we must collapse all of the permanent effects of unemployment into a single shock rather than allow for a series of bad shocks which unfold gradually over time.\(^48\) It also abstracts away from heterogeneity in risks between individuals (e.g., skill heterogeneity, life cycle effects,

\(^{48}\)A similar simplifying assumption is common in the literature on rare macroeconomic disasters.
etc.) and across industries/occupations, which would require a substantially more complicated model.

The remaining assumptions involve the agents’ information and moment restrictions on $\eta_{t+1}$, 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 2.iii guarantees that the idiosyncratic shock is truly idiosyncratic (i.e. doesn’t affect the law of motion for aggregate quantities).

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) $G_{y,t}G_{y,t}' = h_y + \sum_{j=1}^{L} H_{y,j} y_{j,t}$, where $h_y$ and $H_{y,1}, \ldots, H_{y,L}$ are $L \times L$ matrices.

(ii) $\lambda_{y,t} = l_{y0} + l_{y1} y_t$, where $l_0$ and $l_1$ are $L \times 1$ and $L \times L$ matrices,

(iii) The $(1,1)$ element of $G_{\eta,t}G_{\eta,t}'$ equals $h_{\eta0} + H_{\eta1}' y_t$, where $h_\eta$ and $H_{\eta1}$ are a scalar and a $M \times 1$ vector, respectively. All other elements of $G_{\eta,t}G_{\eta,t}'$ are zero.

(iv) $\lambda_{\eta,t} = l_{\eta0} + l_{\eta1} y_t$, where $l_{\eta0}$ and $l_{\eta1}$ 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.49 In the absence of idiosyncratic shocks, our 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.50

As such, we can easily allow for CCSV within our model. Given that GOS find little evidence of CCSV in their Social Security Administration dataset, our analysis will focus more on state-dependence in the “jump” shocks. However, the theoretical implications of time-varying

---

49 The key property we exploit is that $\log E_t[\exp(u'y_{t+1})]$ and $\log E_t[u \exp(\eta_{t+1})|y_{t+1}]$ are affine functions of $y_t$ and $y_{t+1}$, respectively. Therefore, our 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.

50 Without loss of generality, we can concentrate all of the “diffusion risk” in the first element of $\tilde{\eta}_t$, so $M$, the dimension of $\tilde{\eta}_t$, is solely determined by the number of independent sources of jump risk.
volatility of the Gaussian shocks are similar. Assumption 3.iv parallels Assumption 3.ii, allowing for state-dependence in the idiosyncratic jump intensities.

It is sensible to ask whether Assumptions 2 and 3 are compatible with one another. Proposition 1, proved in section 9.1.1 of the Appendix, shows that this is the case, deriving admissible expressions for $\mu_\eta$ and $F_\eta$ which will guarantee that Assumption 2.iii holds.

**Proposition 1.** Let Assumptions 1, 1.i-ii, and 3 hold. Then,

(i) $\mu_\eta = -1/2h_{\eta0}e_1 - l_{\eta0}(\Psi_\eta(1_M) - 1_M)$

(ii) $F_\eta = -1/2e_1 \otimes h'_{\eta1} - l_{\eta1} \otimes [\Psi_\eta(1_M) - 1_M] \otimes 1'_L$,

satisfies $E[\exp(\eta^i_{t+1})|y_{t+1}, y_t] = 1$, where $\otimes$ denotes element-by-element multiplication.

Proposition 1 says that under the distributional assumptions above, there exist choices of $\mu_\eta$ and $F_\eta$ which aggregate properly. Therefore, we 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. For purposes of asset pricing, we can assume without loss of generality that $\mu_\eta$ and $F_\eta$ take the forms given in Proposition 1, significantly reducing the number of free parameters.

## 5 Asset Pricing

This section characterizes the solution for asset prices within our general theoretical framework. Section 5.1 presents the general equilibrium conditions, Section 5.2 presents analytical solutions to a log-linearized model, and Section 5.3 presents an alternative ICAPM characterization of the stochastic discount factor.

### 5.1 Equilibrium Conditions

An equilibrium in this economy is a sequence of state-contingent prices $\{P_1, \ldots, P_{K,t}\}_{t=0}^{\infty}$ and allocations $\{C^i_t, i \in I\}_{t=0}^{\infty}$ which solves agents’ optimization problems and satisfies market clearing in the capital markets. We 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
identical. Market clearing is trivially satisfied. Toda (2014b) establishes the existence and essential uniqueness of a symmetric equilibrium in a similar environment. In this section, we characterize its properties.

Mathematically, asset pricing behavior in our incomplete markets economy is identical to that of a representative agent economy where aggregate consumption is hit with an additional shock with the same distribution as \( \eta^i_t \). Our solution method closely follows Eraker and Shaliastovich (2008) and DY, who present a general solution for representative agent models with jump-diffusion shocks in continuous time and discrete time, respectively. The key difference is an additional interaction between aggregate shocks and the shape of the idiosyncratic risk distribution. Thus, while 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 different.

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:

\[
1 = E_t \left[ \delta^\theta \left( \frac{C_{t+1}^i}{C_t^i} \right)^{-\frac{\theta}{\psi}} (R_{c,t+1}^i)^{-\gamma} \tilde{R}_{t+1} \right] = E_t \left[ \delta^\theta \left( \frac{C_{t+1}^i}{C_t^i} \right)^{-\gamma} \left( \frac{WC_{t+1} + 1}{WC_t} \right)^{-(1-\theta)} \tilde{R}_{t+1} \right], \tag{14}
\]

where \( \theta = \frac{1-\gamma}{1-\psi} \), \( R_{c,t+1}^i \equiv \frac{W_{t+1}^i + C_{t+1}^i}{W_t^i} \) is the return of an (non-traded) asset delivering an arbitrary agent’s consumption stream, and \( WC_t \) is the (ex-dividend) wealth-consumption ratio.\(^{51}\) Since all agents face identical consumption risks by Assumption 2, the distribution of \( R_{c,t+1}^i \) is ex-ante identical across households. As such, the marginal rate of substitution of an arbitrary household is a valid stochastic discount factor. Plugging (11) into (14) yields

\[
1 = E_t \left[ \delta^\theta \left( \frac{C_{t+1}^i}{C_t^i} \right)^{-\gamma} \left( \frac{WC_{t+1} + 1}{WC_t} \right)^{-(1-\theta)} \tilde{R}_{t+1} \right] \equiv E_t[M_{t+1}^i \tilde{R}_{t+1}], \tag{15}
\]

so that the pricing kernel, which we will denote by \( M_{t+1}^i \), 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.\(^{52}\) The second term incorporates idiosyncratic consumption risk, which, since it is undiversifiable and uninsurable to the agent, also affects risk premia.

---

\(^{51}\)In a symmetric equilibrium, \( WC_t \) and \( \omega_t \) are identical across agents so we suppress \( i \) superscripts.

\(^{52}\)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.
With the sole exception of the consumption claim, \( \eta_{t+1}^i \) is independent of \( \tilde{R}_{t+1} \) given the past and current realizations of the aggregate state, \( y_{t+1} \). We will refer to assets satisfying this independence restriction, such as the dividend claims, as “financial assets”. When pricing financial assets, we can use the law of iterated expectations to re-write (15) as

\[
1 = E_t \left[ \delta^\theta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} \left( \frac{W_{t+1}}{W_t} + 1 \right)^{-(1-\theta)} \exp(-\gamma \cdot \eta_{t+1}^i | y_{t+1}, y_t) \right] \equiv E_t[M_{t+1} \tilde{R}_{t+1}] \quad (16)
\]

\[
= E_t \left[ \delta^\theta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} \left( \frac{W_{t+1}}{W_t} + 1 \right)^{-(1-\theta)} \sum_{j=0}^\infty \frac{(-\gamma)^j}{j!} E[\eta_{t+1}^i | y_{t+1}, y_t] \right] \tilde{R}_{t+1}, \quad (17)
\]

where the second line plugs in a Taylor series expansion of \( \exp(-\gamma \cdot \eta_{t+1}^i) \) around zero before taking the (cross-sectional) expectation.\(^{53}\) This expression shows that, in general, the pricing kernel is higher in states where even (odd) moments of the cross-sectional distribution of \( \eta_{t+1}^i \) are larger (smaller). For example, all else constant, assets which perform well when idiosyncratic second (third) moments are high (low) provide a valuable hedging benefit, lowering investors’ required rate or return.

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 idiosyncratic risk term and asset returns. The second is an indirect effect which can come from investors’ preferences over the temporal resolution of uncertainty. 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 distribution of future idiosyncratic shocks. Such a preference affects investors’ hedging demands, changing the term in the pricing kernel involving the wealth-consumption ratio. If agents have CRRA preferences, only contemporaneous covariances are priced, and the indirect effect is zero.

To illustrate the indirect effect, consider the special case when \( \eta_{t+1}^i \) is independent of all of the other random variables in (15) conditional on \( y_t \) (as opposed to the \( y_t \) and \( y_{t+1} \)). Under such an assumption, the distribution of idiosyncratic shocks is known at time \( t \) and is independent of any aggregate shocks which hit between \( t \) and \( t + 1 \). Then, we can pull the idiosyncratic risk term outside of the expectation. In this case, there is no direct effect since the correlation between the idiosyncratic risk term and excess returns is zero.\(^{54}\) Therefore, the idiosyncratic risk term will have the same impact on expected returns and the risk-free rate. However, idiosyncratic risk

\(^{53}\) The projected kernel will not price assets whose payoffs depend on \( \eta_{t+1}^i \) properly.

\(^{54}\) This follows from the identity \( \log E_t(\tilde{R}_{t+1}) - \tau_{t+1} = \log[1 - \text{cov}(M_{t+1}, \tilde{R}_{t+1})] \).
can still affect asset prices. For example, agents may still be willing to pay a premium to hedge against bad news about the distribution of idiosyncratic shocks in periods \( t + 2 \) and beyond.\(^{55}\)

In equilibrium, (14) and (15) must hold for all assets in the investment opportunity set. Plugging the consumption claim and financial asset returns into (15) yields

\[
1 = \mathbb{E}_t \left[ \delta^\theta \left( \frac{C_{t+1}}{C_t} \right)^{1-\gamma} \left( \frac{WC_{t+1} + 1}{WC_t} \right) \exp((1-\gamma)\eta_{t+1}) \right] \\
1 = \mathbb{E}_t \left[ \delta^\theta \left( \frac{C_{t+1}}{C_t} \right)^{-\gamma} \left( \frac{WC_{t+1} + 1}{WC_t} \right)^{-\theta} \exp(-\gamma \cdot \eta_{t+1}) \left( \frac{P_{k,t+1}}{P_{kt}} \right)^{\frac{D_{k,t+1}}{D_{kt}}} \right],
\]

a system of nonlinear equations involving the two key valuation ratios, \( WC_t \) and \( P_{kt}/D_{kt} \). Since all of exogenous quantities in (18-19) are stationary by Assumptions 1-3, the model may be solved numerically by finding the value of \( WC_t \) that satisfies (18) for each value of \( y_{t+1} \) in the state space. Then, given the solution for \( WC_t \), one can use (19) to solve for the equilibrium price-dividend ratios for the financial assets.

### 5.2 Solution

In this section, we briefly outline how to solve for asset prices with incomplete markets and time-varying idiosyncratic risk. Since our model for aggregate dynamics is quite general and has many properties that have been studied elsewhere, we will primarily focus on the incremental effects associated with incomplete markets. For brevity, many technical details may be found in the Appendix.\(^{56}\) The only requisite approximation is the standard Campbell and Shiller (1988) log-linearization of returns.

#### 5.2.1 Log-linearization

Denote continuously compounded returns by lowercase letters (e.g. \( r_{c,t}^i = \log R_{c,t}^i \)). We approximate the return on the consumption claim around a constant log wealth-consumption ratio \( \frac{wc}{c} \), yielding

\[
r_{c,t+1}^i \approx \kappa_c + \Delta c_{t+1}^i + \rho_c wc_{t+1} - wc_t,
\]

\(^{55}\)Under CRRA utility, \( \theta = 1 \) and the term involving the wealth-consumption ratio drops out entirely. Under this assumption, then, there is no means for idiosyncratic risk to affect the equity premium.

\(^{56}\)Additional details and discussion are available in Eraker and Shaliastovich (2008) and DY.
with linearization constants \( \rho_c \equiv \frac{\exp(\bar{w}c)}{\exp(\bar{w}c) + 1} < 1 \) and \( \kappa_c \equiv \log(\exp(\bar{w}c) + 1) - \frac{\exp(\bar{w}c)}{\exp(\bar{w}c) + 1} \bar{w}c \).

57

Combining (14) and (20), the log of the one period pricing kernel approximately equals

\[
m_{t+1} = \theta \log \delta - (1 - \theta) \kappa_c - \gamma (\Delta c_{t+1} + \nu_{t+1}) - (1 - \theta)(\rho_c w_{ct+1} - wc_t) - \gamma \cdot \eta^i_{t+1}.
\]

Equation (21) is the linearized version of (15), where the representative agent pricing kernel is augmented by an additional term capturing idiosyncratic risk.

Analogous to the result in (16), for purposes of pricing financial assets, we can replace \( \eta^i_{t+1} \) with its projection onto the aggregate state, \( \log E_t[\exp(-\gamma \cdot \eta^i_{t+1})|y_{t+1}] \), in the log-linearized pricing kernel. Note that \(-\frac{1}{\gamma} \log E_t[\exp(-\gamma \cdot \eta^i_{t+1})|y_{t+1}] \equiv \nu_{t+1} \) can be interpreted as the log of an expected utility maximizer’s certainty equivalent for lottery \( \exp(\eta^i_{t+1}) \) given \( y_{t+1} \). Thus, we can equivalently consider the following projection of the pricing kernel

\[
m_{t+1} = \theta \log \delta - (1 - \theta) \kappa_c - \gamma (\Delta c_{t+1} + \nu_{t+1}) - (1 - \theta)(\rho_c w_{ct+1} - wc_t) - \gamma \cdot \eta^i_{t+1} - \beta(\cdot) y_{t+1},
\]

where \( \kappa \equiv \theta \log \delta - (1 - \theta) \kappa_c + \beta_0(-\gamma) \). Note that (23) will correctly price financial assets; the solution method for assets whose payoffs depend on \( \eta^i_{t+1} \), namely the consumption claim, is somewhat different.

5.2.2 Valuation Ratios

Proposition 2 gives our key result, namely that the wealth-consumption ratio is an affine function of the aggregate state vector.

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

\[
m_{t+1} = \kappa - \gamma (\Delta c_{t+1} + \nu_{t+1}) - (1 - \theta)(\rho_c w_{ct+1} - wc_t) + \beta(-\gamma) y_{t+1},
\]

where \( \kappa \equiv \theta \log \delta - (1 - \theta) \kappa_c + \beta_0(-\gamma) \). Note that (23) will correctly price financial assets; the solution method for assets whose payoffs depend on \( \eta^i_{t+1} \), namely the consumption claim, is somewhat different.

---

57 Analogously, the returns on the dividend claims approximately satisfy \( r_{k,t+1} \approx \kappa_k + \Delta d_{k,t+1} + \rho_k pd_{k,t+1} - pd_{k,t} \) where \( pd_{k,t} \) 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 \( \bar{w}c \).
(i) \( wc_t = A_0 + A'y_t, \)

(ii) \( pd_k = A_{0,k} + A_k'y_t, \) 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 most of the details are in section 9.1.2 of the Appendix, a brief outline of the solution method is as follows. We begin by conjecturing (and later verifying) that the log of the wealth-consumption ratio is an affine function of \( y_t. \)\(^{58}\) We solve for \( A_0 \) and \( A \) using the Euler equation for the consumption claim and the method of undetermined coefficients. Given our restrictions on the law of motion for the state vector, we can evaluate the Euler equations analytically.\(^{59}\) Since the Euler equations must hold for each \( y_t \) in the state space, we 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 the valuation ratios for the \( K \) other risky assets. We guess, then verify, that the log price-dividend ratio for the \( k^{th} \) risky asset is \( pd_k = A_{0,k} + A_k'y_t. \) Given the solution for the wealth-consumption ratio, we calculate the projected version of the pricing kernel in (23) and use the method of undetermined coefficients to solve for \( A_{0,k} \) and \( A_k. \)

Plugging the affine form into (23), the projected pricing kernel, and subtracting off terms known as of time \( t \) yields

\[
m_{t+1} - E_t(m_{t+1}) = -[\gamma S_c' - \beta(-\gamma)' + (1 - \theta)\rho_cA']\left[y_{t+1} - E_t(y_t)\right] \equiv -\Lambda'[y_{t+1} - E_t(y_t)], \tag{24}
\]

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 consumption risk; the first is the usual representative agent term capturing the aggregate consumption innovation, and the second term captures contemporaneous news about the distribution of idiosyncratic risk (i.e. the certainty equivalent \( \nu_{t+1} \)). 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, incorporating the indirect effect.

\(^{58}\)Since lagged values of \( \eta_t \) 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, \( y_t. \)

\(^{59}\)A general expression for the expectation of an exponential affine function of \( y_{t+1} \) given \( y_t \) is given in Lemma 2 in the Appendix. This expectation is also an exponential affine function of \( y_t \).
5.2.3 Risk Premia

Conditional on the vector with the prices of risk ($\Lambda$), and the dividend-price ratio coefficients ($A_{0,k}$ and $A_k$), the representative agent solutions in DY go through with almost no modifications. For example, the vast majority of the excellent discussion in DY describes the model conditional on the valuation ratios and, as such, is directly applicable here. Thus, our discussion is quite brief. It is worth emphasizing, however, that 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.

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

$$r_{k,t+1} = \kappa_k + (\rho_k - 1)A_{0,k} + (S'_{k} + \rho_k A'_k)\gamma_{t+1} - A'_{k}y_{t} \equiv \kappa_k + (\rho_k - 1)A_{0,k} + B'_k y_{t+1} - A'_k y_{t}. \quad (25)$$

Proposition 3 gives the solution for the equity premium, which is derived in Section 9.1.3 of the Appendix.

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

$$\log(E_t[R_{k,t+1}]) - rf_{t+1} = B'_{k}G_{y,t}G'_{y,t}\Lambda + \lambda'_{y,t}[\Psi_{y}(B_{k}) - 1_L] - \lambda'_{y,t}[\Psi_{y}(B_{k} - \Lambda) - \Psi_{y}(-\Lambda)]. \quad (26)$$

The first term reflects the covariance between the Gaussian innovation to returns and the pricing kernel. The second term is the expected value of the jump component of returns under the physical measure, while the last term subtracts off its expected value under the risk-neutral measure. The difference between the two reflects the compensation for jump risk. The terms are additive because of the independence of Gaussian and jump shocks. If $G_{y,t}$ or $\lambda_{y,t}$ vary over time, then the equity premium can also be time-varying.\(^{60}\)

An immediate corollary to Proposition 3 is that a necessary condition for an extensive margin measure to predict excess returns is that it is correlated with uncertainty about shocks to the aggregate state. In the next section, we present evidence that this condition holds in the data. When employment growth is low, it is generally the case that uncertainty about future employment growth is high.

---

\(^{60}\)For additional details and discussion, we also refer the reader to DY, sections 3.3.3 and A.4.
5.2.4 The Term Structure of Risk Premia

Part of our analysis will study the role of state-dependent idiosyncratic risk in affecting the term structure of risk premia. Understanding the pricing of risky cash flows at different points in time helps to clarify the mechanics of the model, and, in some cases, generate additional testable predictions. 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: \( C^i_{t+h} \), 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 our 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, we 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 we 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 \( R^h_{k,t+1} \equiv P^h_{k,t+1}/P^h_{k,t} \) 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

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

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 4 says that the valuation ratios for the zero-coupon consumption and dividend claims are affine functions of $y_t$.

**Proposition 4.** Let Assumptions 1-3 hold. The log-linearized model satisfies

(i) $\log(P^{h}_{c,t}/C_t) \equiv w^h_c = A^h_0 + y_t^t A^h_c,$

(ii) $\log(P^{h}_{k,t}/D_{k,t}) \equiv p^h_{k,t} = A^h_{0,k} + y_t^t A^h_{k},$ 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, we can study how the term structures of real bond yields, expected dividend growth rates, and risk premia evolve over time.

### 5.3 ICAPM Representation of the Pricing Kernel

Campbell (1991) 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. In this section, we derive an ICAPM representation for our incomplete markets economy. Many of the steps of the derivation follow Campbell et al. (2013), so we highlight the incremental effects of adding incomplete markets. 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, we extend Campbell et al. (2013) to allow for general affine jump-diffusion dynamics for the aggregate state vector, $y_t$.

Define $R_{c,t+1} \equiv E[R^i_{c,t+1}\mid y_{t+1}]$ and $r_{c,t+1} = \log R_{c,t+1}$. Following Campbell (1991), we substitute out $\Delta c_{t+1}$ using the identity

$$\Delta c_{t+1} \approx r_{c,t+1} - \kappa_c - \rho_c w_{c,t+1} + w_c.$$  

(28)

Plugging (28) into the log-linearized pricing kernel (21), one obtains

$$m_{c,t+1} = \text{const.} - \gamma(r_{c,t+1} + \eta_{c,t+1}) + \frac{\theta}{\sigma} (\rho_c w_{c,t+1} - w_c).$$  

(29)
Plugging (29) into the Euler equation for \( r_{c,t+1} \) and projecting out \( \eta_{i,t+1} \) yields

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

where \( \nu_{t+1}^* \equiv \frac{1}{1-\gamma} \log E_{t+1}[\exp(1 - \gamma)\eta_{i,t+1}|y_{t+1}] \). Again, \( \nu_{t+1}^* \) 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 \( \eta_{i,t+1} \). Our distributional assumptions imply

\[
wc_t = \text{const.} + (\psi - 1)[E_t r_{c,t+1} + E_t \nu_{t+1}^*] + \rho_c E_t wc_{t+1} + \frac{1}{2} \theta \right),
\]

where \( \vartheta_t \) is a Jensen’s inequality term. In the absence of jump risk, this term equals \( Var_t [m_{t+1}^i + r_{c,t+1}^i] \), 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 (31) and assuming that \( \lim_{j \to \infty} \rho^j w_{c,t+1} = 0 \), one obtains

\[
w_{c,t} = \text{const.} + E_t \sum_{j=0}^{\infty} \rho^j_c [\psi - 1](r_{c,t+1+j} + \nu_{t+1,j}^*) + \frac{1}{2} \psi \theta_{t+j}]
\]

\[
\rho_c [w_{c,t+1} - E_t wc_{t+1}] = [E_{t+1} - E_t] \sum_{j=1}^{\infty} \rho^j_c [\psi - 1](r_{t+1+j} + \nu_{t+1,j}^*) + \frac{1}{2} \psi \theta_{t+1+j}]
\]

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

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

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

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

The subscript \( \text{FIR} \) is shorthand for future idiosyncratic risk, which captures news about the higher moments of idiosyncratic shocks.\(^{61}\) Equation (34) 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.

\(^{61}\)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.
Plugging (34-35) into the projected pricing kernel (22), one obtains

\[ m_{t+1} - E_t m_{t+1} = -\gamma N_{CIR,t+1} + (1 - \gamma) N_{FIR,t+1} + N_{DR,t+1} - \gamma N_{CF,t+1} + \frac{1}{2} N_{UNC,t+1}. \]  

(36)

Idiosyncratic risk adds two additional 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, where \( N_{CIR,t+1} \equiv \nu_{t+1} - E_t \nu_{t+1} \). Agents dislike assets that perform badly when the cross-sectional certainty equivalent, \( \nu_{t+1} \), is unexpectedly low. This term provides a rationale for cross-sectional asset-pricing tests such as Jagannathan and Wang (1996) which include the labor income growth rate as a priced factor. This is the primary source of risk considered in the literature on testing household Euler equations using the higher moments of cross-sectional household consumption growth.

The second term provides compensation for news about the future trajectory of idiosyncratic risk. Given the high persistence of extensive margin measures, 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 our theoretical framework.

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. Marginal utility is low when expected future consumption growth is high, so assets that covary positively with cash flow news carry higher risk premiums. 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.\(^{62}\) Intuitively, discount rate shocks are transitory in nature, whereas cash flow shocks are permanent, making the latter more important to an investor with a long time horizon.

Equation (36) 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 \( \nu^*_{t+1} \) is more persistent and/or volatile than aggregate consumption growth, this term can play a very

\[^{62}\text{Standard choices of } \gamma \text{ tend to range between 5 and 10.}\]
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 which govern 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, we refer the interested reader to Campbell et al. (2013).

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 our assumptions, $\vartheta_t$ and $\nu^*_t$ are affine function functions of $y_t$ and $y_{t+1}$, respectively. Therefore, one can express all of the news terms above as linear combinations of the VAR 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 high frequency dynamics of $\Delta c_{t+1}$. We provide explicit expressions for these linear combinations in the Appendix.

6 Interactions between Idiosyncratic Risk and Stock Returns

In this section, we discuss the implications of our incomplete markets model for return predictability, and we consider the dynamic interactions between proxies for idiosyncratic risk and asset returns. In our model, a necessary condition for a variable to predict returns is for it to govern uncertainty about the aggregate state vector. We begin by demonstrating that initial claims for unemployment is a very good proxy for uncertainty about labor market conditions, suggesting that this necessary condition is satisfied. Next, we present new evidence that initial claims for unemployment predicts excess returns on the market portfolio, as well as Fama and French’s (1992) SMB portfolio, outperforming a number of conventional predictors. We also explore the covariance structure between initial claims and these predictor variables.

6.1 Uncertainty about idiosyncratic risk is countercyclical

From Proposition 3 above, in order for a variable to directly predict excess returns, it must govern the degree of uncertainty about the aggregate state vector, $y_t$. In our quantitative model below, a single variable will govern both the level of idiosyncratic risk, as well as uncertainty about idiosyncratic risk. Earlier, we demonstrated that one of the most successful predictors of
the skewness of the idiosyncratic risk distribution is private sector employment growth. A well-established, leading indicator of future changes in employment is the rate of initial claims for unemployment. We divide the number of claims filed in each month by the size of the workforce (from the BEA) in order to get a variable which may be interpreted as the rate of job loss.

While the relationship between employment growth and initial claims is relatively well known, we argue that it also proxies well for uncertainty about labor market conditions. To demonstrate this, we compare initial claims with a more direct uncertainty measure which exploits cross-sectional variation in business cycle conditions across states. While uncertainty about aggregate employment growth is more directly tied to the theory, aggregate employment data are, at best, available at a monthly frequency, implying that realized volatility measures are very imprecisely
estimated. We use quarterly, seasonally-adjusted, state-level employment growth data from Hamilton and Owyang (2011).

For each state, we estimate the following model:

\[
\Delta emp_{s,t} = \alpha_s + \beta_s \Delta emp_t + \sigma_s u_{s,t},
\]

(37)

where \(\overline{emp}_t\) is the cross-sectional average employment growth. Our uncertainty measure is the cross-sectional volatility of the fitted residuals:

\[
Vol_t \equiv \sqrt{\frac{1}{S} \sum_{s=1}^{I} \hat{u}_{s,t}^2}.
\]

Figure 6 plots the two (standardized) measures, \(Vol_t\) and initial claims for unemployment, over the time period for which both series are available. We also plot a smoothed version of the volatility measure, calculated using a 5 month centered moving-average filter. The low frequency components of both series (which carry the most importance within an asset pricing framework) are virtually identical, though there are some non-trivial differences between the two measures during the double-dip recession of 1982. In a cross-sectional sense, then, initial claims is a very good proxy for cross-sectional uncertainty about labor market conditions.

Within our model, risk premium dynamics depend on uncertainty about aggregate shocks, particularly about shocks to persistent state variables. Recall that our cross-sectional measure, \(Vol_t\), already controls for average employment growth. Here, we demonstrate that both measures—\(Vol_t\) and initial claims for unemployment—are good proxies for uncertainty about future idiosyncratic risk. In the labor market, cross-sectional and time series uncertainty are closely related.

To demonstrate this relationship more formally, we consider two tests which are similar to the ARCH test of Engle (1982). We first estimate an AR(4) model for each of the estimated skewness indices, \(\hat{skew}_t\), from the GMM specifications in Panel B on Table 2,

\[
\hat{skew}_t = a_0 + \sum_{j=1}^{4} a_j \hat{skew}_{t-1} + \sigma v_t,
\]

(38)

where \(\sigma\) is the unconditional volatility of the fitted residual. Then, we estimate \(b_1\) and \(c_1\) in

\[
|\hat{v}_t| = b_0 + b_1 Unc_t + \epsilon_{1t} \quad \text{(39)}
\]

\[
\hat{v}_t^2 = c_0 + c_1 Unc_t^2 + \epsilon_{2t}, \quad \text{(40)}
\]

where \(Unc_t\) is one of our uncertainty measures. Given concerns about potential measurement errors in \(Vol_t\), we also report the coefficient from a two-stage least squares regression, where

---

63 We are grateful to James Hamilton for making the data available. We extend the data to the present by aggregating the monthly, seasonally-adjusted series which are now provided by the Bureau of Economic Analysis.
<table>
<thead>
<tr>
<th>Quarterly skewness model</th>
<th>Absolute AR(4) residuals</th>
<th></th>
<th></th>
<th></th>
<th>Squared AR(4) residuals</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\text{Vol}_t$</td>
<td>$\text{Claims}_t$</td>
<td>IV</td>
<td></td>
<td>$\text{Vol}_t^2$</td>
<td>$\text{Claims}_t^2$</td>
<td>IV</td>
<td></td>
</tr>
<tr>
<td>Model 1: $\Delta \text{emp}_t$</td>
<td>0.314***</td>
<td>0.235***</td>
<td>0.392***</td>
<td>0.853***</td>
<td>0.597***</td>
<td>0.971***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.054)</td>
<td>(0.053)</td>
<td>(0.092)</td>
<td>(0.219)</td>
<td>(0.193)</td>
<td>(0.293)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2: $\Delta y_t$</td>
<td>0.189***</td>
<td>0.066</td>
<td>0.126</td>
<td>0.726***</td>
<td>0.290</td>
<td>0.472</td>
<td>0.238 (0.084)</td>
<td>(0.221)</td>
</tr>
<tr>
<td></td>
<td>(0.060)</td>
<td>(0.057)</td>
<td>(0.084)</td>
<td>(0.155)</td>
<td>(0.170)</td>
<td>(0.261)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 3: $\text{pw}_{t-1}$</td>
<td>0.075</td>
<td>0.132**</td>
<td>0.215**</td>
<td>0.198</td>
<td>0.294*</td>
<td>0.493*</td>
<td>0.198</td>
<td>0.294*</td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td>(0.052)</td>
<td>(0.085)</td>
<td>(0.155)</td>
<td>(0.170)</td>
<td>(0.261)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 4: $\Delta c_t$</td>
<td>0.205***</td>
<td>0.174***</td>
<td>0.287***</td>
<td>0.459***</td>
<td>0.426**</td>
<td>0.692**</td>
<td>0.198</td>
<td>0.294*</td>
</tr>
<tr>
<td></td>
<td>(0.046)</td>
<td>(0.053)</td>
<td>(0.098)</td>
<td>(0.156)</td>
<td>(0.204)</td>
<td>(0.316)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 5: $\Delta \text{emp}<em>t$ and $\text{pw}</em>{t-1}$</td>
<td>0.287***</td>
<td>0.211***</td>
<td>0.351***</td>
<td>0.826***</td>
<td>0.678**</td>
<td>1.083***</td>
<td>0.287***</td>
<td>0.211***</td>
</tr>
<tr>
<td></td>
<td>(0.057)</td>
<td>(0.069)</td>
<td>(0.123)</td>
<td>(0.236)</td>
<td>(0.273)</td>
<td>(0.414)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 6: $\Delta y_t$ and $\text{pw}_{t-1}$</td>
<td>0.185***</td>
<td>0.069</td>
<td>0.130</td>
<td>0.683***</td>
<td>0.285</td>
<td>0.459</td>
<td>0.214</td>
<td>0.207</td>
</tr>
<tr>
<td></td>
<td>(0.058)</td>
<td>(0.057)</td>
<td>(0.084)</td>
<td>(0.214)</td>
<td>(0.207)</td>
<td>(0.345)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 7: $\Delta y_t$ and $\Delta c_t$</td>
<td>0.196***</td>
<td>0.078</td>
<td>0.150</td>
<td>0.762***</td>
<td>0.355</td>
<td>0.551</td>
<td>0.226</td>
<td>0.245</td>
</tr>
<tr>
<td></td>
<td>(0.059)</td>
<td>(0.063)</td>
<td>(0.099)</td>
<td>(0.226)</td>
<td>(0.245)</td>
<td>(0.380)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 5:** Cross-sectional employment growth volatility and initial claims proxy for volatility of innovations to idiosyncratic skewness indices

We estimate an AR(4) model for each quarterly third central moment estimate from the GMM specifications in Panel B of Table 2. For each skewness index, the left panel present the coefficient from an OLS regressions of the absolute residuals on $\text{Vol}_t$, then $\text{Claims}_t$. To aid in interpretation, we standardize the residuals and the regressors to have unit variance. The “IV” specification is a two-stage least squares regression, where $\text{Claims}_t$ is used as an instrument for $\text{Vol}_t$. The right columns repeat the analysis for squared residuals. The sample period is 1967:1 through 2012:3. Robust standard errors are in parentheses.

Table 5: Cross-sectional employment growth volatility and initial claims proxy for volatility of innovations to idiosyncratic skewness indices

We estimate an AR(4) model for each quarterly third central moment estimate from the GMM specifications in Panel B of Table 2. For each skewness index, the left panel present the coefficient from an OLS regressions of the absolute residuals on $\text{Vol}_t$, then $\text{Claims}_t$. To aid in interpretation, we standardize the residuals and the regressors to have unit variance. The “IV” specification is a two-stage least squares regression, where $\text{Claims}_t$ is used as an instrument for $\text{Vol}_t$. The right columns repeat the analysis for squared residuals. The sample period is 1967:1 through 2012:3. Robust standard errors are in parentheses.

We standardize the right hand side variable in each regression to have variance one, so $b_1$ can be interpreted as the effect of a one standard deviation increase in $\text{Unc}_t$ on the conditional volatility of $v_t$. Analogously, $c_1$ is the effect of a one standard deviation increase in $\text{Unc}_t^2$ on the conditional variance of $v_t$.

The results are qualitatively identical regardless of the specification considered. In increase in either of our proxies is associated with an increase in the variance of the forecast errors for each of our skewness proxies. For our cross-sectional measure, $\text{Vol}_t$, our estimates are significant at the 1% level for all models except for model 3. The coefficients suggest that a one standard deviation increase in $\text{Vol}_t$ is associated with a 20-30% increase in the standard deviation (45-85% increase in the variance) of the forecast errors from our skewness models. The instrumental variables estimates are generally higher, consistent with a moderate level of measurement error in $\text{Vol}_t$. While we motivated initial claims for its ability to track low-frequency variation in $\text{Vol}_t$, it also has substantial forecasting power for the variance of the residuals. The coefficients
are statistically significant in models 1 and 3-5, and it appears to be better capable of capturing uncertainty about shocks to the profit-compensation ratio (model 3).

6.2 Labor market uncertainty predicts returns

One of our primary objectives is to study the ability of an asset pricing model with incomplete markets to generate large, time-varying equity premia. In this section, we test a necessary condition for such a model by considering the ability of our labor market uncertainty measure, initial claims, to predict returns in the data.\textsuperscript{64} We also consider its covariance with and compare its forecasting power with other leading predictor variables from the extant literature.

We demonstrate below 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, we 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.

Since a wealth of potential predictors have been suggested, we 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, we 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, we 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. We 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. As we will see below, 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.

We 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,\textsuperscript{64}

---

\textsuperscript{64}We emphasize initial claims over our cross-sectional volatility measure because it is available at a higher frequency, does not require the estimation of any parameters, and is less likely to be prone to measurement errors.
we 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 6 presents a number of pairwise correlations between initial claims, each of our predictors, and a measure of employment declines over the next three months—a monthly measure of the first of our skewness proxies. 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

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Initial claims</th>
<th>1928-2012</th>
<th>1967-2012</th>
<th>1967-2012 (orth.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employment cuts</td>
<td>0.38*</td>
<td>0.10*</td>
<td>0.35*</td>
<td>0.41*</td>
</tr>
<tr>
<td>Equal-weighted equity premium forecast combinations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1928 - 2012</td>
<td>0.67*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1967 - 2012</td>
<td>0.58*</td>
<td>0.77*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1967 - 2012 (orth.)</td>
<td>0.00</td>
<td>0.29*</td>
<td>0.72*</td>
<td></td>
</tr>
<tr>
<td>Goyal and Welch (2008) predictors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dividend yield (dy)</td>
<td>0.74*</td>
<td>0.72*</td>
<td>0.41*</td>
<td>-0.27*</td>
</tr>
<tr>
<td>Earnings-price ratio (ep)</td>
<td>-0.46*</td>
<td>-0.59*</td>
<td>-0.10*</td>
<td>0.47*</td>
</tr>
<tr>
<td>Book-to-market ratio (bm)</td>
<td>0.76*</td>
<td>0.67*</td>
<td>0.28*</td>
<td>-0.41*</td>
</tr>
<tr>
<td>Stock market realized variance (svar)</td>
<td>0.01</td>
<td>0.13*</td>
<td>0.35*</td>
<td>0.38*</td>
</tr>
<tr>
<td>3 month T-bill rate (tbl)</td>
<td>0.41*</td>
<td>0.23*</td>
<td>-0.11*</td>
<td>-0.72*</td>
</tr>
<tr>
<td>Term spread (tms)</td>
<td>0.09*</td>
<td>0.23*</td>
<td>0.63*</td>
<td>0.79*</td>
</tr>
<tr>
<td>Default yield: BAA - AAA spread (dfy)</td>
<td>0.69*</td>
<td>0.62*</td>
<td>0.71*</td>
<td>0.31*</td>
</tr>
<tr>
<td>Long term yield (lty)</td>
<td>0.57*</td>
<td>0.43*</td>
<td>0.24*</td>
<td>-0.41*</td>
</tr>
<tr>
<td>Net issuance (ntis)</td>
<td>0.08</td>
<td>-0.39*</td>
<td>-0.26*</td>
<td>-0.36*</td>
</tr>
<tr>
<td>Inflation (infl)</td>
<td>0.24*</td>
<td>0.09*</td>
<td>-0.37*</td>
<td>-0.69*</td>
</tr>
<tr>
<td>Corporate - govt bond return (dfr)</td>
<td>0.07</td>
<td>-0.05</td>
<td>0.04</td>
<td>0.00</td>
</tr>
<tr>
<td>Long term bond return (ltr)</td>
<td>0.08</td>
<td>0.24*</td>
<td>0.31*</td>
<td>0.28*</td>
</tr>
</tbody>
</table>

Table 6: 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 is our 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.
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.  

Next, we 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 7 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 the combination forecast less closely. During this period, the pairwise correlation between initial claims and the combination forecast is still higher than any of the other univariate predictors.

Finally, when we 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 were 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 7 summarizes the forecasting performance of each of our predictor variables for cumulative returns. We report the $R^2$ and the $t$-statistic on $\beta_h$ from the following predictive regression:

$$\sum_{j=1}^{h} r_{t+j} \equiv r_{t:t+h} = \alpha_h + \beta_h x_t + u_{t:t+h},$$  \hspace{1cm} (41)

65 An even stronger negative correlation (-72%) arises between the dividend yield and the earnings price ratio.
<table>
<thead>
<tr>
<th>Predictor</th>
<th>3 mo</th>
<th>1 yr</th>
<th>2 yr</th>
<th>3 mo</th>
<th>1 yr</th>
<th>2 yr</th>
</tr>
</thead>
<tbody>
<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>
</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>
</tr>
</tbody>
</table>

Equal-weighted forecast combinations

<table>
<thead>
<tr>
<th>Period</th>
<th>Market Excess Return</th>
<th>SMB Return</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3 mo</td>
<td>1 yr</td>
</tr>
<tr>
<td>1928 - 2012</td>
<td>2.22**</td>
<td>8.96***</td>
</tr>
<tr>
<td></td>
<td>(2.07)</td>
<td>(2.71)</td>
</tr>
<tr>
<td>1967 - 2012</td>
<td>3.39**</td>
<td>14.88***</td>
</tr>
<tr>
<td></td>
<td>(2.05)</td>
<td>(4.19)</td>
</tr>
<tr>
<td>1967 - 2012</td>
<td>2.41</td>
<td>9.49***</td>
</tr>
<tr>
<td>(orth. predictors)</td>
<td>(1.64)</td>
<td>(2.82)</td>
</tr>
</tbody>
</table>

Univariate regressions with Goyal and Welch (2008) predictors

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Market Excess Return</th>
<th>SMB Return</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3 mo</td>
<td>1 yr</td>
</tr>
<tr>
<td>dy</td>
<td>0.97</td>
<td>3.12</td>
</tr>
<tr>
<td></td>
<td>(1.44)</td>
<td>(1.47)</td>
</tr>
<tr>
<td>ep</td>
<td>0.32</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>(-0.48)</td>
<td>(-0.60)</td>
</tr>
<tr>
<td>bm</td>
<td>0.29</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>(0.50)</td>
<td>(0.63)</td>
</tr>
<tr>
<td>svar</td>
<td>0.31</td>
<td>0.85*</td>
</tr>
<tr>
<td></td>
<td>(-0.44)</td>
<td>(1.75)</td>
</tr>
<tr>
<td>tbl</td>
<td>0.67</td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td>(-1.10)</td>
<td>(-0.84)</td>
</tr>
<tr>
<td>tms</td>
<td>1.76**</td>
<td>8.04***</td>
</tr>
<tr>
<td></td>
<td>(2.01)</td>
<td>(2.67)</td>
</tr>
<tr>
<td>dfy</td>
<td>1.29</td>
<td>3.82*</td>
</tr>
<tr>
<td></td>
<td>(1.30)</td>
<td>(1.88)</td>
</tr>
<tr>
<td>lty</td>
<td>0.19</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>(-0.17)</td>
<td>(0.40)</td>
</tr>
<tr>
<td>ntis</td>
<td>0.21</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>(-0.20)</td>
<td>(-0.31)</td>
</tr>
<tr>
<td>infl</td>
<td>0.70</td>
<td>2.95**</td>
</tr>
<tr>
<td></td>
<td>(-0.94)</td>
<td>(-2.09)</td>
</tr>
<tr>
<td>dfr</td>
<td>0.92</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>(1.46)</td>
<td>(0.88)</td>
</tr>
<tr>
<td>ltr</td>
<td>0.62</td>
<td>1.64***</td>
</tr>
<tr>
<td></td>
<td>(1.27)</td>
<td>(3.41)</td>
</tr>
</tbody>
</table>

Table 7: 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. We consider the market excess return as well as the Fama and French (1993) SMB portfolio. We 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.
Figure 7: 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.

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. We consider forecasts of the log excess return on the CRSP value-weighted index, as well as the Fama and French (1993) SMB portfolio. We consider forecast horizons ($h$) of 3, 12, and 24 months, though results are similar at other horizons. Our sample period is 1967-2012. In order to make an apples-to-apples comparison, we limit our attention to the period for which initial claims data are available.

The results in Table 7 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

---

66 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 we only find weak evidence that initial claims forecasts HML at long horizons.
achieves an $R^2$ of 2.4%. Our measure 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 7. 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 we are 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, we 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. Our 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.

Table 8 repeats the analysis where each of the variables from Table 7 is orthogonalized with respect to initial claims prior to running the predictive regressions. Initial claims effectively captures the predictive content of many of the variables, particularly the dividend yield (dy), the earnings-price ratio, and the default yield (dfy). Variables related to inflation and/or the yield curve appear have additional predictive content. However, as discussed above, these variables are not unrelated to the extensive margin, given the 41% correlation between the combination forecast constructed with orthogonalized predictors and future cuts in employment.
### Table 8: Predictive regressions for excess returns on Market and SMB portfolios with orthogonalized predictors

This table plots the $R^2$ values (in percentage points) from predictive regressions of cumulative returns on a number of univariate state variables. We consider the market excess return as well as the Fama and French (1993) SMB portfolio. All of the forecast combinations and Goyal and Welch (2008) predictors have been orthogonalized with respect to initial claims. We 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></th>
<th>SMB Return</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<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></td>
<td>2.40**</td>
<td>3.73**</td>
<td>4.96</td>
<td>4.29***</td>
<td>10.82***</td>
<td>9.47***</td>
<td></td>
</tr>
<tr>
<td></td>
<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, orthogonalized</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1928 - 2012</td>
<td></td>
<td>0.52</td>
<td>3.23</td>
<td>4.62</td>
<td>0.47</td>
<td>0.22</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.81)</td>
<td>(1.34)</td>
<td>(1.53)</td>
<td>(0.90)</td>
<td>(0.19)</td>
<td>(-0.42)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.50</td>
<td>8.82***</td>
<td>12.27***</td>
<td>1.03*</td>
<td>0.30</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.22)</td>
<td>(2.90)</td>
<td>(3.50)</td>
<td>(1.76)</td>
<td>(0.43)</td>
<td>(-0.65)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></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></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, orthogonalized</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dy</td>
<td></td>
<td>0.28</td>
<td>0.21</td>
<td>0.56</td>
<td>0.75</td>
<td>3.22*</td>
<td>0.57*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-0.50)</td>
<td>(-0.14)</td>
<td>(0.40)</td>
<td>(-1.17)</td>
<td>(-1.66)</td>
<td>(-1.86)</td>
<td></td>
</tr>
<tr>
<td>ep</td>
<td></td>
<td>0.31</td>
<td>0.35</td>
<td>0.39</td>
<td>0.52</td>
<td>1.31</td>
<td>1.57</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.49)</td>
<td>(0.36)</td>
<td>(0.35)</td>
<td>(1.19)</td>
<td>(1.27)</td>
<td>(0.87)</td>
<td></td>
</tr>
<tr>
<td>bm</td>
<td></td>
<td>1.70**</td>
<td>3.15</td>
<td>3.62</td>
<td>0.69</td>
<td>0.66</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-1.98)</td>
<td>(-1.44)</td>
<td>(-1.33)</td>
<td>(-1.16)</td>
<td>(-0.59)</td>
<td>(-0.28)</td>
<td></td>
</tr>
<tr>
<td>svar</td>
<td></td>
<td>0.31</td>
<td>0.81*</td>
<td>1.67**</td>
<td>0.45</td>
<td>1.57*</td>
<td>2.08*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-0.45)</td>
<td>(1.70)</td>
<td>(2.23)</td>
<td>(1.57)</td>
<td>(2.06)</td>
<td>(1.82)</td>
<td></td>
</tr>
<tr>
<td>tbl</td>
<td></td>
<td>2.25**</td>
<td>5.01**</td>
<td>4.86**</td>
<td>2.90***</td>
<td>4.64**</td>
<td>2.88</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-2.30)</td>
<td>(-2.25)</td>
<td>(-2.08)</td>
<td>(-3.15)</td>
<td>(-2.37)</td>
<td>(-1.16)</td>
<td></td>
</tr>
<tr>
<td>tms</td>
<td></td>
<td>1.50*</td>
<td>6.94**</td>
<td>11.95***</td>
<td>0.72</td>
<td>1.03</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.82)</td>
<td>(2.44)</td>
<td>(3.35)</td>
<td>(1.28)</td>
<td>(1.02)</td>
<td>(-0.06)</td>
<td></td>
</tr>
<tr>
<td>dfy</td>
<td></td>
<td>0.18</td>
<td>0.25</td>
<td>0.42</td>
<td>0.31</td>
<td>0.18</td>
<td>0.26</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.05)</td>
<td>(0.24)</td>
<td>(0.42)</td>
<td>(0.72)</td>
<td>(0.05)</td>
<td>(0.19)</td>
<td></td>
</tr>
<tr>
<td>lty</td>
<td></td>
<td>1.49*</td>
<td>1.45</td>
<td>0.40</td>
<td>3.21***</td>
<td>5.24***</td>
<td>5.59*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-1.80)</td>
<td>(-1.09)</td>
<td>(-0.34)</td>
<td>(-3.90)</td>
<td>(-2.93)</td>
<td>(-1.84)</td>
<td></td>
</tr>
<tr>
<td>ntis</td>
<td></td>
<td>0.26</td>
<td>0.66</td>
<td>0.63</td>
<td>0.31</td>
<td>0.35</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-0.34)</td>
<td>(-0.42)</td>
<td>(-0.46)</td>
<td>(-0.66)</td>
<td>(-0.35)</td>
<td>(-0.30)</td>
<td></td>
</tr>
<tr>
<td>infl</td>
<td></td>
<td>1.40</td>
<td>5.65***</td>
<td>3.27***</td>
<td>1.03**</td>
<td>0.18</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-1.38)</td>
<td>(-3.07)</td>
<td>(-3.38)</td>
<td>(-1.98)</td>
<td>(0.09)</td>
<td>(0.76)</td>
<td></td>
</tr>
<tr>
<td>dfr</td>
<td></td>
<td>0.71</td>
<td>0.21</td>
<td>0.18</td>
<td>0.25</td>
<td>0.22</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.22)</td>
<td>(0.45)</td>
<td>(0.13)</td>
<td>(0.74)</td>
<td>(0.73)</td>
<td>(-0.33)</td>
<td></td>
</tr>
<tr>
<td>ltr</td>
<td></td>
<td>0.48</td>
<td>1.21***</td>
<td>0.54**</td>
<td>0.33</td>
<td>0.18</td>
<td>0.31</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.07)</td>
<td>(2.80)</td>
<td>(2.00)</td>
<td>(0.93)</td>
<td>(-0.09)</td>
<td>(-0.98)</td>
<td></td>
</tr>
</tbody>
</table>
6.3 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, we explore the covariance structure between market return innovations and our proxies for the level of and uncertainty about idiosyncratic risk. Empirically, we 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, we estimate model-free impulse response functions. Our method closely relates to the local projection method of Jorda (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, we 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 \( y_t \), our definition of the impulse response is

\[
\begin{align*}
\begin{align*}
E[y_{t+k}|r_{t+1}] & = v, F_t - E[y_{t+k}|r_{t+1}] - E[r_{t+1}] = 0, F_t
\end{align*}
\end{align*}
\]

for different values of \( k \). Given a set of conditioning variables, \( z_t \), we decompose \( y_{t+k} \) as

\[
y_{t+k} = \text{proj}(y_{t+k}|z_t) + [E_t(y_{t+k}) - \text{proj}(y_{t+k}|z_t)] + \epsilon_{t:t+k} = \beta_k z_t + \xi_t + \epsilon_{t:t+k}, \tag{42}
\]

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. If the conditional mean of returns takes the linear form

\[
r_{t+1} = \gamma' z_t + v_{t+1}, \tag{43}
\]

where \( v_{t+1} \) has mean zero. The impulse response may then be rewritten as \( E[\epsilon_{t:t+k}|v_{t+1} = v] \). One obtains consistent estimates of \( \xi_t + \epsilon_{t:t+k} \) 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, we estimate \( \text{proj}(\epsilon_{t:t+k}|v_{t+1} = 1) \equiv \alpha_h \) by regressing \( \hat{\xi}_t + \hat{\epsilon}_{t:t+k} \) on \( \hat{v}_{t+1} \). Inference is straightforward, since the estimate of \( \alpha_h \) 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 \( \xi_t \), so misspecification of the conditional mean adds noise to the dependent variable \( \hat{\xi}_t + \hat{\epsilon}_{t:t+k} \) of the second stage regression. As long as we have estimated the return innovation correctly, we need not have...
Figure 8: 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. We calculate Newey-West standard errors, where the number of lags equals the horizon minus 1.

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 forecasting method places practical constraints on the maximum lag length which can be considered.

Figure 8 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 our 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. 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, we cannot rule out its presence, since our regression-based test likely has very low power to detect news about an extremely persistent component if its variance is sufficiently small.

Next, we consider the responses of three variables which are tightly linked with cross-sectional skewness: real compensation growth, employment growth, and the skewness index from model 6 from Panel B of Table 2.67 All three pictures are qualitatively similar. The responses are small and insignificant on impact, generally peaking 3 to 4 quarters after the return innovation is observed. The responses become statistically insignificant around 7-8 quarters later, suggesting today’s innovation is most informative about labor market conditions over the next 18-24 months. The point estimates are slightly negative in quarters 9-12. Such a result is suggestive of the existence of a transitory shocks which reverse themselves, though the statistical evidence is relatively weak.68

As was the case with consumption growth, the associated magnitudes for real compensation growth are relatively small. The cumulative response of real compensation growth over the first year is about 80 basis points, or about 30% of its annualized standard deviation. Integrating over the next two quarters adds an additional 40 basis points, which subsequently reverses itself from quarter 7-12. While positive return innovations signal good news about the mean of future compensation growth, the size of the identified shock does not appear to be sufficiently large to generate a large risk premium within a representative agent context. However, as we will demonstrate below, relatively small shocks to the conditional mean can be associated with large risk premia if the aggregate shocks are highly concentrated among a small fraction of individuals.

Finally, we plot the responses of our labor market uncertainty proxies, initial claims for unem-

---

67Model 6 had the best performance among our bivariate specifications. In our earlier estimation, we used the series from model 5 because it was less noisy, while still capturing the salient business cycle variation of model 6. Here, measurement error is less of a concern, so we go with the best-fitting model. Theoretically, it is more attractive because shocks to wages can have a permanent component, whereas all shocks to employment growth must ultimately be temporary if the unemployment rate is stationary. At longer horizons, responses from model 6 are more likely to be informative. The picture for Model 5 closely resembles the response of employment growth.

68If we 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.
ployment and the cross-sectional volatility measure $Vol_t$, to return innovations. The responses of both measures are hump-shaped and unambiguously negative. A positive return innovation is associated with a decrease in future labor market uncertainty, where the news is most informative about labor market uncertainty 6-18 months in the future. Here, the magnitudes are fairly substantial, given that both uncertainty measures are quite persistent.

7 Quantitative Model

Previously, we demonstrated how to integrate our incomplete markets mechanism into a general, jump diffusion model for aggregate cash flows. In this section, we deliberately shut down almost all sources of aggregate risk so as to highlight the potential importance of incomplete markets, emphasizing a stylized model with a single state variable. Despite its simplicity, we find that this model is capable of matching the time series properties of the data quite well.

7.1 Setup

We perturb the representative agent model so that agents are exposed to idiosyncratic, uninsurable event risk. For parsimony, we abstract away from diffusion (Gaussian) shocks and assume that all uninsurable risk comes from compound Poisson shocks. Further, while our empirical results provide evidence of state dependence in both tails of the idiosyncratic risk distribution, our stylized model only emphasizes downside risk in order to minimize on the number of free parameters. As such, the novel mechanism in our model is the inclusion of a time-varying probability of uninsurable idiosyncratic disasters within an otherwise standard endowment economy.

In our model, jump component, $J_{\eta,t+1}$ is driven by a single source of risk ($M = 1$), which captures the risk of large downward adjustments—“falling off the ladder”—driven by events such as job loss. A sufficient statistic for the higher moments of $\eta_{t+1}$ is the (univariate) Poisson intensity, $\lambda_{\eta,t+1}$, i.e. the personal disaster probability. $\lambda_{\eta,t+1} \equiv \bar{\lambda} \cdot x_t$, where $x_t$ follows a square root process,

$$x_{t+1} = 1 - \rho x + \rho x x_t + \sqrt{x_t} \phi x z_{x,t+1}, \quad (44)$$

where we have normalized $E[x_t] = 1$. $\bar{\lambda}$ equals the average Poisson intensity. Given that $\lambda_{\eta,t}$ is very close to zero, it equals the disaster probability to a first approximation. We also assume

\[ \text{An i.i.d diffusion component primarily affects the risk-free rate and thus has little effect on excess returns. Adding such a component is similar to changing the discount rate} \delta. \]
that the distribution of jumps from $t$ to $t+1$ is known at time $t$, which shuts off the direct effect discussed above. We assume that jump sizes are normally distributed with mean $\mu_b$ and variance $\sigma_b^2$. In our calibration, $\mu_b$ is a large, negative number, and $\sigma_b^2$ is fairly large relative to uncertainty about aggregate consumption. Analogously with rare macroeconomic disasters, these infrequent labor market events are associated with extremely high marginal utilities, allowing them to have a disproportionately large impact on asset prices despite their relative infrequency.

In order to make as stark of a contrast with the representative agent model as possible, aggregate consumption is unpredictable and homoskedastic. Shocks to dividend growth are also homoskedastic, but dividend growth is assumed to be predictable.\textsuperscript{70} Specifically, aggregate dynamics evolve according to

$$
\begin{bmatrix}
\Delta c_{t+1} \\
\Delta d_{t+1} \\
\Delta x_{t+1}
\end{bmatrix} =
\begin{bmatrix}
\mu_c & 0 & 0 \\
\mu_d - \phi & 0 & 0 \\
1 - \rho_x & 0 & \rho_x
\end{bmatrix}
\begin{bmatrix}
\Delta x_t \\
\Delta d_t \\
x_t
\end{bmatrix} + G_{y,t}z_{y,t+1},
$$

(45)

where

$$
G_{y,t}G'_{y,t} =
\begin{bmatrix}
\varphi_e^2 & \varphi_e \varphi_d \chi & 0 \\
\varphi_e \varphi_d \chi & \varphi_d^2 & 0 \\
0 & 0 & \varphi_x^2 x_t
\end{bmatrix}.
$$

(46)

The parameters $\mu_c$ and $\mu_d$ govern the means of $\Delta c_t$ and $\Delta d_t$, respectively. We set $\phi < 0$ to capture the intuition that firms may be more reluctant and/or unable to raise dividends when labor market uncertainty / the risk of job loss is high. We also allow for a contemporaneous correlation between $\Delta c_t$ and $\Delta d_t$.

From Proposition 3, the equity premium in the stylized model equals

$$
\log E_t[R_{m,t}] - r_{f,t} = \gamma \chi \varphi_c \varphi_d + (1 - \theta)\rho_c A_x \varphi_d x_t.
$$

(47)

We give explicit expressions for $A_x$ and $A_{d,x}$ in the Appendix, but, in our calibration with $\gamma > \psi > 1$, both $A_x$ and $A_{d,x}$ are both negative. Since aggregate consumption is i.i.d., it immediately follows that, in the representative agent model, $A_x = 0$ and the equity premium equals $\gamma \chi \varphi_c \varphi_d$, which is about 1.7% per annum in our calibration. Any residual increase in the equity premium is solely driven by incomplete markets and idiosyncratic disaster risk.

Our square root specification implicitly assumes that uncertainty about idiosyncratic risk is

\textsuperscript{70}For purposes of generating the pricing kernel, we only need to keep track of the claim on aggregate dividends, so we suppress $k$ subscripts.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma$</td>
<td>10</td>
<td>Relative risk aversion coefficient</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.9950</td>
<td>Rate of time preference</td>
</tr>
<tr>
<td>$\psi$</td>
<td>1.5</td>
<td>Intertemporal elasticity of substitution</td>
</tr>
<tr>
<td>$\phi$</td>
<td>-0.0125</td>
<td>Loadings of expected dividend growth on $x_t$</td>
</tr>
<tr>
<td>$\rho_x$</td>
<td>0.99</td>
<td>Persistence of $x_t$ process</td>
</tr>
<tr>
<td>$\mu_c$</td>
<td>0.00125</td>
<td>Drift of consumption growth</td>
</tr>
<tr>
<td>$\mu_d$</td>
<td>0.00074</td>
<td>Drift of dividend growth</td>
</tr>
<tr>
<td>$\varphi_c$</td>
<td>0.008</td>
<td>Standard deviations of shock to $\Delta c_t$</td>
</tr>
<tr>
<td>$\varphi_d$</td>
<td>0.032</td>
<td>Standard deviations of shock to $\Delta d_t$</td>
</tr>
<tr>
<td>$\chi$</td>
<td>0.55</td>
<td>Correlation between $z_{c,t+1}$ and $z_{d,t+1}$</td>
</tr>
<tr>
<td>$\bar{\lambda}$</td>
<td>0.04/12</td>
<td>Average idiosyncratic jump intensity</td>
</tr>
<tr>
<td>$\mu_b$</td>
<td>-0.15</td>
<td>Average consumption decline given a disaster</td>
</tr>
<tr>
<td>$\sigma_b$</td>
<td>0.1</td>
<td>Standard deviation of disaster magnitude</td>
</tr>
</tbody>
</table>

Table 9: 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 monthly. The additional free parameters, $\mu_\eta$ and $F_\eta$, are assumed without loss of generality to equal the expressions given in Proposition 1.

high when the probability of a bad shock is high. It also guarantees, under suitable parametric restrictions, that the Poisson intensity is nonnegative in the continuous time limit. As was demonstrated earlier, such an assumption is also supported by the data. This aspect of our model is symmetric with Wachter (2011)’s assumption within the context of a model emphasizing rare macroeconomic disasters. In addition, while this simple structure abstracts away from higher frequency lead-lag dynamics between dividends and extensive margin measures, it does capture the basic intuition that dividends fall when unemployment risk is high. Our baseline specification also shuts off compound Poisson shocks to the aggregate state vector in the interest of parsimony.

### 7.2 Calibration and Results

Table 9 provides an overview of the parameters in the quantitative model, along with the calibrated values. Our choices of $\gamma = 10$ and $\psi = 1.5$ are relatively standard choices in the literature.
with Epstein-Zin preferences. The discount factor $\delta$ is chosen in order to generate a real risk free rate of 1% per annum.$^{71}$ The persistence of $x_t$ is 0.99, which is very close to the monthly autocorrelation of initial claims for unemployment. We set the correlation $\chi$ equal to the correlation between $\Delta c_t$ and $\Delta d_t$ in the annual data. $\varphi_x$, $\varphi_d$ and $\phi$ were chosen to approximately match the volatility of the market return and annual dividend growth, as well as the level of the dividend yield.

A different way to assess the reasonableness of our assumption on $\varphi_x$ is to compare the mean and minimum values of initial claims. In the data, the minimum value is 0.024, which is 1.4 standard deviations below the mean value of initial claims, which is 0.039. Our specification implies that the minimum is 2 standard deviations below the mean.

Finally, we choose idiosyncratic risk parameters with an eye towards conservatism. Both the probability ($\bar{\lambda}$) and severity ($\mu_b$ and $\sigma_b^2$) are considerably lower than the values implied by the income data. Stockholders have some means with which to smooth their consumption over time, so we are hesitant to assume that income shocks translate one-for-one into consumption shocks. This is particularly likely to be true in households with multiple earners. We believe that our choices for these parameters, which we continue to refine in subsequent analysis, likely understate the true level of risk faced by labor market participants.

Table 10 demonstrates the ability of the quantitative model to match a number of key asset pricing moments. Data moments are taken from Bansal, Kiku, and Yaron (2012), who calculate statistics using annual time series of real returns and cash flow growth rates from 1930-2008. We refer the reader to their paper for further details about the underlying data sources. Next, we use the model to simulate 20,000 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, we note that our model with idiosyncratic disaster risk generates a large and time-varying equity premium of about 6% per year. It easily replicates the excess volatility puzzle; the volatility of the market return is double that of dividend growth. Most other values, including the volatility of the price-dividend ratio, are well within the model-implied confidence bounds. However, it is worth noting that the first order autocorrelation of consumption and dividend growth are significantly lower and higher than the corresponding values in the data, respectively. The former is by construction, while the latter is somewhat more troubling.

$^{71}$ Matching observed risk-free rates the presence of idiosyncratic risk generally necessitates the use of lower discount factors relative to standard choices in representative agent models.
<table>
<thead>
<tr>
<th>Moment</th>
<th>Data Estimate</th>
<th>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>$E[\Delta c]$</td>
<td>1.93</td>
<td>1.47</td>
<td>0.91</td>
<td>0.99</td>
<td>1.27</td>
<td>1.69</td>
<td>2.04</td>
<td>2.14</td>
</tr>
<tr>
<td>$\sigma(\Delta c)$</td>
<td>2.16</td>
<td>2.25</td>
<td>1.86</td>
<td>1.92</td>
<td>2.11</td>
<td>2.37</td>
<td>2.57</td>
<td>2.63</td>
</tr>
<tr>
<td>AC1($\Delta c$)</td>
<td>0.45</td>
<td>0.23</td>
<td>0.02</td>
<td>0.06</td>
<td>0.16</td>
<td>0.30</td>
<td>0.38</td>
<td>0.41</td>
</tr>
<tr>
<td>$E[\Delta d]$</td>
<td>1.15</td>
<td>0.97</td>
<td>-7.00</td>
<td>-5.42</td>
<td>-1.37</td>
<td>3.43</td>
<td>6.16</td>
<td>6.86</td>
</tr>
<tr>
<td>$\sigma(\Delta d)$</td>
<td>11.05</td>
<td>10.95</td>
<td>8.86</td>
<td>9.15</td>
<td>10.08</td>
<td>11.98</td>
<td>13.99</td>
<td>14.80</td>
</tr>
<tr>
<td>AC1($\Delta d$)</td>
<td>0.21</td>
<td>0.45</td>
<td>0.20</td>
<td>0.24</td>
<td>0.36</td>
<td>0.53</td>
<td>0.67</td>
<td>0.71</td>
</tr>
<tr>
<td>Corr($\Delta c, \Delta d$)</td>
<td>0.55</td>
<td>0.46</td>
<td>0.21</td>
<td>0.26</td>
<td>0.39</td>
<td>0.53</td>
<td>0.61</td>
<td>0.63</td>
</tr>
<tr>
<td>$E[R_m]$</td>
<td>7.66</td>
<td>7.10</td>
<td>0.97</td>
<td>1.92</td>
<td>5.14</td>
<td>9.11</td>
<td>11.64</td>
<td>12.29</td>
</tr>
<tr>
<td>$\sigma(R_m)$</td>
<td>20.28</td>
<td>22.47</td>
<td>18.63</td>
<td>19.21</td>
<td>21.05</td>
<td>23.93</td>
<td>26.10</td>
<td>26.87</td>
</tr>
<tr>
<td>$E[R_f]$</td>
<td>2.86</td>
<td>1.13</td>
<td>-1.80</td>
<td>-1.17</td>
<td>0.29</td>
<td>1.89</td>
<td>2.83</td>
<td>3.01</td>
</tr>
<tr>
<td>$\sigma(R_f)$</td>
<td>0.57</td>
<td>1.13</td>
<td>-1.80</td>
<td>-1.17</td>
<td>0.29</td>
<td>1.89</td>
<td>2.83</td>
<td>3.01</td>
</tr>
<tr>
<td>Corr($R_m, R_f$)</td>
<td>0.70</td>
<td>6.03</td>
<td>2.03</td>
<td>2.63</td>
<td>4.64</td>
<td>7.39</td>
<td>9.21</td>
<td>9.78</td>
</tr>
<tr>
<td>$\sigma(pd)$</td>
<td>0.45</td>
<td>0.34</td>
<td>0.22</td>
<td>0.24</td>
<td>0.29</td>
<td>0.40</td>
<td>0.54</td>
<td>0.59</td>
</tr>
<tr>
<td>AC1(pd)</td>
<td>0.87</td>
<td>0.83</td>
<td>0.64</td>
<td>0.67</td>
<td>0.77</td>
<td>0.87</td>
<td>0.92</td>
<td>0.93</td>
</tr>
</tbody>
</table>

Table 10: Comparison of data with 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 20,000 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.

Nonetheless, the overall performance of the model is quite strong, considering its highly stylized nature. Importantly, our results obtain without relying on extremely persistent volatility, as in Bansal, Kiku, and Yaron (2012), who select a monthly autocorrelation coefficient of 0.999. Nor does it rely on assumptions about aggregate tail risk, which are more difficult to verify.

8 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. Our analysis suggests that risks associated with redistribution...
of consumption across across agents are likely to be just as important, if not more important, than aggregate consumption risks. We view our 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. Our model can be easily embedded within a production setting, and we plan to explore these interactions in future work.

In the data, aggregate and idiosyncratic risks are tightly linked with one another. While our general model easily accommodates the study of these interactions, we deliberately downplay risks associated with aggregate consumption so as to highlight the potential of the incomplete markets mechanism. Our 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 us to address a larger number of policy questions.

Our 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 increases in the quality of panels of earnings records, we 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*.

9 Appendix [Under Construction]

9.1 Proofs of Propositions

9.1.1 Proof of Proposition 1 (Aggregation Restrictions)

Using the independence on $z_{\eta,t+1}^i$ and $J_{\eta,t+1}^i$, it follows that
\[
E[\exp(\eta_{t+1}^i | y_{t+1}, y_t)] = \exp(1' \hat{\mu}_\eta + 1' F_\eta y_t + 1/2(h_{\eta0} + h'_{\eta1} y_t)) E[\exp(1' J_{\eta,t+1}^i | y_{t+1}, y_t)]
\]
\[
= \exp(1' \hat{\mu}_\eta + 1' F_\eta y_t + 1/2(h_{\eta0} + h'_{\eta1} y_t)) \times \exp(1' (l_{\eta0}(\Psi_\eta(1_M) - 1_M) + l_{\eta1} \otimes [(\Psi_\eta(1_M) - 1_M) \otimes 1']) y_t)),
\]
where we used the moment-generating function of the normal distribution and a compound Poisson process to go from the first to the second line. In order to satisfy Assumption 2.iii, 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.

9.1.2 Proof of Proposition 2 (Wealth-Consumption Ratios)

We will begin by solving for the wealth-consumption ratio coefficients, then proceed to solve for the price-dividend ratios. However, before working with the Euler Equations, we 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 1.** Let Assumptions 2 and 3.iii-iv hold. Then,
\[
E[\exp(u \cdot \eta_{t+1}^i | y_{t+1}, y_t)] = \exp[\beta_0(u) + \beta'(u)y_{t+1}],
\]
where $\beta_0(u): \mathbb{R} \rightarrow \mathbb{R}$ and $\beta(u): \mathbb{R} \rightarrow \mathbb{R}^L$ for $u \in \mathbb{R}$ are given by

(i) $\beta_0(u) = \mu'_{\eta} 1_M u + \frac{1}{2} u^2 h_{\eta0} + l_{\eta0}'(\Psi_\eta(u1_M) - 1_M)$ and
(ii) $\beta(u) = F'_\eta 1_M u + \frac{1}{2} u^2 H_{\eta1} + l_{\eta1}'(\Psi_\eta(u1_M) - 1_M)$.
Proof. By definition, \( \eta_{t+1} = 1_M \tilde{\eta}_{t+1} \). By the conditional independence of \( z_{\eta,t+1} \) and \( J_{\eta,t+1} \),

\[
\log E[\exp(u \cdot \eta_{t+1}) | y_{t+1}] = u_1' \mu_\eta + F_{\eta,y_{t+1}} + \log E[\exp(u_1' M_{\eta,t+1} z_{\eta,t+1}^i) | y_{t+1}]
\]

\[
\log E[\exp(u_1' M_{\eta,t+1} z_{\eta,t+1}^i) | y_{t+1}] = \frac{1}{2} u_1^2 M_{\eta,t+1} z_{\eta,t+1}^i 1_M = \frac{1}{2} u^2 [h_{\eta,0} + H_{\eta} y_{t+1}]
\]

\[
\log E[\exp(u_1' M_{\eta,t+1} J_{\eta,t+1}) | y_{t+1}] = \lambda_{\eta,t+1}' [\Psi_{\eta}(u1_M) - 1_M] = [l_{\eta 0} + l_{\eta 1} y_{t+1}] [\Psi_{\eta}(u1_M) - 1_M],
\]

where we 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.

\[\square\]

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

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

where \( f(u) : \mathbb{R}^L \to \mathbb{R} \) and \( g(u) : \mathbb{R}^L \to \mathbb{R}^L \) 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) \( g(u) = F_y u + \frac{1}{2} [u' H_{yi} u]_{i \in \{1,\ldots,L\}} + l_{y1}' (\Psi_y(u) - 1_L) \)

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

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

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

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

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

As before, we use moment generating functions to evaluate expectations, plug in the functional forms from Assumptions 3.1i-ii, then collect terms. See also DY section A.1.

\[\square\]

We will assume that the wealth-consumption ratio \( w_c = A_0 + A'y_t \). By Assumptions 1-2, we can write consumption growth in vector notation as \( \Delta c_t^i = S_i^t y_t + \eta_i^t \). Combining (21) with our
assumption, the log-linearized Euler equation for the consumption claim is

\begin{align*}
1 &= E_t \left[ \exp \{ \theta \log \delta + \theta \kappa_c + \theta (\rho_c - 1) A_0 + (1 - \gamma)(S_c y_{t+1} + \eta_{t+1}^1) + \theta (\rho_c A' y_{t+1} - A' y_t) \} \right] \\
&= \exp(\theta \log \delta + \theta \kappa_c + \theta (\rho_c - 1) A_0 - \theta A' y_t) E_t \left[ \exp \{ [(1 - \gamma) S'_c + \theta \rho_c A'] y_{t+1} + (1 - \gamma) \eta_{t+1}^1 \} \right] \\
&= \exp(\theta \log \delta + \theta \kappa_c + \theta (\rho_c - 1) A_0 - \theta A' y_t) \\
\times & E_t \left[ \exp \{ [(1 - \gamma) S'_c + \theta \rho_c A'] y_{t+1} + \log E_t [(1 - \gamma) \eta_{t+1}^1 | y_{t+1}] \} \right] \\
&= \exp(\theta \log \delta + \theta \kappa_c + \theta (\rho_c - 1) A_0 - \theta A' y_t) \\
\times & E_t \left[ \exp \{ [(1 - \gamma) S'_c + \theta \rho_c A'] y_{t+1} + \beta_0 (1 - \gamma) + \beta (1 - \gamma)' y_{t+1} \} \right],
\end{align*}

where the second and third equalities use the law of iterated expectations and the last line follows from Lemma 1. Using Lemma 2 to evaluate the expectation yields

\begin{align*}
\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.
\end{align*}

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

\begin{align*}
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) \quad (50) \\
g((1 - \gamma) S_c + \theta \rho_c A + \beta (1 - \gamma) &= \theta, \quad (51)
\end{align*}

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

In addition to the primitive parameters governing preferences and cash flows, the system (50-51) also depends on the log-linearization constants \( \kappa_c \) and \( \rho_c \). Following DY and Eraker and Shaliastovich (2008), we choose the linearization point to equal the unconditional mean of the wealth-consumption ratio. In particular, we choose \( \bar{wc} \) so that

\begin{align*}
\log(\rho_c) - \log(1 - \rho_c) = \bar{wc} = E(wc_t) = A_0 + A' E(y_t), \quad (52)
\end{align*}

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

\begin{align*}
\kappa_c + (\rho_c - 1) A_0 &= -\log \rho_c - (\rho_c - 1) A' E(y_t). \quad (53)
\end{align*}
We can then substitute (53) into (50), yielding
\[
 f((1 - \gamma)S_c + \theta \rho_c A + \beta(1 - \gamma)) + \beta_0 (1 - \gamma) = -\theta (\log \delta - \log \rho_c - (\rho_c - 1)A'E(y_t)),
\] (54)
leaving (51) and (54), an exactly identified system of equations in \(A\) and \(\rho_c\). Then, given these solutions, we can use the expressions above to derive \(A_0, \kappa_c,\) and \(\kappa\).

We will assume that the price-dividend ratio for asset \(k\), \(pd_{k,t} = A_{0,k} + A_k'y_t\). By Assumption 1, we can write dividend growth as \(\Delta d_{kt} = S_k'y_t\). Since the dividend claims are financial assets, we can price them using the projected pricing kernel in (23). Plugging in the projected kernel, the log-linearized Euler equation for the \(k^{th}\) dividend claim is
\[
1 = \exp\left[\kappa - (1 - \theta)(\rho_c - 1)A_0 + \kappa_k + (\rho_k - 1)A_{0,k} + (1 - \theta)A_k'y_t - A_k'y_t\right] \\
\times E_t \left[\exp \left\{\left[-\Lambda' + S_k'y_t + \rho_k A_k'y_t\right]y_{t+1}\right\}\right].
\] As before, using Lemma 2 to evaluate the expectation and taking logs yields the \((L + 1)\)-dimensional system of equations
\[
 f(-\Lambda + S_k + \rho_k A_k) = -[\kappa - (1 - \theta)(\rho_c - 1)A_0 + \kappa_k + (\rho_k - 1)A_{0,k}]
\] (55)
\[
g(-\Lambda + S_k + \rho_k A_k) = A_k - (1 - \theta)A.
\] (56)

Once again, we choose the linearization constants in order to linearize around the unconditional mean log price-dividend ratio. In order to obtain a more accurate solution, we allow the linearization constants \(\kappa_k\) and \(\rho_k\) to differ across assets. This amounts to replacing equation (55) with
\[
 f(-\Lambda + S_k + \rho_k A_k) = -[\kappa - (1 - \theta)(\rho_c - 1)A_0 - \log \rho_k - (\rho_k - 1)A_k'E(y_t)].
\] (57)

9.1.3 Proof of Proposition 3 (Risk Premia)

Following DY, we decompose the projected pricing kernel and the return on a risky asset into jump and Gaussian components
\[
m_{t+1} = \kappa - (1 - \theta)(\rho_c - 1)A_0 + (1 - \theta)A_k'y_t - \Lambda' \left[\mu_y + F_y y_t + G_y z_{y,t+1}\right] + -\Lambda' \left[\mu_y + F_y y_t + G_y z_{y,t+1}\right] + -\Lambda' \left[\mu_y + F_y y_t + G_y z_{y,t+1}\right]
\equiv m_{t+1}^{m}
\]
\[
r_{k,t+1} = \kappa_k + (\rho_k - 1)A_{0k} - A_k'y_t + B_k'(\mu_y + F_y y_t + G_y z_{y,t+1}) + B_k' \left[\mu_y + F_y y_t + G_y z_{y,t+1}\right]
\equiv r_{k,t+1}^{q}
\]
DY show that the risk premium may be decomposed as

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

\[
+ \log(E_t[\exp(m^J_{t+1})]) - \log(E_t[\exp(r^J_{k,t+1} + m^J_{t+1})])
\]

\[
= B'_k G_{y,t} G'_y \Lambda + \lambda'_y \left[ \Psi_y(B_k) - 1 \right] - \lambda'_y \left[ \Psi_y(B_k - \Lambda) - \Psi_y(-\Lambda) \right].
\]

See DY, section A.4, for further details.

9.1.4 Proof of Proposition 4 (Term Structure)

We will begin by establishing the result for the dividend claim, then proceed to the consumption claim. The proof is by induction. First, we establish that

\[
pd^h_{k,t} = A^h_{0,k} + (A^h_k)'y_t.
\]

By our no arbitrage restriction that

\[
D_{k,t} = P^0_{k,t},
\]

implying that

\[
pd^h_{k,t} = 0
\]

and thus that

\[
A^0_{0,k} = 0
\]

and

\[
A^0_k = 0.
\]

Next, we must show that

\[
pd^h_{k,t} = A^h_{0,k} + (A^h_k)'y_t
\]

implies

\[
pd^{h+1}_{k,t} = A^{h+1}_{0,k} + (A^{h+1}_k)'y_t.
\]

Combining the Euler equation with (??) yields

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

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

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

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

where

\[
m_0 \equiv \kappa - (1 - \theta)(\rho_c - 1)A_0.
\]

Matching coefficients yields the recursions

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

(58)

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

(59)

which establishes the claim. One obtains coefficients for the real risk-free asset, by setting

\[
S_k = 0
\]

in (58-59). Analogously, the coefficients for expected real dividend growth are obtained by setting

\[
m_0 = 0
\]

and

\[
\Lambda = 0.
\]

Next, we turn to the consumption claim. The only substantive difference is that, since the return on the consumption claim depends on

\[
\eta^i_{t+1},
\]

we cannot use the projected version of the pricing kernel. Instead, we 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, we show that \( wc^h_t = A^h_0 + (A^h)'y_t \) implies \( wc^{h+1}_t = A^{h+1}_0 + (A^{h+1})'y_t \):

\[
\begin{align*}
wc^{h+1}_t &= \log E_t[\exp\{m^i_{t+1} + \Delta c_{t+1} + \eta^i_{t+1} + wc^h_{t+1}\}] \\
A^{h+1}_0 + (A^{h+1})'y_t &= \log E_t[\exp\{m^i_{t+1} + \eta^i_{t+1} + A^h_0 + (S_c + A^h)'y_{t+1}\}] \\
&= \theta \log \delta - (1 - \theta)(\kappa_c + (\rho_c - 1)A_0) + A^h_0 + (1 - \theta)A'y_t \\
&\quad + \log E_t[\exp\{[(1 - \gamma)S_c - (1 - \theta)\rho_c A + A^h]'y_{t+1} + (1 - \gamma)\eta^i_{t+1}\}] \\
&= \tilde{m}_0 + A^h_0 + (1 - \theta)A'y_t \\
&\quad + \log E_t[\exp\{[(1 - \gamma)S_c - (1 - \theta)\rho_c A + \beta(1 - \gamma) + A^h]'y_{t+1}\}] \\
&= \tilde{m}_0 + A^h_0 + f(S_c + A^h - \tilde{\Lambda}) + [g(S_c + A^h - \tilde{\Lambda}) + (1 - \theta)A'y_t]
\end{align*}
\]

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

\[
\begin{align*}
A^{h+1}_0 &= \tilde{m}_0 + A^h_0 + f(S_k + A^h_k - \tilde{\Lambda}) \\
A^{h+1} &= g(S_k + A^h_k - \tilde{\Lambda}) + (1 - \theta)A,
\end{align*}
\]

so the only substantive difference between (58-59) and (60-61) comes from the definitions of \( \tilde{m}_0 \) and \( \tilde{\Lambda} \). Further note that \( \tilde{m}_0 = m_0 - \beta_0(1 - \gamma) + \beta_0(1 - \gamma) \) and \( \tilde{\Lambda} = \Lambda - \beta(-\gamma) + \beta(1 - \gamma) \), so the difference between the recursions comes entirely from the projection terms.

### 9.1.5 Risk-free rate

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

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

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

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

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