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

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

This version: October 2014
First version: June 2013

Abstract

Discrete events, such as transitions between jobs, often induce changes in wages that can be quite large, highly persistent, and largely uninsurable, effectively exposing labor market participants to *idiosyncratic* tail risk. The nature of these events is highly cyclical; transitions are more likely to be favorable if they occur during expansions relative to recessions. In incomplete markets, agents require a premium to invest in assets which underperform when labor market event risk is high, a feature absent from leading asset pricing models. I provide new empirical evidence on the plausibility of event risk in explaining the shape of the idiosyncratic distribution of income growth rates and its evolution over time. Next, I formalize its role within a general affine, jump-diffusion asset pricing framework with heterogeneous agents and incomplete markets, making my results immediately applicable to a wide class of existing models for aggregate dynamics. In addition, I propose a model where agents are exposed to a small, time-varying probability of experiencing a rare, idiosyncratic disaster. The model, whose key parameters are calibrated to new data on the cross-sectional distribution of labor income, quantitatively matches the level and dynamics of the equity premium. 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 92093, lschmidt@ucsd.edu. I am particularly indebted to my advisor, Allan Timmermann, for his help and support throughout the writing of this paper. I am also grateful to seminar participants at the SITE conference on "Interrelations between Labor Markets and Financial Markets", the UCSD applied, macroeconomics, and finance workshops, Brendan Beare, Jonathan Berk, Jules Van Binsbergen, John Campbell, Steven Davis, Martin Eichenbaum, Marjorie Flavin, James Hamilton, Simon Gilchrist, Fatih Guvenen, Bryan Kelly, Ralph Koijen, Arthur Korteweg, David Lagakos, Hanno Lustig, Serdar Ozkan, Chris Parsons, David Romer, Martin Schneider, Frank Schorfheide, Eric Swanson, 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.[1] 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.[2]

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.[3] 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.[4] These events resemble rare but *idiosyncratic* disasters, since they need not be accompanied by large declines in aggregate consumption or output.[5] 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

---

[1] See, *e.g.* Gourio (2012), Gourio (2013), Gabaix (2012), Wachter (2013), Tsai and Wachter (2013).

[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.

I propose and test a number of implications of a model where predictable changes in idiosyncratic tail risk induce predictable changes in risk premia. First, I 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. My analysis builds upon statistics from Guvenen, Ozkan, and Song (2014c, 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. My 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.

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

Second, I propose a tractable asset pricing framework which integrates heterogeneous agents, incomplete markets, and state-dependent cross-sectional consumption moments into a Lucas (1978) endowment economy. The key mechanism is that the shape of the distribution of idiosyncratic shocks to consumption growth is linked to aggregate state variables. I solve the model with arbitrary, jump diffusion 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.[6]

To complement my solution for a fully-specified endowment-based model, I also derive an intertemporal capital asset pricing (ICAPM) representation of the stochastic discount factor from my incomplete markets economy, 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 I emphasize labor income throughout, my asset pricing framework provides a tractable way of pricing risks associated with the *redistribution* of wealth more generally. These shocks could also come from households' idiosyncratic exposures to firms' capital–e.g. private equity and entrepreneurial investments. In my framework, redistribution risk varies over time and enters as a priced state variable. Moreover, this incomplete markets mechanism, which is largely absent in production-based asset pricing models, is likely to generate an amplified response to aggregate shocks. If unfavorable redistributions become more likely when productivity is low and/or uncertainty is high (e.g. because default risk is higher), the associated increase in discount rates will affect firms' incentives to invest.[7]

Third, I test one of the key implications of my incomplete markets model, namely that the equity premium is high when labor market uncertainty is high. I 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. Using a semi-parametric method, I show that stock returns are highly informative about future labor market conditions, whereas they convey little information about aggregate consumption growth.

[6]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).

[7]For example, a recent literature emphasizes the link between uncertainty and economic growth. In representative agent models, uncertainty affects risk premia indirectly (e.g. by changing the distribution of aggregate consumption. In my model, 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.

The novel mechanism is that agents are exposed to rare, idiosyncratic disasters, where the idiosyncratic disaster probability is time-varying. In this section, I work with a fairly standard specification for aggregate risk so as to highlight the amplification in risk premia associated with incomplete markets.

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

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. My 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").[8]

GOS, using a larger panel of income records from the Social Security Administration, find little

[8]Beginning 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.

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.[9] My analysis builds heavily on calculations in GOS, and I discuss their findings in greater detail in the next section of the paper. While I emphasize time-varying skewness throughout, my 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

[9] Davis 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.

[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).

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 operates in models which generate large risk premia with persistent variation in the probability or severity of experiencing a macroeconomic disaster.

My theoretical model also features long-run risks, and the intuition behind it is quite similar. Moreover, my assumptions on aggregate cash flow dynamics are sufficiently general to encompass the vast majority of models considered to-date. I allow for additional dimensions of priced risk which do not affect the distribution of aggregate consumption. Quantitatively speaking, my measures are 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. I find that the price-dividend ratio has substantial predictive content for my idiosyncratic risk index.

While I 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.[14] 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.[15] 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.

[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.

[14]Berk and Walden (2013) 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. See also Guiso et al. (2005) for empirical evidence on risk-sharing via the labor market.

[15]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.

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. I also present evidence of procyclical 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 right tail of the income growth distribution. Thus, my 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 my framework, they assume that a representative agent can pool income from employed and unemployed workers.[16] 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 for market tightness, the key state variable in search-and-matching models.

Two contemporaneous working papers 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 my 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).[17] 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. 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.

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

[17] These empirical findings compliment mine, which are derived from the cross-sectional distribution of income growth rates. While CEX consumption measures have a more direct link with the theory, they require the use of much smaller sample sizes. Also, survey-based consumption measures are more susceptible to measurement errors relative to administrative earnings data.

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 90^{th} and 10^{th} 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.

A key difference between these two papers and mine is the way in which idiosyncratic risk is modeled. I 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. My process can generate arbitrarily high levels of skewness.

Moreover, my general solution clarifies the theoretical mechanisms at play. I 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, I 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, I also allow for a much richer specification of aggregate dynamics, which enables me to study the interactions between aggregate and idiosyncratic risk factors. To clarify the mechanics associated with time-varying idiosyncratic risk, I downplay these dynamics in my quantitative model. However, such an extension is likely to be useful going forward, as I find a 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. I 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 my general theoretical asset pricing model. Section 6 introduces my proxy for labor market uncertainty and tests the key implication of my model for return predictability. Section 7 presents my stylized, quantitative model, and Section 8 concludes.

2 Motivating Evidence

In this section, I briefly review several key implications of statistics calculated by GOS, whose calculations are used as the basis for my analysis. GOS report a number of statistics for the cross-section of real income growth rates and demonstrate how these distributions evolve over time. The key, highly robust result from their analysis is that the variance of idiosyncratic shocks is almost acyclical, while there is quantitatively important cyclical variation in skewness/asymmetry of the distribution. I highlight that this cyclical variation in asymmetry takes a very particular form; while the center of the distribution is relatively insensitive to the business cycle, its conditional tails are highly state-dependent.

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). The data provide uncapped (i.e. not top-coded), nominal annual earnings for each individual from 1978-2011, which are adjusted to real terms using the personal consumption expenditure deflator.[18] My calculations rely upon a number of statistics from GOS about the distribution of changes in log earnings at 1, 3, and 5-year horizons.

One obtains an intuitive measure of the asymmetry of a distribution by considering three conditional quantiles. A robust measure of skewness is Kelley's skewness, which is defined as $\frac{[Q90_t - Q50_t]-[Q50_t - Q10_t]}{Q90_t - Q10_t}$.[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, the denominator is almost constant. The numerator splits $Q90_t - Q10_t$ into two pieces. The first, $Q90_t - Q50_t$, measures the width of the right tail, while the latter, $Q50_t - Q10_t$ 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. These statistics pool all observations in their sample, giving a snapshot of the entire cross-sectional distribution of wage changes across the U.S. population. As the economy moves from an expansion to a recession, the left tail of the distribution ($Q50 - Q10$, in Panel A)

[18] According to GOS, the earnings data "include wages and salaries, bonuses, and exercised stock options as reported on the W-2 form (box 1)." See GOS for additional details about the data and sample selection criteria.

[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.

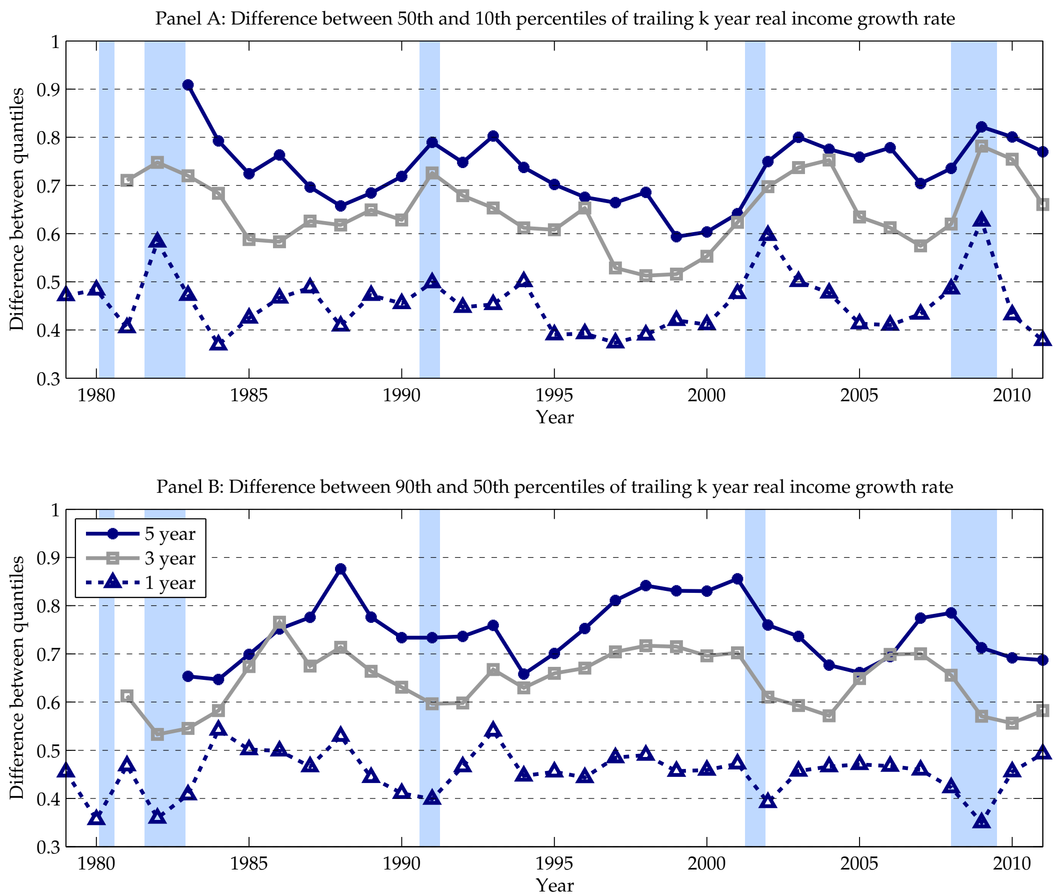

Figure 1: Dynamic evolution of the cross-section of income growth rates over time

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

expands, indicating an increased likelihood of experiencing large decreases in wages, while the width of the right tail ($Q90 - Q50$, in Panel B) shrinks. There are more big losers in recessions and fewer big winners. Note that my use of trailing growth rates in the graphs means that the long horizon measures will tend to lag the recession bars.

Table 1 summarizes a number of GOS's results on the distribution of 5-year wage changes, which control for cohort and life-cycle fixed effects and individuals' previous earnings.[21] 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 periods and recession periods, and compare the average levels of the different quantiles in expansions with those in recessions. In their classification, recession periods begin one year prior to the start of the recession and end several years after the recession has ended, in order to emphasize persistent changes in wages from recessions, as opposed to more temporary declines in income such as lost wages during unemployment spells.[22] Expansions are 5-year periods which do not include a recession year.

Given my interest in asset pricing, I 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. I summarize this information by averaging the reported statistics over the 91^{st} through 95^{th} percentiles of the lagged earnings distribution. These individuals have sufficiently high earnings that they are likely to participate in stock markets. However, labor income is still likely to be their primary source of non-housing wealth.[23] Section A.1 in the Appendix shows that the same results hold for different segments of the earnings distribution.

The left panel of Table 1 reports the median, 10^{th} percentile, 90^{th} percentile, and Kelley's skewness of five year income growth rates in expansions and recessions, respectively. The right panel reports the changes in quantile-based measures of scale–the inter-quartile range and the 90 -

[21] 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. 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 controlling for age. I refer the reader to GOS for further details on this procedure.

[22] 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.

[23] In 2011, GOS report that the 90th and 95th percentiles of wages are \$98k and \$135k per year, respectively, in 2005 dollars. GOS also emphasize the high cyclicality of the incomes of extremely high earners, particularly those in the top 1%. See also Guvenen, Kaplan, and Song (2014a). While the higher cyclicality is interesting, these individuals likely possess substantial financial wealth, making it difficult to characterize the extent to which income shocks translate to consumption. For example, existing evidence on partial insurance is unlikely to be representative of their behavior.

Statistic	Period	Value	Statistic	Period	Value
Median	E	1.53	Scale measures		
	R	-2.41	Inter-Quartile Range	R - E	1.30
	R - E	-3.81	90-10 Percentile Spread	R - E	-0.62
10th Percentile	E	-62.62	Left tail width measures		
	R	-74.23	50-25 Percentile Spread	R - E	3.01
	R - E	-11.51	50-10 Percentile Spread	R - E	7.70
90th Percentile	E	53.19	Right tail width measures		
	R	40.96	75-50 Percentile Spread	R - E	-1.72
	R - E	-11.93	90-50 Percentile Spread	R - E	-8.12
			95-50 Percentile Spread	R - E	-15.47
Kelley's Skewness	E	-10.85	99-50 Percentile Spread	R - E	-26.71
	R	-24.73			
	R - E	-13.89			

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. I report the average of each statistic over the 91^{st} through 95^{th} percentiles of the 5-year average income distribution (see GOS for a detailed definition) and over time. The second column indicates the period over which the average value of the statistic is calculated, where "E", "R", and "R - E" denote expansions, recessions, and the difference between recessions and expansions, respectively.

10 percentile spread–in recessions versus expansions. The right panel also reports 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.

Several features of Table 1 are particularly striking. First, high earners face a substantial degree of idiosyncratic labor income risk, even in expansions. The average 90-10 spread is 115 log percentage points. Second, the entire distribution shifts to the left in bad times; all of the quantiles are strictly lower in recessions relative to expansions. This shift is not specific to the three quantiles in the left panel. All of the reported quantiles are lower in recessions.

Third, the change in the 10^{th} and 90^{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 (as was the case in Figure 1. This cyclical asymmetry is reflected by the change in Kelley's skewness, which decreases by 14 percentage points. In contrast, both measures of the overall spread of the distribution changes very little over the cycle. GOS demonstrate that a similar result holds for second moments, particularly at longer horizons.

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

The results in Table 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 Conditional skewness index

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, I develop a mixed-frequency approach to extract an empirical proxy for the conditional skewness of the idiosyncratic income growth distribution from a large cross section of macroeconomic and financial time series. My skewness index is available at a higher frequency (quarterly) and is available over a longer time period, making it easier to understand its time series properties. In the final section, I calibrate my theoretical model to match its dynamics.

3.1.1 Statistical framework

My reduced-form model for labor income is a version of the canonical permanent income life-cycle model. Let w_t^i and age_t^i be individual i's log labor income (after subtracting common shocks and a deterministic life-cycle component) and age in period t, respectively. I assume

$$
\begin{aligned}
w_t^i &= \underbrace{\alpha_i + \beta_i age_t^i}_{\substack{\text{profile}\\\text{heterogeneity}}} + \underbrace{\xi_t^i}_{\substack{\text{permanent}\\\text{component}}} + \underbrace{\rho(L)\cdot\eta_t^i}_{\substack{\text{state dependent}\\\text{transitory shock}}} + \underbrace{\epsilon_t^i}_{\substack{\text{state independent}\\\text{transitory shock}}} && (1)\\
\xi_t^i &= \xi_{t-1}^i + \eta_t^i, \quad \eta_t^i|y_t \sim F_\eta(\eta;y_t), \quad \epsilon_t^i \perp y_t, \quad (\alpha_i,\beta_i) \overset{iid}{\sim} G(\alpha,\beta), && (2)
\end{aligned}
$$

where η_t^i is a shock that is independently distributed over time conditional on the aggregate state, which I assume can be characterized by a finite-dimensional random vector, y_t. α_i and β_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. ξ_t^i is the permanent component to wages. As I discuss below, my theoretical model requires that ξ_t^i be a random walk process. Alternatively, one could allow for ξ_t^i to follow a persistent, stationary process such as an AR(1). I impose this restriction throughout, noting that empirical estimates of the AR(1) parameter in other studies are generally close to 1.[24]

I also allow for a transitory component in labor income. The first term, which can depend on current and past η_t^i via the lag polynomial $\rho(L)$, allows permanent shocks to have additional temporary effects on measured income. For example, large negative realizations of η_t^i could be accompanied by unemployment spells, leading to temporary interruptions in the flow of labor income.[25] The second term, ϵ_t^i, 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 ϵ_t^i, I maintain this assumption for parsimony.[26]

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. I use macroeconomic time series which are sampled at a quarterly frequency in my analysis, then I derive the implications for the higher moments of idiosyncratic wage changes, where wages are measured annually. To do so, I make use of a simple log-linear approximation for time-aggregated wages.

[24] 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.

[25] Signing bonuses could generate similar effects for large positive shocks.

[26] My aggregation result goes through if, given the aggregate state, the third central moment of ϵ_t^i is constant.

Define $W^i_{A,t} = \sum_{j=0}^{3} \exp(W^i_{t-j})$ and $W^i_{A,t} \equiv \log W^i_{A,t}$, so that $W^i_{A,t}$ is a four quarter moving average of labor income, measured at the end of quarter t. A time series of annual wages can equivalently be expressed as a quarterly time series of $W^i_{A,t}$, where $W^i_{A,t}$ is only observed in the fourth quarter each year. In this notation, the year-on-year change in log wages is $w^i_{A,t} - w^i_{A,t-4}$. A first-order Taylor expansion yields that, for $k \geq 4$,

$$w^i_{A,t} - w^i_{A,t-k} \approx \tfrac{1}{4}\Delta w^i_t + \tfrac{1}{2}\Delta w^i_{t-1} + \tfrac{3}{4}\Delta w^i_{t-2} + \sum_{j=3}^{k-1} \Delta w^i_{t-j} + \tfrac{3}{4}\Delta w^i_{t-k} + \tfrac{1}{2}\Delta w^i_{t-k-1} + \tfrac{1}{4}\Delta w^i_{t-k-2}. \quad (3)$$

This approximation replaces an arithmetic mean with a geometric mean, a solution proposed in a related context by Mariano and Murasawa (2003).[27] If, in addition, the third central moment of η^i_t takes the affine form, for $k \geq 4$, $E[(\eta^i_t)^3 | y_t] = a + b' y_t$, then (1-3) imply

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

where the coefficients c_k and the lag polynomial $\phi_k(L;\rho)$ are derived and defined in Appendix A.2.1, and $\mathscr{F}_t$ is a filtration capturing aggregate information up to time t (which includes $\{y_{t-j}\}_{j=0}^{\infty}$). Crucially, $\phi_k(\cdot)$ is a *known* function that only depends upon the horizon k and the specification of state-dependent transitory risk $\rho(L)$.

Equation (4) says that, when the third central moment of the permanent shock η^i_t is affine in an observable vector y_t, under my assumptions on the income process, I 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 the third central moment of $w^i_{A,t} - w^i_{A,t-k}$ on a constant and $\phi_k(L;\rho) y_t$. Further, I can pool the information from time-aggregated moments which are measured over different horizons k, improving efficiency by imposing the cross-equation restriction that b for all k. 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 $E[(\eta^i_t)^3 | y_t] = a + b' y_t$? A sufficient condition is that the cumulant-generating function (the log of the moment-generating function) of η^i_t is linear in y_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

[27] See also Camacho and Perez-Quiros (2010) for a similar approach. Simulation results in Appendix A.2.2 demonstrate that these errors are negligible for the parametric model in Section 3.2, particularly over longer horizons. Henceforth, I ignore approximation errors, treating (3) as the "true model" for time-aggregated labor income measures.

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, sums of these processes also satisfy this property. For example, if η_t^i has a compound Poisson distribution with arrival intensity $\lambda_t = \lambda_0 + \lambda_1' y_t$, my regression recovers λ_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 Empirical implementation

Given a parametric specification for $\rho(L)$, (4) implies that a simple regression can yield a high-frequency estimate of the conditional skewness of permanent income shocks, i.e. $b' y_t$. The dependent variable is the time series of 1 and 5-year third central moments from GOS. I allow $\rho(L)$ a restricted MA(1) structure: $\rho(L) = \rho \cdot [1 + L]$. This specification implies that permanent shocks can have additional temporary effects that last about 6 months, though I find similar results with different lag lengths.

Next, I must specify the state variables y_t to include, as well as an estimation method. I explore two approaches. One option is to estimate b directly via GMM (OLS when ρ is known, non-linear least squares otherwise). The length of the sample precludes the estimation of a large number of coefficients, so the risk of overfitting is nontrivial. Thus, such an approach works well only if the dimension of y_t is relatively low. Appendix A.3.1 adopts this approach, estimating univariate and bivariate specifications involving several theoretically-motivated regressors.

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

I include 109 quarterly time series in the vector y_t, all of which are available from 1960-2013. I begin with 97 monthly macroeconomic and financial time series considered in Bernanke et al.

[28]Otherwise, one could still motivate our specification using standard linear projection arguments.

Category	Number of Variables	R^2: 3^{rd} moment 1 year	5 year	Correlation w/ overall index
All variables (3PRF)	109	70.1	80.8	100.0
All variables (Combination)	109	66.4	72.2	99.6
Real output and income	18	66.6	68.8	96.3
Employment and hours	24	57.7	63.3	95.8
Real inventories, orders, and unfilled orders	5	56.8	65.1	89.1
Stock returns and predictability state variables	14	51.4	74.4	68.9
Money and credit quantity aggregates	6	53.9	66.0	63.9
Interest rates	11	30.7	47.0	62.1
Housing starts and sales	7	11.2	13.4	49.7
Consumption	4	28.9	5.1	48.6
Consumer expectations	1	7.0	17.6	34.0
Price indices	15	7.6	16.9	25.6
Average hourly earnings	2	3.2	7.2	19.8
Exchange rates	2	12.5	15.1	12.5

Table 2: Goodness-of-fit statistics and variance decomposition of skewness index

This table presents the results from estimating equation (4) for different subsets of y_t. The dependent variables are 1 and 5-year 3^{rd} central moments of the cross-sectional distribution of income growth rates for 1978-2011 from GOS. Columns report R^2 values at each horizon, as well as the correlation between the implied skewness measure and the overall measure in the top row. Parameters are estimated so as to minimize the sum of squared residuals for both 1 and 5-year measures. The specification in the first line uses the 3PRF, while the remaining specifications construct a linear index as a weighted average of the univariate forecasts from each variable in the category, using the inverse of the mean-squared error as weights. Variables and categories are listed in Appendix A.3.2. The parameter (ρ) governing the state dependent, temporary shock is 0.45.

(2005).[29] I augment these series with 12 additional variables from the literature on stock return predictability, which are updated regularly by Goyal and Welch (2008). I then construct quarterly series by averaging the underlying monthly series within each calendar quarter. To allow for potential lead-lag relationships, I additionally include several lags and (weighted and unweighted) moving averages of each variable. Further details are in Appendix A.3.2.

3.1.3 Skewness index estimates

Table 2 documents the forecasting performance of the overall index, which is constructed using the 3PRF, as well as other indices which are constructed using different subsets of the 109 variables in y_t. These subindices are constructed using an alternative method for estimating linear models in a data-rich environment: forecast combination methods–see, e.g. Timmermann (2006). The 3PRF uses a series of regressions to optimally combine the information from each univariate model, whereas the forecast combination approach takes a weighted average of the univariate models, using the inverse of the mean-squared error (IMSE) as weights.[30]

Looking at the first row of Table 2, the overall performance of my skewness index is quite strong. A single factor extracted from the 109 variables achieves R^2's of 70% and 81% at the 1-year and 5-year horizons, respectively. My estimate of ρ, which governs transitory risk, is 0.45, suggesting that a 10% decline in permanent income in quarter t is associated with an additional 4.5% transitory decline in income in quarters t and $t+1$. For additional discussion of the role of state-dependent transitory risk, which much easier to see within the context of the univariate and bivariate specifications explored there, I refer the interested reader to Appendix A.3.1.

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

Figure 2 plots the estimated quarterly skewness indices obtained using the 3PRF and combination approaches, respectively. Both measures are visually indistinguishable from one another and highly cyclical, peaking in expansions and bottoming out during (or immediately after) recessions. Note however that skewness dynamics appear to considerably richer than the two-state (expansion and recession) Markov process typically assumed in the empirical literature on estimating earnings process.[31] Moreover, the measures are quite persistent, exhibiting substan-

[29] I obtain the data from Global Insight and transform them as in Bernanke et al. (2005) to ensure stationarity. Of the 120 variables included in Bernanke et al. (2005), Wu and Xia (2014) identify a subset of 97 variables which are available through the end of 2013. Wu and Xia (2014) also verify that the findings of Bernanke et al. (2005) are replicable using the subset with available data.

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

[31] In unreported results, I find that a quarterly NBER recession indicator has little explanatory power for the

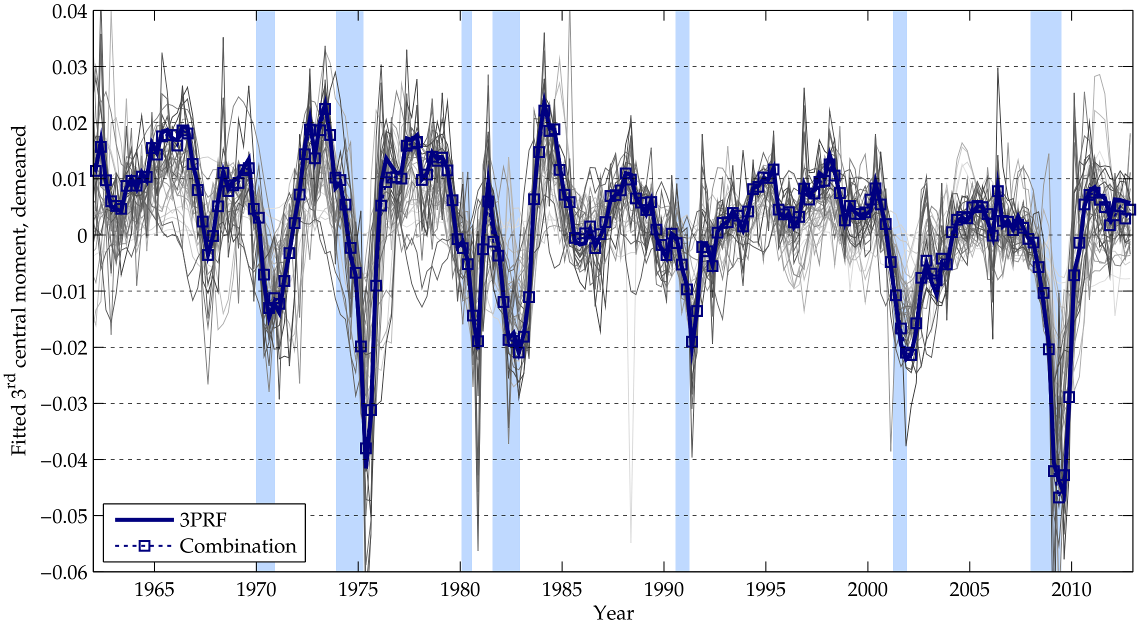

Figure 2: Conditional skewness estimates from combination and univariate models

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

tial variation at business cycle frequencies. The first-order autocorrelation, when expressed as a monthly number, is around 96% for each measure.

In addition to the overall skewness indices, Figure 2 plots fitted values from the 40 best-performing univariate regression models, which are shaded from light to dark according to goodness-of-fit. Darker lines indicate better fit. These univariate forecasts are highly correlated with one another and generally track the overall indices quite closely, indicating a strong factor structure in the data. Note that the estimation sample begins in 1978, while the overall and univariate skewness measures track one another quite closely prior to 1978, suggesting that the identified factor is a genuine feature of the data, rather than an artifact of over-fitting.

time aggregated skewness measures. However, such a result could partially reflect a timing mismatch between NBER recession dates and labor market peaks and troughs, which can lag the business cycle.

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

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

Perhaps more surprisingly, an index constructed using aggregate consumption measures does not capture the variation in idiosyncratic skewness relative to other measures of real activity. The consumption-based subindex achieves a modest R^2's of 29% and 5% at the 1-year and 5-year horizons, respectively, though the consumption-based index maintains a 50% correlation with the overall skewness index. Much of the disconnect in performance relative to other economic activity measures is certainly due to the substantial measurement errors in high frequency consumption data.[32] In addition, such a result could reflect temporal instabilities in lead-lag relationships between aggregate consumption and other measures of real activity.

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

[32] Recall that, so as to use the same data as in Bernanke et al. (2005), I construct quarterly series by averaging the monthly measures over each quarter. In Appendix A.3.1, I estimate univariate regressions using quarterly NIPA data on real consumption of nondurables and services. Using these data, aggregate consumption works somewhat better. However, its performance is inferior to that of other univariate indicators, such as real compensation growth, employment growth, or a measure of overall profitability of corporations.

3.2 Parametric model with labor market event risk

The analysis in the previous section was semi-parametric, allowing me to make statements about the time-series behavior of idiosyncratic skewness without any distributional assumptions. In order to close a quantitative asset pricing model, such assumptions are required. This section adopts a more parametric approach. I 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.

I maintain my assumptions on the labor income process from the previous section. For ease of notation, I will suppress i superscripts here. Equation (41), in Appendix A.2.1, shows that the growth rate of annual labor income at horizon k is the sum of three pieces: profile heterogeneity, $(\beta \cdot k)$, a state-independent transitory shock $(\epsilon_{A,t} - \epsilon_{A,t-k})$, and a weighted moving average of permanent shocks $(\theta_k(L;\rho)\eta_t)$. For the first two terms, I adopt the fairly standard assumptions that $\beta \sim N(0, \sigma_b^2)$ and $\epsilon_{A,t} \sim N(0, \sigma_\epsilon^2)$.

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

$$J_{g,t} = \sum_{j=1}^{N_{g,t}} [\mu_s + Exponential(\sigma_s) - \sigma_s], \qquad N_{g,t} \sim Possion(\lambda_{0g} + \lambda_1 x_t) \tag{5}$$

$$J_{b,t} = \sum_{j=1}^{N_{b,t}} [-\mu_s + \sigma_s - Exponential(\sigma_s)], \qquad N_{b,t} \sim Possion(\lambda_{0b} - \lambda_1 x_t), \tag{6}$$

where x_t is a scalar, and $J_{g,t} = 0$ and $J_{b,t} = 0$ when $N_{g,t} = 0$ and $N_{b,t} = 0$, respectively. The first component, $J_{g,t}$, is a good shock, capturing 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 infrequent, large downward adjustments–"falling off the ladder"–likely driven by events such as job loss. I 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 $\lambda_{0g} + \lambda_1 x_t$ and $\lambda_{0b} - \lambda_1 x_t$ as the quarterly probabilities of experiencing good and bad shocks, respectively. State dependence manifests itself via time variation in these probabilities. The conditional skewness of η_t is proportional to $\lambda_1 x_t$. When

$\lambda_1 x_t$ is high, large positive shocks become more likely while large negative shocks become less likely, shifting probability mass from the left to the right tail.

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

As discussed earlier, if the data are generated according to (5-6), the quarterly skewness index from Figure 2 is a consistent estimate of x_t up to a constant of proportionality. Therefore, when calibrating the idiosyncratic shock distribution, I set x_t equal to my skewness index, normalized it to have mean zero and variance one. After this normalization, λ_{0g} and λ_{0b} capture the unconditional probabilities of experiencing good and bad shocks, while λ_1 captures the marginal effect of a 1 standard deviation increase in x_t on the probability of a good shock.

My calibration tries to match a number of statistics Table 1, plus several additional moment conditions from statistics reported by GOS. I estimate the parameters so as to minimize a weighted sum of squared errors between a number of model-implied and data-implied moments.[33] First, I try to match the average distance between the median and the 10^{th} percentile of 5-year income growth rates in expansions and recessions. I also target the average distance between the 90^{th} percentile and the median in expansions and recessions. Second, I target the changes from recessions versus expansions in the left tail and right tail width measures. Third, I target the average distance between the 90^{th} and 10^{th} percentiles of 1-year growth rates in expansions. Finally, I add information about the standard deviation, skewness, and kurtosis from GOS.

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

[33] I also place some practical constraints on the parameters, which restrict the variance of jump shocks, σ_s, and guarantee that the fitted poisson intensities are non-negative for most of the sample. If the fitted intensity is negative, I truncate $\lambda_1 x_t$ so that the minimum intensity is zero the sum of the fitted intensities is $\lambda_{0b} + \lambda_{0g}$.

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

$$f_{k,t}(z;\beta) = \frac{1}{\pi} \cdot \text{Real}\left[\int_0^\infty [\varphi_{k,t}(\omega;\beta)e^{-i\omega z}]d\omega\right], \tag{7}$$

which involves a single numerical integration. I use the fractional fast Fourier transform to efficiently evaluate $f_{k,t}(z;\beta)$ on a fine grid over the support of $w_{A,t} - w_{A,t-k}$. By integrating $f_{k,t}(z;\beta)$, I quickly and accurately recover the conditional cdf and quantile functions.[34] Expressions for $\varphi_{k,t}(z;\beta)$ and further details about the procedure are in Appendix A.3.3.

Table 3 presents estimates of the parameters governing the labor income process. $\lambda_{0g} + \lambda_{0b}$ is about 2%, suggesting that the probability of receiving a large shock within a given year is about 8%. λ_{0b} is larger than λ_{0g}, implying that large negative shocks are more likely to occur than large positive shocks. λ_1 is 0.26%, implying that, on an annualized basis, a 1 standard deviation increase in x_t shifts 1% of the probability mass from bad to good shocks.

The magnitudes associated with the state-dependent shocks are extremely large. Conditional on receiving a large shock, the average absolute change in log wages is 77.5% and the standard deviation is 51.7%. Incorporating Jensen's inequality, the average decline in wages from a large negative shock is about 49%! Positive shocks induce an average increase of 165%. Before moving forward, recall that these magnitudes are for declines in pre-tax labor earnings for a single individual. Associated declines in household consumption are likely to be smaller. In my quantitative model, I assume that less than 25% of income declines pass through to consumption, which is on the low end of estimates in Blundell et al. (2008) and Heathcote et al. (2014).

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

Figure 3, Panel A plots the evolution of the fitted probabilities of good and back shocks, respec-

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

Parameter	Value	Description
λ_{0g}	0.75%	Average quarterly intensity of large positive shocks
λ_{0b}	1.25%	Average quarterly intensity of large negative shocks
λ_1	0.26%	Sensitivity of quarterly intensity of large shocks to a one standard deviation shock to business cycle factor
μ_s	77.5%	Absolute value of average change in log wages given a large shock
σ_s	51.7%	Standard deviation of a large shock to wages
σ_n	1.86%	Standard deviation of quarterly state-independent permanent shock
σ_ϵ	13.5%	Standard deviation of annual state-independent temporary shock
ρ	45%	State-dependent mean of transitory shock (MA parameter)

Table 3: Calibrated income process parameters

This table presents the estimated parameters governing the distribution of idiosyncratic labor income shocks. Estimates are obtained by minimizing a weighted sum of squared errors between model-implied and data moments, both of which are displayed in Table 4. These moments are averages of quantiles or moments of the cross-sectional distribution of income from GOS.

tively, 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.

Note that the fitted good intensity goes slightly negative during the financial crisis. For purposes of calculating the model-implied moments, I truncate the bad intensity so that the sum of the two intensities is always $\lambda_{0g} + \lambda_{0b}$. The dashed line shows the untruncated path of λ_{bt}. These untruncated estimates suggest that, during the financial crisis, idiosyncratic risk reached unprecedentedly high levels relative to the rest of the period where the index is available. Therefore, from an incomplete markets perspective, the Great Recession could easily be considered a "disaster" in spite of the relatively moderate observed decline in aggregate consumption.[35]

Figure 3, 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, I subtract the median from each quantile.

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

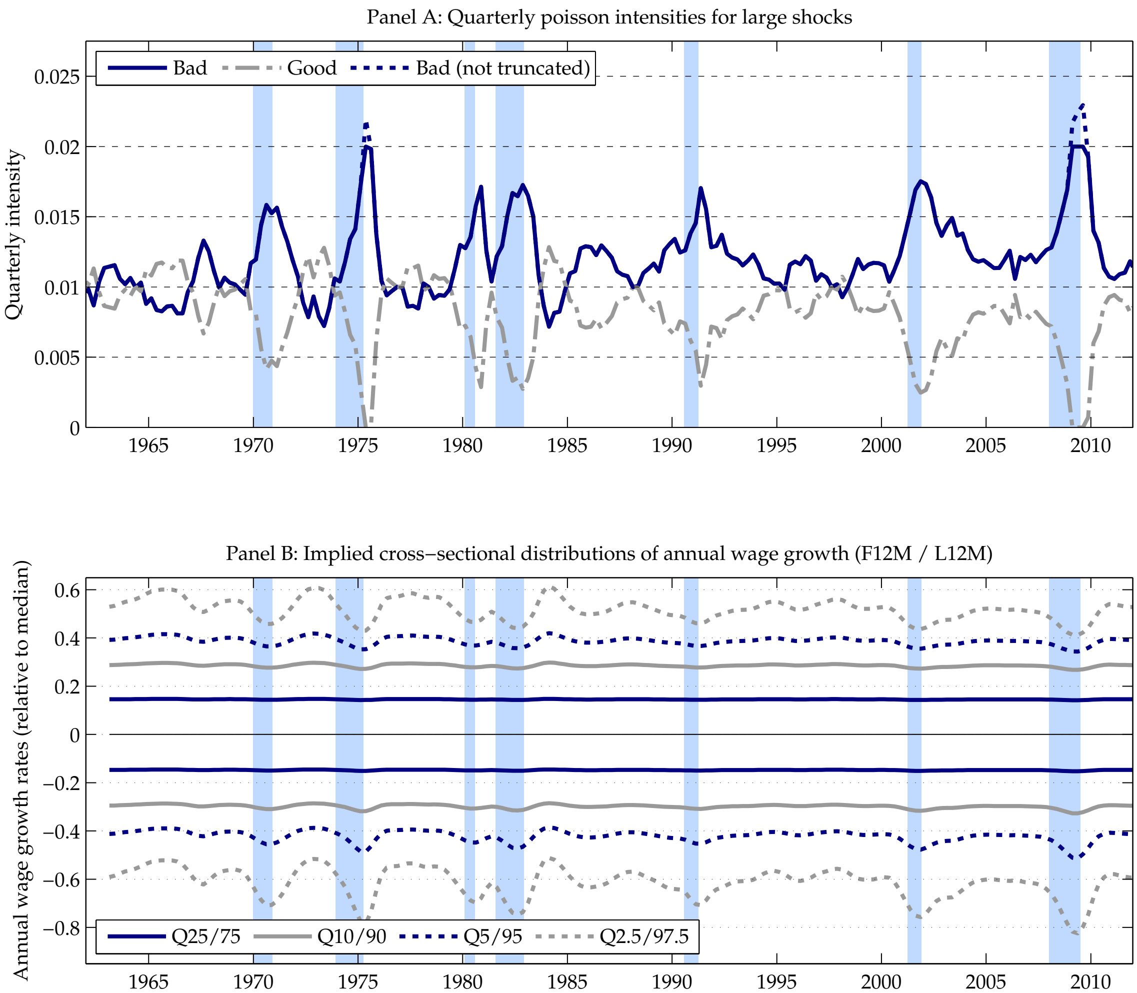

Figure 3: Fitted dynamics of idiosyncratic distributions

Panel A plots the poisson intensities for good and bad shocks from the estimated model for the income process. Panel B plots the difference between the median and several quantiles of the model-implied distributions of year-on-year changes in income.

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 4 plots several densities associated with the fitted model. Panel B characterizes the densities of the permanent component of year-on-year changes in wages in expansions and recessions,

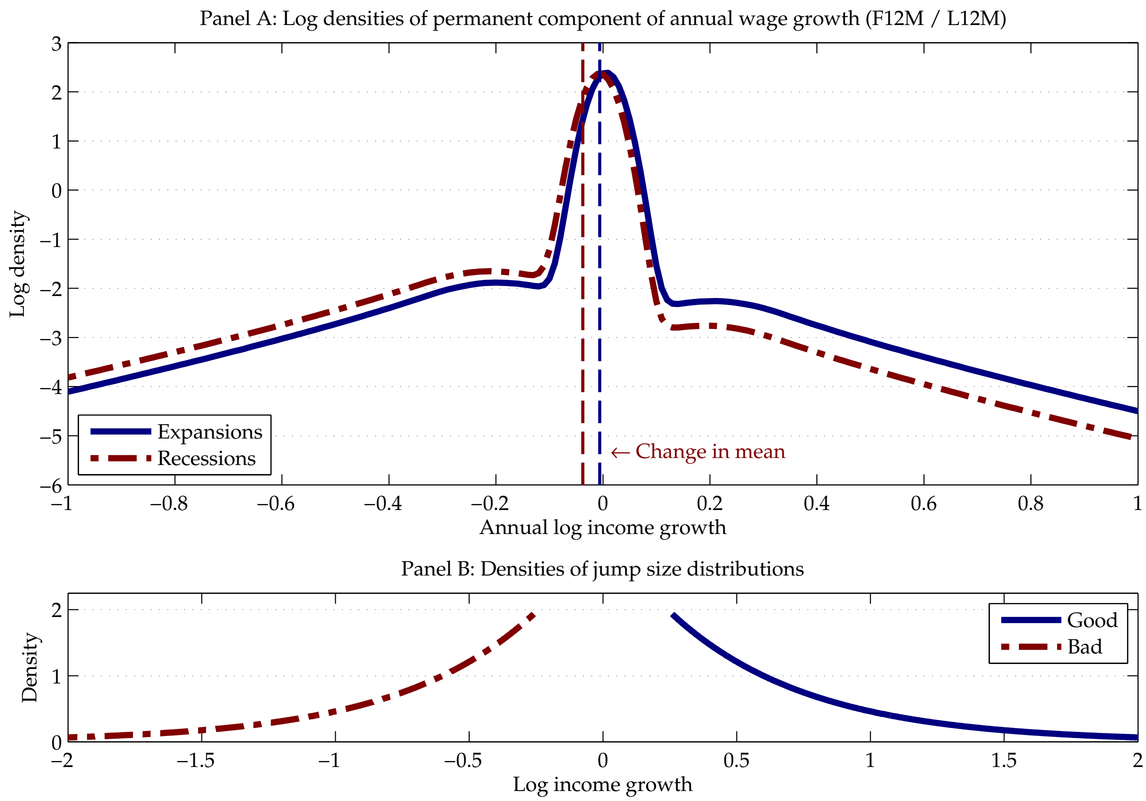

Figure 4: Densities from fitted model

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.

respectively. To generate the figures, I randomly sample with replacement from the observed values of x_t in expansion quarters and recession quarters, respectively, then plot the densities of year-on-year changes in the permanent component of wages (i.e. I strip out transitory shocks). I also allow for a time variation in the average logarithmic growth rate of labor income, which is assumed to be an affine function of the change in aggregate private sector real compensation, choosing the slope and intercept to exactly match the median in expansions and recessions from Table 1. I use a log scale on the vertical axis in order to better show the changes in the tails.[36]

Dashed vertical lines indicate the average change in log wages in expansions and recessions,

[36]On this scale, the pdf of a normal distribution is a quadratic.

Cross-sectional distribution of income growth rates				Time series		
Statistic	Period	Data	Model	Statistic	Horizon	R^2
50-10 Percentile Spread	5 yr E	64.14	63.29	Mean - 50 Percentile	1 yr	0.790
	5 yr R	71.82	72.79		3 yr	0.801
	5 yr R-E	7.68	9.50		5 yr	0.793
90-50 Percentile Spread	5 yr E	51.66	51.64	50-10 Percentile Spread	1 yr	0.520
	5 yr R	43.37	44.27		3 yr	0.702
	5 yr R-E	-8.30	-7.37		5 yr	0.671
50-25 Percentile Spread	5 yr R-E	3.01	1.81	90-50 Percentile Spread	1 yr	0.659
75-50 Percentile Spread	5 yr R-E	-1.72	-0.79		3 yr	0.695
95-50 Percentile Spread	5 yr R-E	-15.47	-15.15		5 yr	0.667
99-50 Percentile Spread	5 yr R-E	-26.71	-18.62	Kelley skewness	1 yr	0.709
90-10 Percentile Spread	1 yr E	58.67	58.31		3 yr	0.762
Standard Deviation	5 yr E	66.52	62.36		5 yr	0.789
Skewness (Moment)	5 yr E	-1.18	-0.29			
	5 yr R	-1.80	-0.76			
	5 yr R-E	-0.62	-0.47			
Kurtosis	5 yr	≈ 10	10.53			

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

This table presents goodness-of-fit statistics for the calibrated model for idiosyncratic labor income risk. The left panel compares model-implied moments with their counterparts in the data. These moments are averages of quantiles or moments of the cross-sectional distribution of income growth rates for individuals in the 91^{st} through 95^{th} 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.

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.

Figure 4, Panel B 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 26% in the calibrated model, about seven times the annualized standard deviation of the permanent gaussian shock.

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 replicates the negative unconditional asymmetry which is observed in the data. The model comes pretty close to matching the cyclical variation in the 50-10 and 90-10 spreads, slightly overestimating variation in the former and underestimating the latter. It matches the level of the 90-10 spread at 1 and 5 year horizons almost perfectly. While the exact magnitudes of changes in the other quantiles do not match perfectly, particularly for the 99-50 percentile spread, the fit is reasonably close. The presence of rare, large shocks generates substantial cyclical variation in the tails of the distribution, while leaving the central quantiles essentially unchanged.

In addition to quantile-based measures from Table 1, I target several time series averages of cross-sectional moments from GOS. First, the model-implied standard deviation is relatively close, though slightly lower, to its value in the data.[37] GOS provide a number of reasons to prefer quantile-based skewness measures such as those reported in Table 1. I also compare the calibrated (moment-based) skewness measures to their counterparts in the data. The fitted model has some trouble matching the level of unconditional skewness, though it does a better job of matching the observed change in skewness from expansions.[38]

I also calculate the kurtosis of 5-year earnings growth. Guvenen et al. (2014b) emphasize the extremely high degree of excess kurtosis observed in the SSA earnings data, which is at odds with the assumption of normally distributed shocks that is ubiquitous in the literature on calibrating earnings distributions. Their estimates, which are constructed using similar methods to GOS, suggest that the kurtosis of 5-year income growth rates for relatively high earners is about 10. My model with jump shocks generates an average kurtosis of 10.5, consistent with this evidence.

The right panel of Table 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, my calibrated model targets the level of income risk faced by relatively high earners. However, I can still assess whether our fitted model qualitatively matches the time series dynamics of these cross-sectional moments. At 1, 3, and 5-year horizons, I report the R^2 from a univariate regression of the GOS data on model-implied statistics. I consider four different time series, two of which are plotted in Figure 1: the difference between the mean and the median, the 50-10 spread, the 90-10 spread, and Kelley's skewness.

[37] It is worth noting that the fitted standard deviation is a bit further off at a 1-year horizon, suggesting that there is room for improvement in my specification of transitory risk.

[38] Recall from the discussion above that the constant term, a_k, in (4) capturing unconditional skewness is more difficult to identify relative to the slope coefficient b. The same difficulty applies here as well.

Overall, my model matches essentially all of these time series moments quite well. The R^2's tend to be highest at a 3-year horizon, which is somewhat reassuring given that no data from that horizon were used to estimate the linear index, x_t. I have some trouble matching the transitory variation in the 50-10 spread in recessions, suggesting that I am omitting some important sources of transitory risk. The high R^2's on the Kelley's skewness measure suggest that we do a reasonable job of capturing the cyclical variation in the overall asymmetry of the distribution.

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

4 Theoretical Framework

In this section, I embed an endowment-based asset pricing model with heterogeneous agents and incomplete markets within a general affine, jump-diffusion framework. My setup most closely resembles the model in Toda (2014b), itself based upon the seminal contribution of Constantinides and Duffie (1996). I 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.

The model is a Lucas (1978) endowment economy with incomplete markets. Agents' consumption stream derives from two types of assets (trees), each of which delivers an uncertain stream of future cash flows (fruit). Between periods, total fruit output grows stochastically and the growth of each tree is potentially subject to aggregate and idiosyncratic shocks. The first type of tree, human capital (H_t^i), 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 is that labor income is subject to idiosyncratic risk, meaning that different investors will receive different returns over the same holding period because their trees will not all grow at the same rate. Defining the aggregate quantity $C_t \equiv \int_I C_t^i di$, the fruit production of the first type of tree grows at rate $C_t/C_{t-1} \times \exp(\eta_t^i)$, where η_t^i is a shock which is independently and identically distributed across agents satisfying $E[\exp(\eta_t^i)] = 1$.[39] Households are unable to buy or sell human capital, nor trade contingent claims on realizations of η_t^i.[40] 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):

$$U_t^i = \left[(1-\delta)(C_t^i)^{1-1/\psi} + \delta(E_t[(U_{t+1}^i)^{1-\gamma}])^{\frac{1-1/\psi}{1-\gamma}}\right]^{1/(1-1/\psi)}, \tag{8}$$

where ψ governs the elasticity of intertemporal substitution (EIS) and γ is the coefficient of relative risk aversion.[41] At time 0, each agent begins with an initial endowment H_0^i. Thereafter, each agent chooses her consumption (the numeraire) and investment N_t^i to maximize (8). All financial assets are in zero net supply, so market clearing will imply $N_{kt}^i = 0$ for all i, k, and t.[42]

These assumptions imply the following budget constraint

$$C_t^i + \sum_{k=1}^{K} P_{kt} N_{kt}^i \quad = \quad C_t \exp(\eta_t^i) H_{t-1}^i + \sum_{k=1}^{K} (P_{kt} + D_{kt}) N_{k,t-1}^i \tag{9}$$

subject to $\sum_{k=1}^{K} P_{kt} N_{kt}^i > -\underline{W}$, where P_{kt} is the price of the k^{th} financial asset. The borrowing constraint, which will not bind in equilibrium, is simply present in order to rule out Ponzi schemes. In the section that follows, I will restrict attention to symmetric, no-trade equilibria. Under my assumptions, $H_t^i = \exp(\eta_t^i) H_{t-1}^i$. This will imply that

$$C_t^i = H_t^i C_t \quad \Rightarrow \quad \Delta c_t^i = \Delta c_t + \eta_t^i, \tag{10}$$

[39] I will formalize my assumptions about η_t^i later in the section.

[40] Since the financial assets are in zero net supply, no-trade will be an equilibrium. Therefore, I could assume that households are able to buy and sell human capital without affecting any of the results below. The key friction is the inability to write contingent claims on η_t^i.

[41] 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 δ.

[42] Without loss of generality, we normalize the total supply of human capital to equal 1.

where $c_t = \log(C_t)$. I denote levels with capital letters and logs with lower case letters, e.g. $d_{k,t} = \log(D_{k,t})$. This is a special case of the 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.[43] 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.[44]

Assumption 1 gives my general model for aggregate dynamics.

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

$$\underset{(L\times 1)}{y_{t+1}} = \mu_y + F_y y_t + G_{y,t} z_{y,t+1} + J_{y,t+1}, \tag{11}$$

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 $S_1, \ldots, S_K$

I summarize the jump size distribution for the compound Poisson shocks with $\Psi_y(u)$, the $L \times 1$ vector-valued function whose j^{th} element is the moment-generating function of the size distribution for the j^{th} jump component. I need little structure on $\Psi_y(u)$ beyond the existence of such a function (and boundedness for certain values of u).

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.[45] 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 $G_{y,t}$.

[43] 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.

[44] See, e.g. Bansal et al. (2010) and Dew-Becker and Giglio (2013).

[45] The disaster model obtains if I assume that both S_c and S_k have a common exposure to one of the Poisson jump components, whose probability varies over time.

My next assumption places some structure on the idiosyncratic shocks, η^i_{t+1}. Assumption 2 allows the distribution of the idiosyncratic shock to depend on the realization of the aggregate state vector, y_{t+1}. This structure means that, ex post, aggregate shocks (e.g. consumption declines) need not be distributed equally across agents. Denote agent i's private information by the filtration $\mathscr{F}^i_t$ and public information by $\mathscr{F}_t = \bigcap_i \mathscr{F}^i_t$.

Assumption 2. *The following statements are true.*

(i) $\eta^i_t = 1'_M \tilde{\eta}^i_t$*, and, conditional on* y_{t+1}*,* $\tilde{\eta}^i_t$ *is generated according to*

$$\underset{(M\times 1)}{\tilde{\eta}^i_{t+1}} = \mu_\eta + F_\eta y_{t+1} + G_{\eta,t+1} z^i_{\eta,t+1} + J^i_{\eta,t+1}, \tag{12}$$

where $z^i_{\eta,t+1}$ *is a vector of standard normal random variables that is i.i.d. across agents and over time,* $G_{\eta,t}G'_{\eta,t}$ *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* $\lambda_{\eta,t+1}$*.*

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

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

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

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 I 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 can capture unemployment-type events, except that I must collapse all of the effects of unemployment into a single shock rather than allow for a series of bad shocks which unfold gradually over time.[46] It also abstracts away from

[46] A similar simplifying assumption is common in the literature on rare macroeconomic disasters.

heterogeneity in risks between individuals (e.g. skill heterogeneity, life cycle effects, etc.) and across industries/occupations, which would require a substantially more complicated model.

The remaining assumptions involve the agents' information and moment restrictions on η_{t+1}^i, 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* $h_{\eta 0} + H_{\eta 1}' y_t$*, where* h_η *and* $H_{\eta 1}$ *are a scalar and a* $M \times 1$ *vector, respectively. All other elements of* $G_{\eta,t}G_{\eta,t}'$ *are zero.*

(iv) $\lambda_{\eta,t} = l_{\eta 0} + l_{\eta 1} y_t$*, where* $l_{\eta 0}$ *and* $l_{\eta 1}$ *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.[47] In the absence of idiosyncratic shocks, my framework nests long-run risk-type representative agent models with Poisson jumps, such as DY.

Assumptions 3.iii-vi parameterize the idiosyncratic shocks. Assumption 3.iii allows for a normally-distributed "diffusion" shock, and the variance of this shock is allowed to be state-dependent.[48] As such, I can easily allow for CCSV within our model. Given that GOS find little evidence of CCSV in their Social Security Administration dataset, my analysis will focus more on state-dependence in the "jump" shocks. However, the theoretical implications of and intuition for

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

[48]Without loss of generality, I can concentrate all of the "diffusion risk" in the first element of $\tilde{\eta}_t^i$, so M, the dimension of η_t^i, is solely determined by the number of independent sources of jump risk.

time-varying volatility of the Gaussian shocks are similar. Assumption 3.iv parallels Assumption 3.ii, allowing for state-dependence in the idiosyncratic jump intensities.

It is sensible to ask whether Assumptions 2 and 3 are compatible with one another. Proposition 1, proved in Appendix B.2.1, shows that this is the case, deriving admissible expressions for μ_η and F_η which will guarantee that Assumption 2.iii holds.

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

(i) $\mu_\eta = -1/2 h_{\eta 0} e_1 - l_{\eta 0}(\Psi_\eta(1_M) - 1_M)$

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

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

Proposition 1 says that under the distributional assumptions above, there exist choices of μ_η and F_η which aggregate properly. Therefore, I can fully decouple any assumptions about the time-variation in the distribution of idiosyncratic consumption risk from aggregate consumption while preserving linearity of the process. For purposes of asset pricing, I can assume without loss of generality that μ_η and F_η 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 my 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 $\{P_{1,t}, \ldots, P_{K,t}\}_{t=0}^{\infty}$ and allocations $\{C^i_t, i \in I\}_{t=0}^{\infty}$ which solves agents' optimization problems and satisfies market clearing in the capital markets. I restrict attention to symmetric (no-trade) equilibria where all agents consume their endowments. Such an equilibrium can obtain since all agents have identical homothetic preferences and access to the same investments, making their first order conditions 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, I characterize its properties.

Mathematically, asset pricing behavior in my 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 η_t^i. My 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:

$$1 = E_t\left[\delta^\theta \left(\frac{C_{t+1}^i}{C_t^i}\right)^{-\frac{\theta}{\psi}} (R_{c,t+1}^i)^{-(1-\theta)} \tilde{R}_{t+1}\right] = E_t\left[\delta^\theta \left(\frac{C_{t+1}^i}{C_t^i}\right)^{-\gamma} \left(\frac{WC_{t+1}+1}{WC_t}\right)^{-(1-\theta)} \tilde{R}_{t+1}\right], \quad (13)$$

where $\theta = \frac{1-\gamma}{1-1/\psi}$, $R_{c,t+1}^i \equiv \frac{W_{t+1}^i + C_{t+1}^i}{W_t^i}$ 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.[49] 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 (10) into (13) yields

$$1 = E_t\Bigg[\underbrace{\delta^\theta \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} \left(\frac{WC_{t+1}+1}{WC_t}\right)^{-(1-\theta)}}_{\text{“representative agent” pricing kernel}} \quad \underbrace{\exp(-\gamma \cdot \eta_{t+1}^i))}_{\substack{\text{additional term from} \\ \text{uninsurable idiosyncratic risk}}} \quad \tilde{R}_{t+1}\Bigg] \equiv E_t[M_{t+1}^i \tilde{R}_{t+1}], \quad (14)$$

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.[50] The second term incorporates idiosyncratic consumption risk, which, since it is undiversifiable and uninsurable to the agent, also affects risk premia.

[49] In a symmetric equilibrium, WC_t is identical across agents so I suppress i superscripts.

[50] 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, η^i_{t+1} is independent of $\tilde{R}_{t+1}$ given the past and current realizations of the aggregate state, y_{t+1}. I will refer to assets satisfying this independence restriction, such as the dividend claims, as "financial assets". When pricing financial assets, I use the law of iterated expectations to re-write (14) as

$$1 = E_t\left[\delta^\theta \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} \left(\frac{WC_{t+1}+1}{WC_t}\right)^{-(1-\theta)} E_t[\exp(-\gamma \cdot \eta^i_{t+1}))|y_{t+1}]\tilde{R}_{t+1}\right] \equiv E_t[M_{t+1}\tilde{R}_{t+1}] \quad (15)$$

$$= E_t\left[\delta^\theta \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} \left(\frac{WC_{t+1}+1}{WC_t}\right)^{-(1-\theta)} \left(\sum_{j=0}^{\infty} \frac{(-\gamma)^j}{j!} E[\eta^j_{t+1}|y_{t+1}, y_t]\right) \tilde{R}_{t+1},\right] \quad (16)$$

where the second line plugs in a Taylor series expansion of $\exp(-\gamma \cdot \eta^i_{t+1})$ around zero before taking the (cross-sectional) expectation.[51] This expression shows that, in general, the pricing kernel is higher in states where even (odd) moments of the cross-sectional distribution of η^i_{t+1} 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 comes from investors' preferences over the temporal resolution of uncertainty. When the EIS (ψ) 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 η^i_{t+1} is independent of all of the other random variables in (14) 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, I 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.[52] Therefore, the idiosyncratic risk term will have the same impact on expected returns and the risk-free rate. However, idiosyncratic risk

[51] The projected kernel will not price assets whose payoffs depend on η^i_{t+1} properly.

[52] This follows from the identity $\log E_t(\tilde{R}_{t+1}) - rf_{t+1} = \log[1 - \mathrm{cov}(M^i_{t+1}, \tilde{R}_{t+1})]$.

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.[53]

In equilibrium, (13) and (14) must hold for all assets in the investment opportunity set. Plugging the consumption claim and financial asset returns into (14) yields

$$1 = E_t\left[\delta^\theta \left(\frac{C_{t+1}}{C_t}\right)^{1-\gamma} \left(\frac{WC_{t+1}+1}{WC_t}\right)^{\theta} \exp((1-\gamma)\eta^i_{t+1})\right] \tag{17}$$

$$1 = E_t\left[\delta^\theta \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} \left(\frac{WC_{t+1}+1}{WC_t}\right)^{-(1-\theta)} \exp(-\gamma\cdot\eta^i_{t+1}) \left(\frac{\frac{P_{k,t+1}}{D_{k,t+1}}+1}{\frac{P_{kt}}{D_{kt}}}\right) \frac{D_{k,t+1}}{D_{k,t}}\right], \tag{18}$$

a system of nonlinear equations involving the two key valuation ratios, WC_t and P_{kt}/D_{kt}. Since all of exogenous quantities in (17-18) are stationary by Assumptions 1-3, the model may be solved numerically by finding the value of WC_t that satisfies (17) for each value of y_{t+1} in the state space. Then, given the solution for WC_t, one can use (18) to solve for the equilibrium price-dividend ratios for the financial assets.

5.2 Solution

In this section, I briefly outline how to solve for asset prices with incomplete markets and time-varying idiosyncratic risk. Since my model for aggregate dynamics is quite general and has many properties that have been studied elsewhere, I will primarily focus on the incremental effects associated with incomplete markets. For brevity, many technical details may be found in Appendix B.2.[54] The only requisite approximation is the standard Campbell and Shiller (1988) log-linearization of returns, which features in the most common representative agent solution method. Adding idiosyncratic risk does not necessitate any additional approximations.

5.2.1 Log-linearization

Denote continuously compounded returns by lowercase letters (e.g. $r^i_{c,t} = \log R^i_{c,t}$). I linearize the return on the consumption claim around a constant log wealth-consumption ratio $\overline{wc}$, yielding

$$r^i_{c,t+1} \approx \kappa_c + \Delta c^i_{t+1} + \rho_c wc_{t+1} - wc_t, \tag{19}$$

[53]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.

[54]Additional details and discussion are available in Eraker and Shaliastovich (2008) and DY.

with linearization constants $\rho_c \equiv \frac{\exp(\bar{wc})}{\exp(\bar{wc})+1} < 1$ and $\kappa_c \equiv \log(\exp(\bar{wc}) + 1) - \frac{\exp(\bar{wc})}{\exp(\bar{wc})+1}\bar{wc}$.[55] Combining (13) and (19), the log of the one period pricing kernel approximately equals

$$m_{t+1}^i = \underbrace{\theta \log \delta - (1-\theta)\kappa_c - \gamma \Delta c_{t+1} - (1-\theta)(\rho_c wc_{t+1} - wc_t)}_{\text{linearized ``representative agent'' term}} - \underbrace{\gamma \cdot \eta_{t+1}^i}_{\substack{\text{incomplete} \\ \text{markets term}}} . \tag{20}$$

Equation (20) is the linearized version of (14), where the representative agent pricing kernel is augmented by an additional term capturing idiosyncratic risk.

Analogous to the result in (15), for purposes of pricing financial assets, I replace η_{t+1}^i with its projection onto the aggregate state, $\log E_t[\exp(-\gamma \cdot \eta_{t+1}^i)|y_{+1}]$, in the log-linearized pricing kernel. Note that $-\frac{1}{\gamma}\log E_t[\exp(-\gamma \cdot \eta_{t+1}^i)|y_{+1}] \equiv \nu_{t+1}$ can be interpreted as the log of an expected utility maximizer's certainty equivalent for lottery $\exp(\eta_{t+1}^i)$ given y_{t+1}. Thus, I can equivalently consider the following projection of the pricing kernel

$$m_{t+1} = \theta \log \delta - (1-\theta)\kappa_c - \gamma(\Delta c_{t+1} + \nu_{t+1}) - (1-\theta)(\rho_c wc_{t+1} - wc_t), \tag{21}$$

so an asset's covariance with the cross-sectional certainty equivalent ν_{t+1} affects its risk premium. Lemma 1 in the Appendix shows that, under my distributional assumptions, $\log E_t[\exp(u \cdot \eta_{t+1}^i)|y_{+1}] = \beta_0(u) + \beta(u)' y_{t+1}$, where $\beta(\cdot)$ is a known function whose expression is given in Lemma 1. Thus, the certainty equivalent ν_{t+1} is an affine function of y_{t+1}. Then, (21) becomes

$$m_{t+1} = \kappa - \gamma \Delta c_{t+1} - (1-\theta)(\rho_c wc_{t+1} - wc_t) + \beta(-\gamma)' y_{t+1}, \tag{22}$$

where $\kappa \equiv \theta \log \delta - (1-\theta)\kappa_c + \beta_0(-\gamma)$. Note that (22) will correctly price financial assets; the solution method for assets whose payoffs depend on η_{t+1}^i, namely the consumption claim, is somewhat different.

5.2.2 Valuation Ratios

Proposition 2 gives my key result, namely that valuation ratios are affine functions of the aggregate state vector y_t.

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

[55] 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 $\overline{wc}$.

(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 Appendix B.2.2, a brief outline of the solution method is as follows. I begin by conjecturing (and later verifying) that the log of the wealth-consumption ratio is an affine function of y_t.[56] I solve for A_0 and A using the Euler equation for the consumption claim and the method of undetermined coefficients. Given my restrictions on the law of motion for the state vector, I can evaluate the Euler equations analytically.[57] Since the Euler equations must hold for each y_t in the state space, I get a system of $L+1$ nonlinear equations which pin down the coefficients. Analytical solutions are available in special cases, but, in general, the system must be solved numerically.

A similar procedure yields solutions for the valuation ratios for the K other risky assets. I 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, I calculate the projected version of the pricing kernel in (22) and use the method of undetermined coefficients to solve for $A_{0,k}$ and A_k.

Plugging the affine form into (22), the projected pricing kernel, and subtracting off terms known as of time t yields

$$m_{t+1} - E_t(m_{t+1}) = -[\gamma S_c' - \beta(-\gamma)' + (1-\theta)\rho_c A'][y_{t+1} - E_t(y_t)] \equiv -\Lambda'[y_{t+1} - E_t(y_t)], \quad (23)$$

a multi-factor CAPM-like formula. Λ captures the sensitivity of investors' intertemporal marginal rate of substitution to shocks to the vector of aggregate state variables. The first two terms in Λ 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 ν_{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 Λ captures investors' hedging demands, incorporating the indirect effect.

In Appendix B.1, I provide affine expressions which can be used to price a single risky payment,

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

[57] An expression for the the conditional moment generating function of y_{t+1} given y_t is given in Lemma 2 in Appendix B.2.2. This expectation is an exponential affine function of y_t.

such as a single dividend $D_{k,t+h}$, as of time t. These expressions can be used to derive the term structure of (real or nominal) interest rates as well as the term structure of risk premia–see, e.g. Lettau and Wachter (2007) and Van Binsbergen et al. (2012). Understanding the pricing of risky cash flows at different points in time can help to clarify the mechanics of the model, and, in some cases, generate additional testable predictions. See Appendix B.1 for further details.

5.2.3 Risk Premia

Conditional on the vector with the prices of risk (Λ), 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, my 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 my solution for the price-dividend ratio, the log-linearized market return is

$$r_{k,t+1} = \kappa_k + (\rho_k - 1)A_{0,k} + (S_k' + \rho_k A_k')y_{t+1} - A_k' y_t \equiv \kappa_k + (\rho_k - 1)A_{0,k} + B_k' y_{t+1} - A_k' y_t. \quad (24)$$

Proposition 3 gives the solution for the equity premium, which is derived in Appendix B.2.4.

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

$$\log(E_t[R_{k,t+1}]) - rf_{t+1} = B_k' G_{y,t} G_{y,t}' \Lambda + \lambda_{y,t}' [\Psi_y(B_k) - 1_L] - \lambda_{y,t}' [\Psi_y(B_k - \Lambda) - \Psi_y(-\Lambda)]. \quad (25)$$

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.[58]

An immediate corollary to Proposition 3 is that a necessary condition for a variable to predict excess returns is that it is correlated with uncertainty about shocks to the aggregate state. In the next section, I present evidence that this condition holds in the data. When the distribution of

[58] For additional details and discussion, I also refer the reader to DY, sections 3.3.3 and A.4.

idiosyncratic shocks is particularly negatively skewed (i.e. large negative shocks are more likely), it is generally the case that uncertainty about future skewness is high.

5.3 ICAPM Representation

Campbell (1993) proposes an alternative method to derive the pricing kernel which substitutes out consumption growth, therefore relying only on returns data. The result is an intertemporal capital asset pricing model (ICAPM), which can be implemented empirically when the return of aggregate wealth is observable. Even if the return on wealth is unobservable, such a representation highlights the key sources of priced risk in a model. In addition to allowing for incomplete markets, I generalize Campbell et al. (2012) to allow for general affine jump-diffusion dynamics in the aggregate state vector, y_t.

Proposition 4, proved in Appendix B.2.3, provides an ICAPM representation of the pricing kernel. Define $R_{c,t+1} \equiv E\left[R^i_{c,t+1}|y_{t+1}\right]$ and $r_{c,t+1} = \log R_{c,t+1}$.

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

$$m_{t+1} - E_t m_{t+1} = -\gamma \underbrace{(\nu_{t+1} - E_t[\nu_{t+1}])}_{\substack{\textbf{contemporaneous}\\ \text{idiosyncratic}\\ \text{risk news}}} + (1-\gamma) \underbrace{N_{FIR,t+1}}_{\substack{\textbf{future}\\ \text{idiosyncratic}\\ \text{risk news}}} + \underbrace{N_{DR,t+1} - \gamma N_{CF,t+1} + \tfrac{1}{2} N_{UNC,t+1}}_{\text{representative agent terms}}, \tag{26}$$

where $\nu^*_{t+1} \equiv \frac{1}{1-\gamma} \log E_{t+1}[\exp(1-\gamma)\eta^i_{t+1}|y_{t+1}]$,

$$N_{FIR,t+1} \equiv [E_{t+1} - E_t] \sum_{j=1}^{\infty} \rho_c^j \nu^*_{t+1+j}, \qquad N_{DR,t+1} \equiv [E_{t+1} - E_t] \sum_{j=1}^{\infty} \rho_c^j r_{c,t+1+j},$$

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

and ϑ_t *is defined in Appendix B.2.3.*

Relative to the representative agent model, idiosyncratic risk adds two news terms to the pricing kernel, which are likely to be positively correlated in practice. As discussed above, the first term captures the "direct effect", news about *contemporaneous idiosyncratic risk*. Agents dislike assets that perform badly when the cross-sectional certainty equivalent, ν_{t+1}, is unexpectedly low. 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–the indirect effect, which is more transparent in the ICAPM representation. Again, ν^*_{t+1} is a certainty equivalent, but the associated power $(\gamma - 1)$ is lower, reflecting the fact that $r^i_{c,t+1}$ is also exposed to the idiosyncratic shock $\eta_{i,t+1}$. Given the high persistence of my skewness measure, this term is likely to be substantially larger in magnitude than the contemporaneous term. The additional hedging demands associated with this second term provide the primary amplification mechanism in my 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 γ times larger than discount rate news. 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 (26) 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 γ, 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.[59] Moreover, if the cross-sectional certainty equivalent ν^*_{t+1} is more persistent and/or volatile than aggregate consumption growth, this term can play a very important quantitative role in amplifying risk premia.

Finally, in the presence of stochastic volatility and/or jumps, there is a final representative agent term which captures news about state variables governing the higher moments of aggregate shocks. The Jensen's inequality term ϑ_t is high when uncertainty is high. All else constant, risk averse agents are willing to pay a premium for assets which hedge against increases in uncertainty. Thus, the price of risk on $N_{UNC,t+1}$ is negative. For additional discussion, I refer the interested reader to Campbell et al. (2012).

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, ϑ_t and ν^*_{t+1} are affine function

[59]Popular choices for γ in the theoretical literature with Epstein-Zin preferences often range between 5 and 15.

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 Λ, which can potentially be more robust to misspecification of the high frequency dynamics of Δc_{t+1}. Analytical expressions for these linear combinations are straightforward to derive and are available upon request.

6 Interactions between Idiosyncratic Risk and Stock Returns

In this section, I discuss the implications of our incomplete markets model for return predictability, and I consider the dynamic interactions between proxies for idiosyncratic risk and asset returns. In my model, a necessary condition for a variable to predict returns is for it to govern uncertainty about the aggregate state vector. I 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, I 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. I also explore the covariance structure between initial claims and these predictor variables.

6.1 Uncertainty about idiosyncratic risk is countercyclical

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

My preferred uncertainty measure is initial claims for unemployment insurance (UI). I divide the number of claims filed in each month by the size of the workforce (from the BEA) to obtain a stationary measure. An individual is only eligible for UI if he/she becomes "unemployed through no fault of his/her own" (e.g. laid off).[60] Thus, the normalized series may be interpreted as the rate of involuntary job loss in the cross-section of employed individuals in the private sector. Earlier, I demonstrated that my skewness index is most closely linked with measures of real activity and employment growth. Initial claims is a leading indicator of future labor market

[60] Source: http://www.edd.ca.gov/unemployment/Eligibility.htm.

conditions (Barnichon and Nekarda (2012)), which is available on a very timely basis and is subject to little measurement error.

While initial claims plausibly proxies for expected future labor market conditions, it is reasonable to ask why it should proxy for *uncertainty*. One would expect more layoffs when aggregate productivity is low. An elegant justification for a link between aggregate productivity and labor market uncertainty comes from Ilut et al. (2014). Using establishment-level Census data, they find strong evidence that firm-level hiring and firing decisions respond more strongly to bad news relative to good news about productivity. Ilut et al. (2014) show that, when this condition holds, the conditional volatility of aggregate employment growth is higher in bad times (i.e. when average firm productivity is low), even if all productivity shocks are iid. The same condition also implies that cross-sectional dispersion of employment growth is countercyclical.

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

I also consider two alternative uncertainty measures. The first is, x_t, the level of my skewness index. If, for example, x_t follows a square-root process, then its conditional mean and volatility are perfectly correlated with one another. The argument from Ilut et al. (2014) also applies to x_t. The second is a measure of cross-sectional volatility of employment growth 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.

To construct the cross-sectional measure, I use quarterly, seasonally-adjusted, state-level employment growth data from Hamilton and Owyang (2011).[61] For each state, I estimate

$$\Delta emp_{s,t} = \alpha_s + \beta_s \Delta \overline{emp}_t + \sigma_s u_{s,t}, \tag{27}$$

where $\overline{emp}_t$ is the cross-sectional average employment growth. My uncertainty measure is the cross-sectional volatility of the fitted residuals: $Vol_t \equiv \sqrt{\frac{1}{S}\sum_{s=1}^{I} \hat{u}_{s,t}^2}$.

Figure 5 plots the three (standardized) measures, initial claims for unemployment ($Claims_t$), the skewness index (x_t), and Vol_t over the time period for which both series are available. x_t

[61] I am grateful to James Hamilton for making the data available. I extend the data to the present by aggregating the monthly, seasonally-adjusted series which are now provided by the Bureau of Economic Analysis.

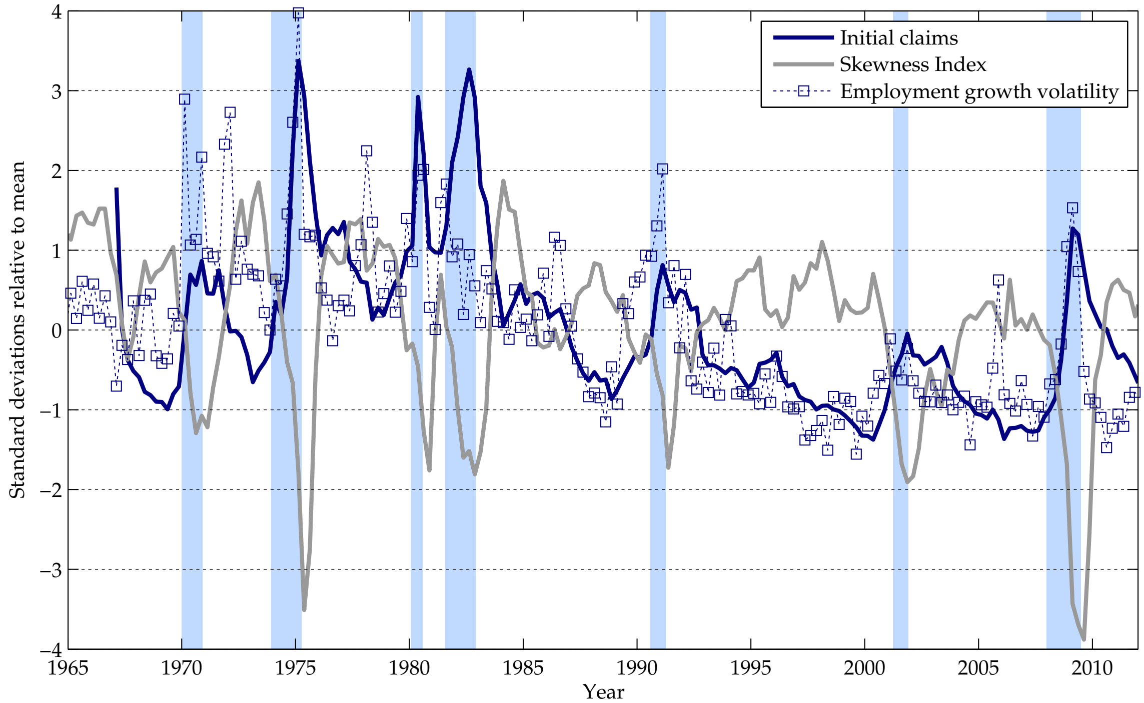

Figure 5: Co-movement of labor market uncertainty proxies

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

and $Claims_t$ are strongly negatively correlated and highly cyclical. Further, $Claims_t$ slightly leads x_t; bivariate Granger-causality tests provide strong evidence that $Claims_t$ Granger-causes x_t when 2 or more lags are included in the VAR. The business-cycle frequency movements of all three series are quite similar, though the cross-sectional measure differs somewhat during the double-dip recession of 1982. Initial claims is a very good proxy for cross-sectional uncertainty about labor market conditions.

Within my model, risk premium dynamics depend on uncertainty about *aggregate shocks*, particularly about shocks to persistent state variables. Recall that the cross-sectional measure, Vol_t, already strips out the effect of average employment growth. Next, I show that all three measures are good proxies for time series uncertainty about future idiosyncratic risk. In the labor market, cross-sectional and time series uncertainty are closely related.

To demonstrate this relationship more formally, I perform two tests which are similar to the ARCH test of Engle (1982). Using x_t and $Claims_t$, I estimate

$$x_t = a_0 + \sum_{j=1}^{p} a_j \hat{x}_{t-1} + \sum_{k=1}^{q} a_{p+k} Claims_{t-1} + v_t \tag{28}$$

for different choices of p and q. Then, I estimate b_1 and c_1 in

$$\begin{aligned} |\hat{v}_t| &= b_0 + b_1 Unc_t + \epsilon_{1t} && (29) \\ \hat{v}_t^2 &= c_0 + c_1 Unc_t + \epsilon_{2t}, && (30) \end{aligned}$$

where Unc_t is one of our uncertainty measures. Under the null hypothesis that shocks to x_t are homoskedastic, b_1 and c_1 are zero. I test this restriction in the second stage regression.[62] I also repeat the analysis with $Claims_t$ is the dependent variable in (28).

Table 5 presents the results from these heteroskedasticity tests. Rows correspond with different specifications for the conditional mean–i.e. choices of p and q in (28). The third column reports the first stage R^2's for each conditional mean model, which are generally quite high. The remaining columns report Newey-West t-statistics on b_1 and c_1 in (29-30) for different uncertainty measures. We use 4 lags, though results are insensitive to this choice. R^2's from these second-stage regressions are in brackets.

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

6.2 Labor market uncertainty predicts returns

My 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, I test a necessary condition for such a model by considering the ability of my preferred labor market uncertainty measure,

[62] The standard errors, as currently calculated, are conditional on the estimated first stage coefficients.

Dependent variable	Mean specification	First stage R^2	Absolute residuals			Squared residuals		
			$Claims_{t-1}$	x_{t-1}	Vol_t	$Claims_{t-1}$	x_{t-1}	Vol_t^2
Skewness (x_t)	AR(1)	75.0	3.50	-4.31	3.35	2.63	-3.65	-4.31
			[19.1]	[18.1]	[10.6]	[16.8]	[17.6]	[18.1]
Skewness (x_t)	AR(2)	80.6	2.73	-3.21	3.19	2.43	-3.08	-3.21
			[12.6]	[12.9]	[8.9]	[12.8]	[13.7]	[12.9]
Skewness (x_t)	AR(4)	80.8	2.69	-3.22	3.07	2.44	-3.15	-3.22
			[12.6]	[13.2]	[8.4]	[12.4]	[14.1]	[13.2]
Skewness (x_t)	VAR(1)	75.3	3.22	-4.09	3.39	2.55	-3.57	-4.09
			[17.5]	[17.3]	[10.1]	[14.6]	[17.2]	[17.3]
Skewness (x_t)	VAR(2)	88.0	3.73	-3.06	2.85	3.30	-3.27	-3.06
			[12.7]	[8.6]	[5.6]	[12.4]	[10.0]	[8.6]
Skewness (x_t)	VAR(4)	92.2	2.16	-2.35	3.15	2.23	-2.06	-2.35
			[2.8]	[3.5]	[6.4]	[2.6]	[3.8]	[3.5]
$Claims_t$	AR(1)	87.8	6.74	-3.42	4.08	3.94	-2.85	-3.42
			[21.4]	[6.7]	[27.4]	[10.1]	[3.1]	[6.7]
$Claims_t$	AR(2)	89.4	4.93	-2.98	4.25	2.70	-2.38	-2.98
			[23.6]	[5.4]	[27.6]	[9.5]	[1.9]	[5.4]
$Claims_t$	AR(4)	89.5	4.99	-2.95	4.25	2.65	-2.34	-2.95
			[22.8]	[4.9]	[27.6]	[9.3]	[1.7]	[4.9]
$Claims_t$	VAR(1)	87.9	6.50	-2.98	4.24	3.80	-2.57	-2.98
			[20.2]	[5.8]	[28.2]	[9.5]	[2.7]	[5.8]
$Claims_t$	VAR(2)	89.6	4.64	-2.79	4.19	2.68	-2.26	-2.79
			[22.5]	[5.2]	[27.1]	[9.6]	[1.9]	[5.2]
$Claims_t$	VAR(4)	89.9	4.72	-2.83	4.13	2.62	-2.30	-2.83
			[20.3]	[4.9]	[26.6]	[8.5]	[1.6]	[4.9]
Number of residuals			179	179	179	179	179	179

Table 5: Tests for heteroskedasticity of skewness index and initial claims residuals

This table presents the results of a test for the heteroskedasticity of my estimated skewness index (x_t, top panels) as well as initial claims for unemployment, $Claims_t$. Test statistics are generated using a two-step procedure. In the first stage, I estimate an AR(p) or VAR(p) model using x_t and $Claims_t$. I report the R^2 from this regression is in the third column. The left and right panels report Newey-West t-statistics (4 lags) and R^2 (in brackets) from regressions of absolute and squared residuals, respectively. The sample period is 1967:1 through 2012:3.

initial claims, to predict returns in the data. I also consider its covariance with and compare its forecasting power with other leading predictor variables from the extant literature. I find that initial claims for unemployment outperforms essentially all of the univariate predictors at short horizons (3 months to 1 year) and the vast majority of variables at a 2 year horizon. Moreover, I find that common components of the associated univariate forecasts track labor market conditions. Many variables which are motivated as proxies for aggregate consumption risk also contain important information about idiosyncratic risk.

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

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

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

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

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

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, divided by private sector employment, is my proxy for labor market uncertainty. Future employment cuts, is the negative of the logarithmic growth rate in private payroll employment over the next 3 months. The table also includes the Goyal and Welch (2008) predictors and combination forecasts which are constructed from these predictors. The first two combination forecasts are estimated using different sample periods. The last combination forecast uses predictors which are orthogonalized with respect to initial claims. Stars indicate statistical significance at the 1% level.

Table 6 presents a number of pairwise correlations between initial claims, each of the predictors, and a measure of employment declines over the next three months–a simple measure of labor market conditions. Initial claims has a 38% correlation with future employment declines and both of the combination forecasts (67% and 58%, respectively). It is even more strongly correlated with the dividend yield (74%), the book-to-market ratio (76%), and the default yield (69%). It is also positively correlated with the T-bill rate (41%), the long term yield (41%) on government bonds, and the inflation rate (24%), which is primarily driven by the period in the 1970s where both inflation and labor market uncertainty were elevated. More surprising is the negative correlation with the earnings-price ratio (-46%), which appears to be driven by

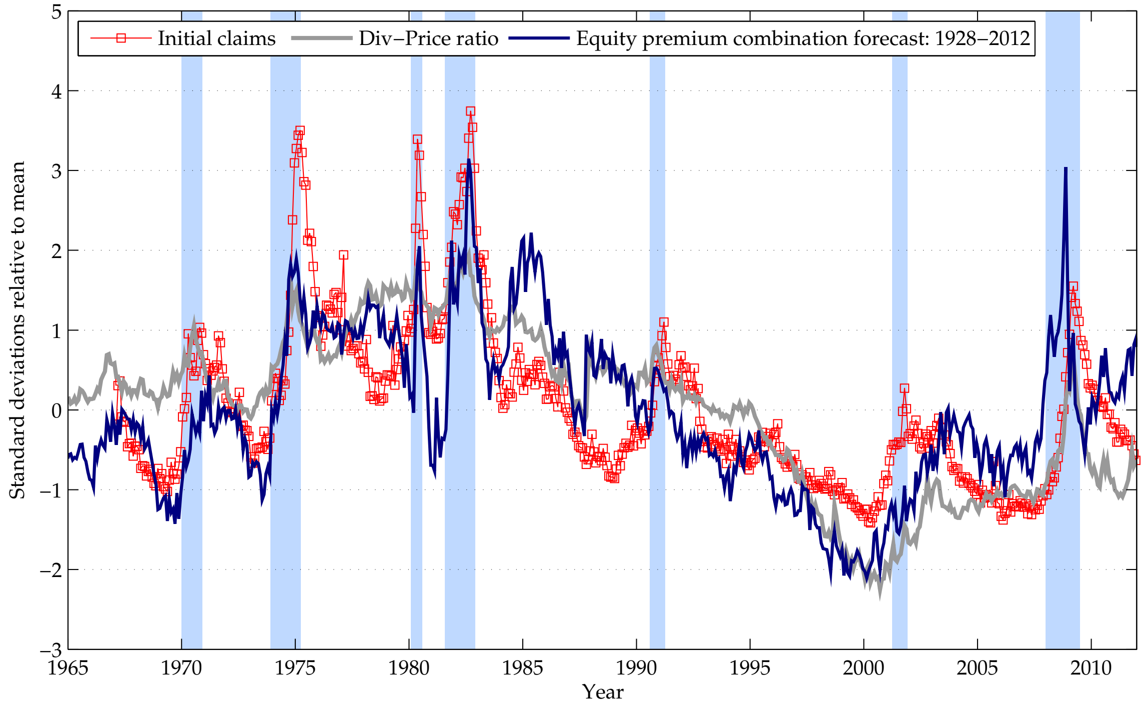

Figure 6: Co-movement of initial claims with representative equity premium forecasts

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.

differences in low frequency variation between the two measures.[63]

Next, I report the pairwise correlations between each of the equity premium combination forecasts and our predictor variables. The first of the combination forecasts is most strongly correlated with the dividend yield (72%), the book-to-market ratio (67%), and the default yield (62%). All three measures are highly correlated with initial claims, suggesting that they are all capturing a common macroeconomic risk factor. Figure 6 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.

[63]An even stronger negative correlation (-72%) arises between the dividend yield and the earnings price ratio.

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

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

Table 7 summarizes the forecasting performance of each of our predictor variables for cumulative returns. I report the R^2 and the t-statistic on β_h from the following predictive regression:

$$\sum_{j=1}^{h} r_{t+j} \equiv r_{t:t+h} = \alpha_h + \beta_h x_t + u_{t:t+h}, \tag{31}$$

where r_t is the log return on a given portfolio, x_t is the predictor variable, and h is the forecast horizon. Rows correspond with different predictors, while columns correspond with different portfolios and forecast horizons. I consider forecasts of the log excess return on the CRSP value-weighted index, as well as the Fama and French (1993) SMB portfolio.[64] I consider forecast horizons (h) of 3, 12, and 24 months, though results are similar at other horizons. My sample period is 1967-2012. In order to make an apples-to-apples comparison, when looking at the other predictors, I limit my attention to the period for which initial claims data are available.

The results in Table 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

[64]Results are qualitatively similar for the HML portfolio, though the statistical evidence is much weaker. None of the variables (including the combinations) forecast HML well at short horizons, though I find weak evidence that initial claims forecasts HML at long horizons.

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

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. I consider the market excess return as well as the Fama and French (1993) SMB portfolio. I use overlapping monthly data for the regressions, and the sample period is 1967-2012, the period for which initial claims data are available. Newey-West t-statistics, with lag length equal to the forecast horizon minus 1, are in parentheses.

achieves an R^2 of 2.4%. Initial claims outperforms every one of the Goyal and Welch (2008) predictors, and its performance is comparable with the first and third combination forecasts. The only other statistically significant univariate predictor is the term spread, which achieves an R^2 of 1.76%. At a 1 year horizon, the R^2 is 6.4%, which is statistically significant. Only the term spread performs better with an R^2 of 8%. At a 2 year horizon, claims performs a bit worse, though the magnitude of the R^2 is still reasonably high. The combination forecasts perform extremely well at the 1-2 year horizons.

A couple of other points are worth noting about the left panel of Table 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 I am taking an average of fitted values from 12 univariate regressions, all of whose coefficients are estimated using data from the period over which evaluation takes place.

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

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 health of the labor market, given the 41% correlation between the combination forecast constructed with orthogonalized predictors and future cuts in employment.

Predictor	Market Excess Return 3 mo	Market Excess Return 1 yr	Market Excess Return 2 yr	SMB Return 3 mo	SMB Return 1 yr	SMB Return 2 yr
Initial claims	2.40**	6.37**	4.96	4.29***	10.82***	9.47***
	(2.42)	(2.44)	(1.58)	(3.63)	(3.90)	(2.58)
Equal-weighted forecast combinations, orthogonalized						
1928 - 2012	0.52	3.23	4.62	0.47	0.22	0.55
	(0.81)	(1.34)	(1.53)	(0.90)	(0.19)	(-0.42)
1967 - 2012	1.50	8.82***	12.27***	1.03*	0.30	0.69
	(1.22)	(2.90)	(3.50)	(1.76)	(0.43)	(-0.65)
1967 - 2012	2.41	9.49***	10.70***	2.10**	1.62	0.32
(orth. predictors)	(1.64)	(2.82)	(3.57)	(2.53)	(1.51)	(0.32)
Univariate regressions with Goyal and Welch (2008) predictors, orthogonalized						
dy	0.28	0.21	0.56	0.75	3.22*	6.57*
	(-0.50)	(-0.14)	(0.40)	(-1.17)	(-1.66)	(-1.86)
ep	0.31	0.35	0.39	0.52	1.31	1.57
	(0.49)	(0.36)	(0.35)	(1.19)	(1.27)	(0.87)
bm	1.70**	3.15	3.62	0.69	0.66	0.39
	(-1.98)	(-1.44)	(-1.33)	(-1.16)	(-0.59)	(-0.28)
svar	0.31	0.81*	1.67**	0.45	1.57**	2.08*
	(-0.45)	(1.70)	(2.23)	(1.57)	(2.06)	(1.82)
tbl	2.25**	5.01**	4.86**	2.90***	4.64**	2.88
	(-2.30)	(-2.25)	(-2.08)	(-3.15)	(-2.37)	(-1.16)
tms	1.50*	6.94**	11.95***	0.72	1.03	0.19
	(1.82)	(2.44)	(3.35)	(1.28)	(1.02)	(-0.06)
dfy	0.18	0.25	0.42	0.31	0.18	0.26
	(0.05)	(0.24)	(0.42)	(0.72)	(0.05)	(0.19)
lty	1.49*	1.45	0.40	3.21***	5.24***	5.59*
	(-1.80)	(-1.09)	(-0.34)	(-3.90)	(-2.93)	(-1.84)
ntis	0.26	0.66	0.63	0.31	0.35	0.38
	(-0.34)	(-0.42)	(-0.46)	(-0.66)	(-0.35)	(-0.30)
infl	1.40	5.65***	3.27***	1.03**	0.18	0.57
	(-1.38)	(-3.07)	(-3.38)	(-1.98)	(0.09)	(0.76)
dfr	0.71	0.21	0.18	0.25	0.22	0.19
	(1.22)	(0.45)	(0.13)	(0.74)	(0.73)	(-0.33)
ltr	0.48	1.21***	0.54**	0.33	0.18	0.31
	(1.07)	(2.80)	(2.00)	(0.93)	(-0.09)	(-0.98)

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. I 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. I use overlapping monthly data for the regressions, and the sample period is 1967-2012, the period for which initial claims data are available. Newey-West t-statistics, with lag length equal to the forecast horizon minus 1, are in parentheses.

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, I explore the covariance structure between market return innovations and my proxies for the level of and uncertainty about idiosyncratic risk. Empirically, I find that while return innovations have little predictive content for contemporaneous measures, they are highly informative about future labor market conditions.

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

For a given observable variable y_t, my definition of the impulse response is $E[y_{t+k}|r_{t+1} - E_t[r_{t+1}] = v, \mathscr{F}_t] - E[y_{t+k}|r_{t+1} - E_t[r_{t+1}] = 0, \mathscr{F}_t]$ for different values of k. Given a set of conditioning variables, z_t, I decompose y_{t+k} as

$$y_{t+k} = \text{proj}(y_{t+k}|z_t) + [E_t(y_{t+k}) - \text{proj}(y_{t+k}|z_t)] + \epsilon_{t:t+k} \equiv \beta_k' z_t + \xi_t^k + \epsilon_{t:t+k}, \qquad (32)$$

where $\xi_{t:t+k}$ is a term reflecting potential misspecification of the conditional mean of y_{t+k} and $\epsilon_{t:t+k}$ is the "true" innovation. I additionally assume that, as is the case in my general theoretical framework, the conditional mean of returns takes the linear form

$$r_{t+1} = \gamma' z_t + v_{t+1}, \qquad (33)$$

where v_{t+1} has mean zero. Then, the impulse response may be rewritten as $E[\epsilon_{t:t+k}|v_{t+1} = v]$. One obtains consistent estimates of $\xi_t^k + \epsilon_{t:t+k}$ and v_{t+1} by taking the residuals from regressions of y_{t+k} and r_{t+1} on z_t, respectively. Given these residuals, I estimate $\text{proj}(\epsilon_{t:t+k}|v_{t+1} = 1) \equiv \alpha_h$ by regressing $\hat{\xi}_t^k + \hat{\epsilon}_{t:t+k}$ on $\hat{v}_{t+1}$. Inference is straightforward, since the estimate of α_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 ξ_t^k, so misspecification of the conditional mean adds noise to the dependent variable ($\hat{\xi}_t^k + \hat{\epsilon}_{t:t+k}$) of the second stage

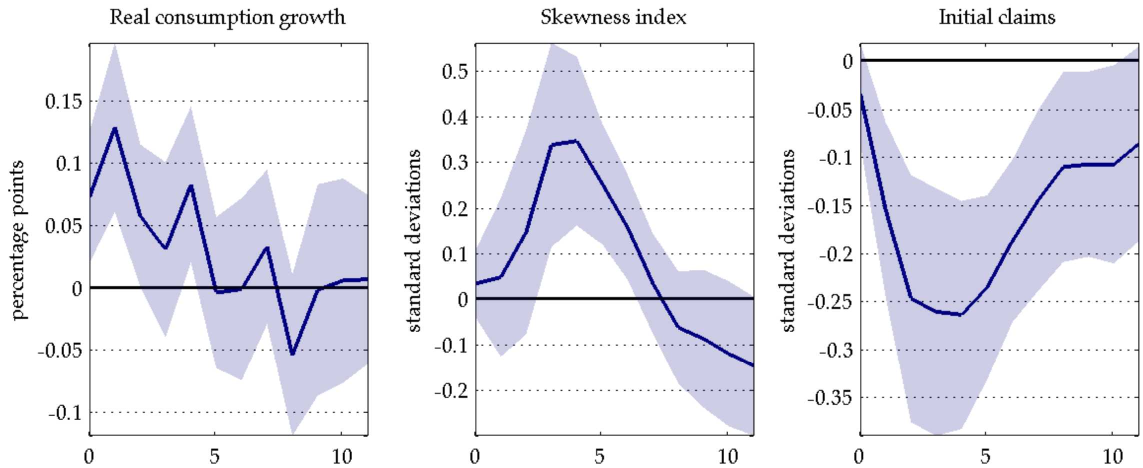

Figure 7: Model-free impulse responses to market excess return innovations

This figure plots model-free impulse responses of key macroeconomic variables to market excess return innovations. The impulse response is the slope coefficient on the market return, $r_{m,t+1}$, from a univariate regression of y_{t+k} on a vector of predictors, x_t, and $r_{m,t+1}$. The vector z_t includes y_t, the dividend yield, initial claims for unemployment, the term spread, and the 3 month T-bill rate. Shaded regions are pointwise 95% confidence bands, calculated using Newey-West standard errors, where the number of lags equals the horizon minus 1.

regression. As long as I have estimated the *return* innovation correctly, I need not have specified the mean of y_{t+k} correctly. The advantage of such an approach is that, in contrast with macroeconomic time series, returns are almost serially uncorrelated. While the conditional mean of returns does vary over time, this variation is second order compared with its highly volatile unforecastable component. However, the use of a direct estimation method places practical constraints on the maximum lag length which can be considered.

Figure 7 shows the estimated impulse response functions to market excess returns for six different macroeconomic variables over twelve quarters. The vector z_t includes 4 lags of the target variable, the dividend yield, initial claims for unemployment, the term spread, and the 3 month T-bill rate. For purposes of identification, it is more important that the variables z_t capture the conditional mean of returns, as opposed to the target variables. Including lags of the target helps to reduce noise in the estimation of the news terms, though, consistent with my identification argument, the results are insensitive to the inclusion of one or more lagged terms.

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

Next, I consider the response of my conditional skewness index. Since x_t is normalized to have an unconditional variance equal to one, the response is measured in standard deviation units. The response is small and insignificant on impact, peaking at about 1/3 of one standard deviation 3 to 4 quarters after the return innovation is observed. The response turns statistically insignificant around 7-8 quarters later. My point estimates are slightly negative, though insignificant, in the last four quarters. Such a result is consistent with a transitory component in idiosyncratic skewness which subsequently reverses itself.[65] The magnitude of the skewness response is quite substantial; the cumulative response over the first two years is 1.35 standard deviations.

Finally, I plot the response of my labor market uncertainty proxy, initial claims for unemployment, to return innovations. The response is hump-shaped and unambiguously negative. A positive return innovation is associated with a decrease in future labor market uncertainty, where the news is most informative about initial claims 6-18 months in the future. Here, the magnitudes are fairly substantial, given that the high persistence of initial claims.

7 Quantitative Model

Sections 4 and 5 show how to integrate my incomplete markets mechanism into a general, jump diffusion model for aggregate cash flows. The novel mechanism is that agents are exposed to rare, *idiosyncratic* disasters, where the idiosyncratic disaster probability is time-varying. In this section, I work with a fairly standard specification for aggregate risk so as to highlight the amplification in risk premia associated with incomplete markets. Despite its simplicity, the stylized model is matches key asset pricing moments quite well, without relying on low-frequency variation in state variable dynamics.

[65] If I reestimate the regressions with only lags of the targets in z_t (which would be valid if returns were unpredictable), these negative point estimates disappear.

7.1 Setup

I perturb the representative agent model so that agents are exposed to idiosyncratic, uninsurable event risk. For parsimony, I abstract away from diffusion (Gaussian) shocks and assume that all uninsurable risk comes from compound Poisson shocks.[66] Further, while my empirical results provide evidence of state dependence in both tails of the idiosyncratic risk distribution, the stylized model only emphasizes downside risk to keep the model as transparent as possible. As such, the novel mechanism in the model is the inclusion of a time-varying probability of uninsurable idiosyncratic disasters within an otherwise standard endowment economy.

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

$$\Delta c_{t+1} = \mu_c + \phi_c x_t + \sigma_t \epsilon_{c,t+1} \tag{34}$$

$$\Delta d_{t+1} = \mu_d + \phi_d x_t + \sigma_t \epsilon_{d,t+1}, \tag{35}$$

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

The persistent component, x_t, plays a second role in my model. In addition to affecting expected cash flow growth rates, it also controls the higher moments of idiosyncratic shocks to consumption. Agents are exposed to a single jump component, $J^i_{\eta,t+1}$, and its Poisson intensity $\lambda_{\eta,t+1}$–the personal disaster probability–is $\lambda_0 - \lambda_1 x_t$. λ_1 is positive, so that personal disasters become more likely in states when expected cash flow growth is low. As above, I normalize x_t to have mean zero and variance one, so that λ_0 and λ_1 can be interpreted as the unconditional disaster probability, and λ_1 is the sensitivity of the conditional disaster probability to a 1 standard deviation change in x_t.

The personal disaster magnitude (the jump size) is normally distributed with mean μ_b and standard deviation σ_b. Note that this choice is inconsistent with the calibration in Section 3.2 assumed exponentially-distributed jumps. In the interest of conservatism, I choose the thin-tailed normal shocks. In my calibration, μ_b is a large, negative number, and σ_b^2 is very large relative to uncertainty about aggregate consumption. Analogously with rare macroeconomic disasters, these infrequent labor market events are associated with extremely high marginal utilities, so they have a large impact on asset prices despite their relative infrequency.

[66] 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 δ.

The state variables evolve according to

$$x_{t+1} = \rho_x x_t + \sigma_t \epsilon_{x,t+1} \tag{36}$$
$$\sigma^2_{t+1} = 1 - \rho_\sigma + \rho_\sigma \sigma^2_t + \sigma_t \epsilon_{\sigma,t+1} \tag{37}$$

where we normalize $E[x_t] = 0$ and $E[\sigma^2_t] = 1$. In Bansal and Yaron (2004), the shock to σ^2_t is homoskedastic. I allow σ^2_t to follow a square-root process, which is consistent with evidence from Table 5. This restriction also guarantees that σ^2_t is nonnegative in the continuous time limit. Whereas popular calibrations of x_t emphasize very low frequency movements in expected growth rates and volatility, both state variables are considerably less persistent in my calibration.

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

$$E_t[\epsilon_{t+1}\epsilon'_{t+1}] = \begin{bmatrix} \varphi^2_c & \pi_c \varphi^2_c & 0 & 0 \\ \pi_c \varphi^2_c & \pi^2_c \varphi^2_c + \varphi^2_d + \pi^2_x \varphi^2_x + \pi^2_\sigma \varphi_\sigma & \pi_x \varphi^2_x & \pi_\sigma \varphi^2_\sigma \\ 0 & \pi_x \varphi^2_x & \varphi^2_x & \chi \varphi_x \varphi_\sigma \\ 0 & \pi_\sigma \varphi^2_\sigma & \chi \varphi_x \varphi_\sigma & \varphi^2_\sigma \end{bmatrix}. \tag{38}$$

The Bansal and Yaron (2004) specification assumes that π_c, π_x, π_σ, and χ are zero. Bansal et al. (2012) allow $\pi_c \neq 0$, which permits for a nonzero covariance between consumption and dividend innovations. Analogously, π_x and π_σ allow for a nonzero correlation between news about dividend growth and the state vector. χ permits a nonzero correlation between x_t and σ^2_t innovations, which has considerable support in the data.

7.2 Amplification

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

$$A_x = \left[\frac{1 - \frac{1}{\psi}}{1 - \rho_x \rho_c}\right] \left[\phi_c + \rho_x \frac{\partial \nu^*_t}{\partial x_t}\right], \tag{39}$$

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

From inspection of (39), the amplification mechanism associated with idiosyncratic risk becomes quite clear. Marginal utility becomes more sensitive to x_t as the cross-sectional certainty equivalent becomes more sensitive to x_t. Households face more risk, so their hedging demands against future increases in that risk are larger. In addition, as in the representative agent case, the sensitivity also increase as x_t becomes more persistent and/or aggregate consumption becomes more predictable. Though the algebraic expression for the coefficient on σ_t^2 is considerably messier, a similar amplification is present. Under my preference configuration, the price of volatility risk is negative. The addition of idiosyncratic risk makes it larger in absolute value relative to the representative agent case.

Turning to the price dividend ratios, one can also show

$$A_{x,m} = \frac{1}{1-\rho_x\rho_m}\left(\phi_x - \tfrac{\phi_c}{\psi} - \rho_x\left[\gamma\left(\frac{\partial\nu_t}{\partial x_t} - \frac{\partial\nu_t^*}{\partial x_t}\right) + \frac{1}{\psi}\frac{\partial\nu_t^*}{\partial x_t}\right]\right). \tag{40}$$

The contribution from the first two terms in parentheses come from the representative agent solution, whereas the last term (in brackets) comes from incomplete markets. The first term within the brackets in (40), which is generally positive, compares the sensitivity of the certainty equivalents of two invididuals, where one is more risk averse than the other, to changes in the higher moments of $\exp(\eta_{t+1}^i)$ given y_{t+1}.[67] The second term within the brackets is the incomplete markets analogue to the $-\frac{\phi}{\psi}$ term coming from the representative agent solution.

Inspection of the bracketed term in (40) reveals one of the potentially counterintuitive implications of the incomplete markets model. Idiosyncratic risk generally increases the sensitivity of the pricing kernel to shocks to the aggregate state vector. However, idiosyncratic risk pushes in the opposite direction for the price dividend ratios. All else constant, for a given degree of dividend predictability, returns will tend to fall less in response to bad news relative to the representative agent case.

[67] Recall that ν_{t+1}, which appears in the projected pricing kernel (21), is the log of a CRRA individual's certainty equivalent for lottery $\exp(\eta_{t+1}^i)$ given x_{t+1} when the risk aversion parameter is γ. Analogously, $\nu_{t+1}^* \equiv \frac{1}{1-\gamma}\log E_t[\exp((1-\gamma)\eta_{t+1}^i)|x_{t+1}]$ is a certainty equivalent for the same lottery when the risk aversion parameter is $\gamma - 1$. When $\gamma \geq 2$, the two partial derivatives will have the same sign, and $\frac{\partial\nu_{t+1}}{\partial x_{t+1}} \geq \frac{\partial\nu_{t+1}^*}{\partial x_{t+1}}$.

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

Taking stock, within the context of this stylized model, idiosyncratic risk increases the sensitivity of household marginal utility (the pricing kernel) to news about the state variables, but it reduces the sensitivity of returns. My quantitative exercise suggests that the increased sensitivity of the pricing kernel is more important. However, one can find parameter configurations where consumption is substantially riskier than dividends, causing stock prices to *increase* in response to bad news, which is counterfactual (see, e.g. the impulse responses in Figure 7). In order to generate a substantial risk premium, dividends must fairly predictable.[68]

7.3 Calibration

Table 9 provides an overview of the parameters in the quantitative model, along with the calibrated values. While the number of parameters is somewhat larger than Bansal and Yaron (2004), most of the key parameters will be tied directly to estimates from the data. I calibrate the model to the quarterly frequency, to match the frequency of my estimated skewness index, as well as the calibrated income process parameters from section 3.2.

I begin with the parameters governing idiosyncratic shocks. I set λ_1, the sensitivity of the disaster probability to a change in x_t, exactly equal to its calibrated value from Table 3. Given my emphasis only on the state-dependent component of idiosyncratic risk, I set the unconditional disaster probability equal to $2.5 \times \lambda_1$, which implies that the probability of the fitted intensity

[68]Constantinides and Ghosh (2014) calibrate a model with a relatively similar structure, except that idiosyncratic risk is driven by a single variable which follows a square root process. In their calibration, aggregate consumption and dividends are i.i.d. When idiosyncratic risk is sufficiently persistent, they generate a large equity premium even in the absence of predictability. This occurs because the level and volatility of idiosyncratic risk are perfectly correlated, so agents' preference for an early resolution of uncertainty causes prices to fall in response to bad news about future idiosyncratic risk.

Parameter	Value	Description
λ_0	0.0065	Average idiosyncratic jump intensity
λ_1	0.0026	Sensitivity of quarterly jump intensity to a one standard deviation change in x_t
μ_b	-0.18	Average consumption decline given a disaster
σ_b	0.115	Standard deviation of disaster magnitude
ρ_x	0.8847	Persistence of x_t process
ρ_σ	0.9446	Persistence of σ_t^2 process
μ_c	0.02 / 4	Drift of consumption growth
μ_d	0.0075 / 4	Drift of dividend growth
ϕ_c	0.000366	Loading of expected consumption growth on x_t
ϕ_x	0.025	Loading of expected dividend growth on x_t
φ_c	0.0125	Standard deviation of shock to Δc_t
φ_d	0.045	Standard deviation of independent shock to Δd_t
φ_x	$\sqrt{1-\rho_x^2}$	Standard deviation of shock to x_t
φ_σ	0.1674	Standard deviation of shock to σ_t^2
π_c	2.5	Loading of dividend innovation on Δc_t innovation
π_x	0.04	Loading of dividend innovation on x_t innovation
π_σ	-0.0896	Loading of dividend innovation on σ_t^2 innovation
χ	-0.66	Correlation of shocks to x_t and σ_t^2
γ	11	Relative risk aversion coefficient
δ	0.9745	Rate of time preference (quarterly)
ψ	2	Intertemporal elasticity of substitution

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 quarterly. The additional free parameters, μ_η and F_η, are assumed without loss of generality to equal the expressions given in Proposition 1.

going negative is very small, though still nonzero.[69]

Turning to the distribution of jump sizes (idiosyncratic disaster magnitudes), I choose parameters with an eye towards conservatism. I set $\mu_b = -18\%$ and $\sigma_b = 11.5\%$, values which are considerably smaller than the values in Table 3. Stockholders have some means with which to smooth their consumption over time, and households often have more than one earner. As such, I do not assume that income shocks translate one-for-one into consumption shocks. Instead, I make an assumption about the extent to which shocks to permanent income "pass through" to consumption–i.e. the elasticity of consumption growth to permanent income growth. These values of μ_b and σ_b translate to an elasticity of about 23%, which is on the lower end of reduced-form estimates from Blundell et al. (2008).[70]

By far, the most controversial parameters in the long-run risk literature govern the persistence of the state variables x_t and σ_t^2, an well as the degree of consumption and/or dividend growth predictability required to generate a sizable risk premium. I choose the persistence parameters, ρ_x and ρ_σ, to match the first order dynamics of my skewness index and initial claims, respectively. I estimate the AR(1) parameters using a regression which is adjusted for finite-sample bias following the approach in Bauer et al. (2012).[71] Analogously, the correlation between the AR(1) innovations, χ, is estimated directly from the bias-corrected regression residuals. I choose the volatility of σ_t^2 to match the Kelley's skewness of initial claims in the data.

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

[69] My calibration embeds a negative correlation between x_t and σ_t^2 shocks, implying that the unconditional distribution of x_t is negatively skewed (consistent with the data). Since the disaster probability decreases in x_t, the likelihood that the fitted intensity goes negative is quite small.

[70] Blundell et al. (2008), Table 7 estimates a 22.5% elasticity for male earnings. All other estimates are higher.

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

I choose the level of consumption predictability ϕ_c so that an agent's consumption is i.i.d. conditional on not receiving a jump shock. This choice implies that, unlike the Bansal and Yaron (2004) calibration, aggregate consumption growth is essentially unpredictable. ϕ_c exactly offsets the location adjustment (F_η) which is subtracted off to ensure proper aggregation. Given this restriction, the only source of predictability in Δc_{t+1} is the conditional expectation of the jump shock–a restriction which approximately holds in the income data (see Figure 4).[72] If, instead, I were to set $\phi_c = 0$, Δc_{t+1} is a random walk. However, the location adjustment would counterintuitively imply that, for all individuals who do not receive jump shocks, the distribution of consumption growth shifts to the right as the personal disaster probability increases.

I set the drift of aggregate dividend growth (μ_d) equal to 75 basis points per year, somewhat lower than the mean of aggregate consumption growth. I assume that a 1 standard deviation increase in x_t increases expected dividend growth by 2.5% in the following quarter. AGiven the degree of persistence in x_t, dividends are fairly predictable at short horizons, but are fairly difficult to predict at longer horizons. I set the loading of the dividend innovation on the consumption growth innovation, π_c, equal to 2.5, which is identical to the value in Bansal et al. (2012). The correlation between consumption growth and dividend growth innovations is 52%. My choice of φ_d implies that the quarterly volatility of the dividend growth innovation is just shy of 6%. The parameters π_x and π_σ imply that dividend innovations are moderately correlated with x_t and σ_t^2 shocks; pairwise correlations are 31% and -25%, respectively.

Table 10 compares my assumptions about state variable dynamics and cash flow predictability with two popular calibrations of the long run risk model: Bansal and Yaron (2004, BY) and Bansal, Kiku, and Yaron (2012, BKY). The top panel compares the persistence coefficients, which are expressed as monthly autocorrelations. In my model, the half life of an x_t shock is 1.4 years, which is considerably shorter than the half lives of 2.7 and 2.3 years in the BY and BKY calibrations, respectively. σ_t^2 is also less persistent. σ_t^2 shocks have a half life of 2.9 years versus 4.4 years in the BY model. The BKY model emphasizes extremely low frequency movements in volatility, so their choice of ρ_σ implies that a σ_t^2 shock has a half life of 57.7 years.

It is useful to clarify my purpose in making such a comparison. The BY and BKY models generate large risk premia which are quite close to those in my baseline calibration. While the first order autocorrelations of my skewness index and initial claims are indeed lower than standard choices of ρ_x and ρ_σ in the long run risk literature, this need not imply that idiosyncratic risk and labor market uncertainty feature important sources of low frequency variation. Rather, my key objective is to illustrate the amplification associated with incomplete markets.

[72]Wachter (2013) makes a similar assumption in a model with rare microeconomic disasters.

Statistic	Baseline	BY	BKY
Persistence of x_t (monthly)	0.96	0.979	0.975
Half-life of x_t shocks	1.4 years	2.7 years	2.3 years
Persistence of σ_t^2 (monthly)	0.980	0.987	0.999
Half-life of σ_t^2 shock	2.9 years	4.4 years	57.7 years
Volatility of expected consumption growth (annualized)	0.15%	2.0%	1.5%
Volatility of expected dividend growth (annualized)	10.0%	6.1%	3.7%
Dividend growth predictability: $E_t\left[\sum_{j=0}^{\infty}\rho^j[\Delta d_{t+1+j} - E(\Delta d_{t+1+j})\vert x_t = \sigma(x_t)\right]$			
0% annual discount rate	21.7%	24.6%	12.6%
5% annual discount rate	19.8%	20.6%	10.9%
10% annual discount rate	18.1%	17.6%	9.5%

Table 10: Comparison with benchmark long-run risk calibrations

This table compares several features of my calibrated model with comparable values implied by the parameters in Bansal and Yaron (2004, BY) and Bansal, Kiku, and Yaron (2012, BKY).

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

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

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

7.4 Performance

Table 11 demonstrates the ability of the quantitative model to match a number of key asset pricing moments. Data moments are taken from BKY, who calculate statistics using annual time series of real returns and cash flow growth rates from 1930-2008. I refer the reader to their paper for further details about the underlying data sources. Next, I use the model to simulate 50,000 annual time series of the same length, then report a number of quantiles of the finite sample distribution of the calibrated model. These quantiles can also be interpreted as robust standard errors for the model-implied moments.

Most importantly, my model with idiosyncratic disaster risk generates a large and time-varying equity premium of about 6.5% per year. It easily replicates the excess volatility puzzle; the volatility of the market return is 10% larger than that of dividend growth. The addition of incomplete markets leads to a fairly volatile real interest rate, whereas long run risk models with complete markets tend to exhibit too little volatility. The model also matches the level of the price-dividend ratio almost exactly, though it understates its volatility (a shortcoming of the BY and BKY models as well). The price dividend ratio also exhibits a lower degree of autocorrelation relative to the data, which is unsurprising given that the state variables in my model are not very persistent.

Looking at the cash flow moments, the biggest differences between the model and the data are the first order autocorrelations of consumption and dividend growth, which are significantly lower and higher than the corresponding values in the data, respectively. The former is by construction, given that I deliberately shut off almost all consumption predictability to highlight the amplification coming from incomplete markets.

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

Moment	Data Estimate	Model Median	2.5%	5%	25%	75%	95%	97.5%
$E[R_m - R_f]$	7.09	6.46	2.19	2.90	5.02	7.91	9.93	10.61
$\sigma(R_m)$	20.28	23.61	18.50	19.27	21.73	25.69	29.12	30.47
$E[R_f]$	0.57	0.46	-1.61	-1.22	-0.19	1.05	1.83	2.07
$\sigma(R_f)$	2.86	3.46	2.50	2.63	3.08	3.87	4.58	4.83
$E[pd]$	3.36	3.37	3.25	3.27	3.33	3.41	3.45	3.46
$\sigma(pd)$	0.45	0.21	0.15	0.16	0.19	0.23	0.28	0.30
$AC1(pd)$	0.87	0.58	0.35	0.39	0.51	0.65	0.73	0.76
$E[\Delta c]$	1.93	2.00	1.43	1.53	1.81	2.19	2.46	2.54
$\sigma(\Delta c)$	2.16	2.04	1.59	1.66	1.87	2.21	2.47	2.57
$AC1(\Delta c)$	0.45	0.21	-0.01	0.02	0.13	0.28	0.39	0.42
$E[\Delta d]$	1.15	0.83	-5.64	-4.46	-1.27	2.77	5.33	6.15
$\sigma(\Delta d)$	11.05	13.78	10.33	10.83	12.50	15.20	17.68	18.61
$AC1(\Delta d)$	0.21	0.50	0.28	0.32	0.43	0.57	0.66	0.68
$Corr(\Delta c, \Delta d)$	0.55	0.41	0.16	0.21	0.33	0.49	0.58	0.61

Table 11: Bootstrapped distribution of model-implied moments

This table presents several moments of aggregate cash flows and asset prices, both from the data and the model. The data moments are reproduced from Bansal, Kiku, and Yaron (2012), who use real, annual data from 1930-2008. The remaining columns show the Monte Carlo distributions of 50,000 simulated paths of analogous quantities, which are simulated from the calibrated model and time-aggregated to an annual frequency. Each simulated path has the same length as the historical data.

to medium-term horizons, increasing the autocorrelation of model-implied dividend growth. As discussed above, the overall level of dividend growth predictability in my model is comparable with Bansal and Yaron (2004).

Table 12 highlights the incremental contribution from incomplete markets by comparing the asset pricing moments from my model with those obtained from a comparable representative agent model. The Markets column indicates whether the relevant moment is obtained from the incomplete markets or representative agent version of the model. It reports the average of the moments in Table 11 of a long simulation of 1 million quarters. I shut off idiosyncratic risk by setting λ_0 and λ_1 equal to zero. In addition, I raise the rate of time preference considerably to 0.996, which generates a risk-free rate of about 2%.[73] In addition, I demonstrate the effects of shutting down several dimensions of risk which are embedded in my baseline calibration.

[73] In some of the specifications, price-dividend ratios approach infinity if I try to match the observed risk-free rate from the data.

Moment	Data Estimate	Simulated Model-Implied Moments				
		Markets	Baseline	No Covariance	Less CF Vol	IID Cons
$E[R_m - R_f]$	7.1	Inc	6.5	5.5	5.8	5.3
		RA	3.2	3.1	2.5	2.0
$\sigma(R_m)$	20.3	Inc	24.2	20.2	23.4	23.1
		RA	26.9	23.5	26.0	25.5
$E[R_f]$	0.6	Inc	0.4	0.4	0.6	0.7
		RA	2.0	2.0	2.0	2.0
$\sigma(R_f)$	2.9	Inc	3.7	3.7	3.5	3.3
		RA	0.3	0.3	0.1	0.0
$E[pd]$	3.4	Inc	3.4	3.4	3.5	3.6
		RA	4.5	4.1	5.1	6.2
$\sigma(pd)$	0.45	Inc	0.22	0.21	0.21	0.20
		RA	0.25	0.25	0.24	0.24
$AC1(pd)$	0.87	Inc	0.62	0.64	0.60	0.60
		RA	0.59	0.62	0.58	0.57
$E[\Delta c]$	1.9	Both	2.0	2.0	2.0	2.0
$\sigma(\Delta c)$	2.2	Both	2.1	2.1	2.1	2.1
$AC1(\Delta c)$	0.45	Both	0.23	0.23	0.23	0.23
$E[\Delta d]$	1.15	Both	0.68	0.75	0.70	0.80
$\sigma(\Delta d)$	11.1	Both	14.5	12.8	14.4	14.4
$AC1(\Delta d)$	0.21	Both	0.53	0.50	0.53	0.53
$Corr(\Delta c, \Delta d)$	0.55	Both	0.40	0.44	0.41	0.36

Table 12: Data and model-implied moments for different specifications

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

In the representative agent version of my baseline specification, the equity premium is generated by three distinct channels. The first is a contemporaneous covariance between consumption and dividend growth innovations. Second, since $\phi_c \neq 0$, there is a small amount of predictability in aggregate consumption growth. Third, there is stochastic volatility about aggregate consumption innovations, as well as shocks to the state vector. The Baseline column of Table 12 indicates that these channels combine to generate an equity premium of 3.2% per annum. Thus, the addition of incomplete markets doubles the risk premium in the baseline specification.

Looking at several of the other asset pricing moments, some general patterns emerge. The risk-free rates are considerably more volatile in the incomplete markets model, and price-dividend ratios are somewhat more autocorrelated. Returns are even more volatile in the representative agent versions of the model. This follows from (40), the expression for $A_{x,m}$. Holding dividend predictability constant, the additional precautionary savings motive associated with incomplete markets reduces the sensitivity of the price-dividend ratio to changes in x_t.

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

The next column reduces the importance of stochastic volatility relative to the baseline model. While shocks to x_t and σ_t^2 continue to be heteroskedastic, the terms involving φ_c and φ_d in the covariance matrix (38) are now assumed to be homoskedastic. This restriction affects both risk premia symmetrically, reducing both by about 0.7%.

The final column of Table 12, "IID Cons", keeps the reductions on the role of stochastic volatility from the "Less CF Vol" column and, in addition, sets $\phi_c = 0$. These restrictions imply that aggregate consumption is i.i.d. Again, the associated reduction in the risk premium of about 0.5% is the same across models. With i.i.d. aggregate consumption, the incomplete markets model generates a risk premium of 5.3% versus a 2% risk premium with complete markets.

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. My analysis suggests that risks associated with redistribution of consumption across across agents can be just as important, if not more important, than aggregate consumption risks. I view this paper's contribution as a "proof of concept"; there remains plenty of room for additional work.

Labor market event risk is likely to provide a novel mechanism for the amplification of aggregate shocks. If the uninsurability of labor market shocks causes discount rates to rise much more sharply in response to bad news than they would if markets were complete, firms' incentives to invest are likely to be substantially distorted. My model can be easily embedded within a production setting, and I plan to explore these interactions in future work.

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

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

References

Angeletos, George Marios, 2007, Uninsured idiosyncratic investment risk and aggregate saving, *Review of Economic Dynamics* 10, 1–30.

Bansal, Ravi, Dana Kiku, and Amir Yaron, 2010, Long run risks, the macroeconomy, and asset prices, in *American Economic Review*, volume 100, 542–546.

Bansal, Ravi, Dana Kiku, and Amir Yaron, 2012, An Empirical Evaluation of the Long-Run Risks Model for Asset Prices, *Critical Finance Review* 1, 183–221.

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

Barnichon, Regis, and Christopher Nekarda, 2012, The Ins and Outs of Forecasting Unemployment: Using Labor Force Flows to Forecast the Labor Market, *Brookings Papers on Economic Activity* 83–132.

Barro, RJ, 2006, Rare disasters and asset markets in the twentieth century, *The Quarterly Journal of Economics* 121, 823–866.

Bauer, Michael D., Glenn D. Rudebusch, and Jing Cynthia Wu, 2012, Correcting Estimation Bias in Dynamic Term Structure Models, *Journal of Business & Economic Statistics* 30, 454–467.

Beeler, Jason, and John Y Campbell, 2012, The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment, *Critical Finance Review* 1, 141–182.

Bekaert, Geert, and Eric Engstrom, 2013, Asset Return Dynamics under Habits and Bad-environment-Good Environment Fundamentals, *Working Paper* .

Berk, J. B., and J. Walden, 2013, Limited Capital Market Participation and Human Capital Risk, *Review of Asset Pricing Studies* 3, 1–37.

Berk, Jonathan B., Richard Stanton, and Josef Zechner, 2010, Human capital, bankruptcy, and capital structure, *Journal of Finance* 65, 891–926.

Bernanke, Ben S, Jean Boivin, and Piotr Eliasz, 2005, Measuring the Effects of Monetary Policy, *Quarterly Journal of Economics* 120, 387–422.

Blundell, By Richard, Luigi Pistaferri, and Ian Preston, 2008, Consumption Inequality and Partial Insurance, *American Economic Review* 98, 1887–1921.

Brav, Alon, George M. Constantinides, and Christopher C. Geczy, 2002, Asset Pricing with Heterogeneous Consumers and Limited Participation: Empirical Evidence, *Journal of Political Economy* 110, 793–824.

Camacho, Maximo, and Gabriel Perez-Quiros, 2010, Introducing the Euro-sting: Short-term indicator of Euro area growth, *Journal of Applied Econometrics* 25, 663–694.

Campbell, John Y, 1993, Intertemporal Asset Pricing without Consumption Data, *The American Economic Review* 83, 487–512.

Campbell, John Y, Stefano Giglio, Christopher Polk, and Robert Turley, 2012, An Intertemporal CAPM with Stochastic Volatility, *Working Paper* .

Campbell, John Y, and Tuomo Vuolteenaho, 2004, Bad Beta, Good Beta, *American Economic Review* 94, 1249–1275.

Campbell, JY, and RJ Shiller, 1988, The dividend-price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* 1, 1–34.

Cochrane, John H, 2005, *Asset Pricing* (Princeton University Press).

Cogley, Timothy, 2002, Idiosyncratic risk and the equity premium: Evidence from the consumer expenditure survey, *Journal of Monetary Economics* 49, 309–334.

Constantinides, George M., 2002, Rational Asset Prices, *The Journal of Finance* 57, 1567–1591.

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

Constantinides, George M, and Anisha Ghosh, 2014, Asset Pricing with Countercyclical Household Consumption Risk, *Working Paper* .

Davis, Steven, and Till Von Wachter, 2011, Recessions and the Costs of Job Loss, *Brookings Papers on Economic Activity* .

Dew-Becker, Ian, and Stefano Giglio, 2013, Asset pricing in the frequency domain : theory and empirics, *Working Paper* .

Drechsler, Itamar, and Amir Yaron, 2011, What's vol got to do with it, *Review of Financial Studies* 24, 1–45.

Engle, Robert F, 1982, Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica* 50, 987–1007.

Epstein, Larry G, and Stanley E Zin, 1989, Susbstitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework, *Econometrica* 57, 937–969.

Eraker, Bjorn, and Ivan Shaliastovich, 2008, An equilibrium guide to designing affine pricing models, *Mathematical Finance* 18, 519–543.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Gabaix, Xavier, 2012, Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance, *Quarterly Journal of Economics* 127, 645–700.

Ghysels, Eric, Alberto Plazzi, and Rossen Valkanov, 2013, Conditional Skewness of Stock Market Returns in Developed and Emerging Markets and its Economic Fundamentals , *Working Paper* .

Gourio, François, 2012, Disaster Risk and Business Cycles, *American Economic Review* 102, 2734–2766.

Gourio, François, 2013, Credit risk and disaster risk, *American Economic Journal: Macroeconomics* 5, 1–34.

Goyal, Amit, and Ivo Welch, 2008, A comprehensive look at the empirical performance of equity premium prediction, *Review of Financial Studies* 21, 1455–1508.

Guiso, Luigi, Luigi Pistaferri, and Fabiano Schivardi, 2005, Insurance within the Firm.

Guvenen, Fatih, Greg Kaplan, and Jae Song, 2014a, How Risky Are Recessions for Top Earners?, *American Economic Review* 104, 148–153.

Guvenen, Fatih, Fatih Karahan, Serdar Ozkan, and Jae Song, 2014b, What Do Data on Millions of U . S . Workers Say About Life Cycle Earnings Risk ?, *Working Paper* .

Guvenen, Fatih, Serdar Ozkan, and Jae Song, 2014c, The Nature of Countercyclical Income Risk, *Journal of Political Economy* 122, 621–660.

Hall, Robert E, 2014, High Discounts and High Unemployment, *National Bureau of Economic Research Working Paper Series* No. 19871.

Hamilton, James D., and Michael T. Owyang, 2011, The Propagation of Regional Recessions.

Harris, Milton, and Bengt Holmstrom, 1982, A Theory of Wage Dynamics, *The Review of Economic Studies* 49, 315.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante, 2014, Consumption and labor supply with partial insurance: An analytical framework, *American Economic Review* 104, 2075–2126.

Herskovic, Bernard, Bryan T Kelly, Hanno Lustig, and Stijn Van Nieuwerburgh, 2014, The Common Factor in Idiosyncratic Volatility: Quantitative Asset Pricing Implications, *Working Paper* .

Huggett, Mark, and Greg Kaplan, 2013, The money value of a man, *Working Paper* .

Ilut, Cosmin, Matthias Kehrig, and Martin Schneider, 2014, Slow to Hire, Quick to Fire: Employment Dynamics with Asymmetric Responses to News, *Working Paper* .

Jordà, Òscar, 2005, Estimation and inference of impulse responses by local projections, *American Economic Review* 95, 161–182.

Julliard, Christian, and Anisha Ghosh, 2012, Can rare events explain the equity premium puzzle, *Review of Financial Studies* 25, 3037–3076.

Kelly, Bryan, and Seth Pruitt, 2014, The Three-Pass Regression Filter : A New Approach to Forecasting Using Many Predictors, *Journal of Econometrics* .

Krebs, Tom, 2003, Human Capital Risk and Economic Growth, *The Quarterly Journal of Economics* 118, 709–744.

Krebs, Tom, 2007, Job Displacement Risk and the Cost of Business Cycles, *American Economic Review* 97, 664–686.

Kuehn, Lars-Alexander, Nicolas Petrosky-Nadeau, and Lu Zhang, 2013, An Equilibrium Asset Pricing Model with Labor Market Search, *National Bureau of Economic Research Working Paper Series* No. 17742.

Lagakos, David, and Guillermo L. Ordoñez, 2011, Which workers get insurance within the firm?, *Journal of Monetary Economics* 58, 632–645.

Lamont, Owen A., 2001, Economic tracking portfolios, *Journal of Econometrics* 105, 161–184.

Lettau, Martin, and Jessica A. Wachter, 2007, Why is long-horizon equity less risky? A duration-based explanation of the value premium, *Journal of Finance* 62, 55–62.

Lucas, Robert E., 1978, Asset Prices in an Exchange Economy, *Econometrica* 46, 1429–1445.

Malloy, Christopher J., Tobias J. Moskowitz, and Annette Vissing-JØ rgensen, 2009, Long-run stockholder consumption risk and asset returns, *Journal of Finance* 64, 2427–2479.

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

Mariano, Roberto S., and Yasutomo Murasawa, 2003, A new coincident index of business cycles based on monthly and quarterly series, *Journal of Applied Econometrics* 18, 427–443.

Mckay, Alisdair, and Tamas Papp, 2012, Accounting for Idiosyncratic Wage Risk Over the Business Cycle, *Working Paper* .

Parker, Jonathan A., and Annette Vissing-Jorgensen, 2010, The Increase in Income Cyclicality of High-Income Households and Its Relation to the Rise in Top Income Shares.

Rapach, David E., Jack K. Strauss, and Guofu Zhou, 2010, Out-of-sample equity premium prediction: Combination forecasts and links to the real economy, *Review of Financial Studies* 23, 821–862.

Rietz, TA, 1988, The Equity Risk Premium - A Solution, *Journal of Monetary Economics* 22, 117–131.

Shephard, N. G., and A C Harvey, 1990, On the Probability of Estimating a Deterministic Component in the Local Level Model, *Journal of Time Series Analysis* 11.

Shleifer, A, and RW Vishny, 1992, Liquidation values and debt capacity: A market equilibrium approach, *The Journal of Finance* 47, 1343–1366.

Storesletten, Kjetil, Chris I Telmer, and Yaro, 2004, Cyclical Dynamics in Idiosyncratic Labor Market Risk, *Journal of Political Economy* 112.

Storesletten, Kjetil, Christopher I. Telmer, and Amir Yaron, 2007, Asset pricing with idiosyncratic risk and overlapping generations, *Review of Economic Dynamics* 10, 519–548.

Timmermann, Allan, 2006, Chapter 4 Forecast Combinations.

Toda, Alexis Akira, 2014a, Asset Prices and Efficiency in a Krebs Economy, *Working Paper* .

Toda, Alexis Akira, 2014b, Incomplete Market Dynamics and Cross-Sectional Distributions, *Working Paper* .

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

Van Binsbergen, Jules, Michael Brandt, and Ralph Koijen, 2012, On the timing and pricing of dividends.

Wachter, Jessica A., 2013, Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?, *Journal of Finance* 68, 987–1035.

Wu, Jing Cynthia, and Fan Dora Xia, 2014, Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound, *Working Paper* .

Appendix [Under Construction]

A Idiosyncratic Risk Process - Calibration and Estimation

A.1 Additional summary statistics for income growth rates

Table 1 reported a number of statistics for the cross-sectional distribution of income growth rates, which were averaged over the 91^{st} through 95^{th} percentiles of the earnings distribution. Table 13 reports the same statistics, averaged over different percentiles of the distribution. These ranges are indicated by different columns, where [96,100] in the first column indicates that we 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.

Regardless of the specific group (column of Table 13) considered, the results are consistent with the discussion in the main text. The overall level of idiosyncratic risk is extremely high, with the level of the 90-10 spread exceeding 100 log percentage points for all groups. The entire distribution shifts to the left in recessions. Scale measures are relatively insensitive to the business cycle, while the extreme tails move much more strongly. Finally, including the highest group of earners increases the overall degree of risk substantially.

A.2 Time Aggregation of Third Central Moments

A.2.1 Log-linear approximation

This section derives expressions for the moments of time-aggregated wages from my quarterly model in (1-2). For notational simplicity, I suppress i subscripts here. As in (3) A first-order

Statistic	Period	Average over percentiles of 5-year average income distribution					
		[96,100]	[91,100]	[91,95]	[76,100]	[76,95]	[76,90]
Median	E	2.71	2.12	1.53	1.22	0.84	0.61
	R	-4.46	-3.44	-2.41	-2.99	-2.62	-2.68
	R - E	-7.50	-5.65	-3.81	-4.19	-3.37	-3.22
10th Percentile	E	-76.48	-69.55	-62.62	-63.99	-60.87	-60.28
	R	-95.34	-84.79	-74.23	-76.09	-71.27	-70.29
	R - E	-18.35	-14.93	-11.51	-11.91	-10.31	-9.90
90th Percentile	E	76.63	64.91	53.19	52.59	46.58	44.37
	R	53.51	47.23	40.96	39.27	35.72	33.97
	R - E	-20.47	-16.20	-11.93	-12.80	-10.88	-10.53
Kelley's Skewness	E	-3.22	-7.03	-10.85	-12.65	-15.01	-16.40
	R	-22.30	-23.52	-24.73	-27.23	-28.46	-29.70
	R - E	-19.09	-16.49	-13.89	-14.58	-13.45	-13.31
Scale Measures							
Inter-Quartile Range	R - E	0.43	0.86	1.30	0.78	0.87	0.73
90-10 Percentile Spread	R - E	-4.26	-2.44	-0.62	-1.22	-0.46	-0.40
Left Tail Width Measures							
50-25 Percentile Spread	R - E	4.08	3.55	3.01	2.67	2.32	2.09
50-10 Percentile Spread	R - E	10.85	9.28	7.70	7.72	6.94	6.69
Right Tail Width Measures							
75-50 Percentile Spread	R - E	-3.66	-2.69	-1.72	-1.89	-1.45	-1.36
90-50 Percentile Spread	R - E	-12.97	-10.54	-8.12	-8.61	-7.52	-7.32
95-50 Percentile Spread	R - E	-19.06	-17.26	-15.47	-15.46	-14.56	-14.26
99-50 Percentile Spread	R - E	-29.81	-28.26	-26.71	-25.82	-24.82	-24.19

Table 13: 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.

Taylor expansion yields that, for $k \geq 4$,

$$
\begin{aligned}
w_{A,t} - w_{A,t-k} &\approx \frac{1}{4}\Delta w_t + \frac{1}{2}\Delta w_{t-1} + \frac{3}{4}\Delta w_{t-2} + \sum_{j=3}^{k-1}\Delta w_{t-j} + \frac{3}{4}\Delta w_{t-k} + \frac{1}{2}\Delta w_{t-k-1} + \frac{1}{4}\Delta w_{t-k-2}. \\
&= \beta \cdot k + \frac{1}{4}\eta_t + \frac{1}{2}\eta_{t-1} + \frac{3}{4}\eta_{t-2} + \sum_{j=3}^{k-1}\eta_{t-j} + \frac{3}{4}\eta_{t-k} + \frac{1}{2}\eta_{t-k-1} + \frac{1}{4}\eta_{t-k-2} \\
&+ \frac{1}{4}\rho(L)(\eta_t + \eta_{t-1} + \eta_{t-2} + \eta_{t-3}) - \frac{1}{4}\rho(L)(\eta_{t-k} + \eta_{t-k-1} + \eta_{t-k-2} + \eta_{t-k-3}) \\
&+ \underbrace{\frac{1}{4}(\epsilon_t + \epsilon_{t-1} + \epsilon_{t-2} + \epsilon_{t-3}) - \frac{1}{4}(\epsilon_{t-k} + \epsilon_{t-k-1} + \epsilon_{t-k-2} + \epsilon_{t-k-3})}_{\equiv \epsilon_{A,t} - \epsilon_{A,t-k}} \\
&\equiv \beta \cdot k + \theta_k(L;\rho)\eta_t + \epsilon_{A,t} - \epsilon_{A,t-k}, \qquad (41)
\end{aligned}
$$

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

$$
\begin{aligned}
M_3[y_{A,t} - y_{A,t-k}] &= k^3 M_3[\beta] + \sum_{j=0}^{\infty}[\theta_{k,j}(\rho)]^3 M_3[\eta_{t-j}] + M_3[\epsilon_{A,t} - \epsilon_{A,t-k}], \\
&\equiv \phi_k(L;\rho)M_3[\eta_t] + k^3 M_3[\beta] + M_3[\epsilon_{A,t} - \epsilon_{A,t-k}], \qquad (42)
\end{aligned}
$$

where $M_3(\cdot)$ denotes the third central moment, conditional on aggregate information.[74] If we further assume that $M_3(\eta_t) = a + b'y_t$, where y_t is a vector of observable state variables, then

$$
M_3[w_{A,t} - w_{A,t-k}] = c_k + b'\phi_k(L;\rho)z_t, \qquad (43)
$$

where $\phi_k(L;\rho)$ is a known lag polynomial and $c_k \equiv \phi_k(1;\rho)a + k^3 M_3[\beta] + M_3[\epsilon_{A,t} - \epsilon_{A,t-k}]$, which is constant given our assumption that the third moments of β_i and ϵ_t are state independent.

A.2.2 Approximation accuracy

[Under Construction]

[74]This follows because, given two independent random variables x and z with $\mu_x \equiv E[x]$ and $\mu_z \equiv E[z]$,

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

where we use independence to replace terms such as $E[(x-\mu_x)^2(z-\mu_z)]$ with $E[(x-\mu_x)^2]E[(z-\mu_z)] = 0$.

A.3 Skewness Index Estimation

A.3.1 Skewness Indices - Parametric Approach

Motivation for Explanatory Variables

I consider four macroeconomic time series for inclusion in the vector y_t. The first variable, Δemp_t, is the quarterly change in the logarithm of private payroll employment. In section 2, I found that the tails of the income growth distribution are much more sensitive to the cycle relative to the center. If tail events are related to transitions between jobs, one would expect to see more large positive shocks and fewer large negative shocks when firms are hiring, generating a positive relation between Δemp_t and cross-sectional skewness.

The second variable, Δy_t, is the quarterly change in real compensation to private sector employees, which is essentially the first moment of the cross-sectional distribution of income growth rates. If changes in the first moment are driven by changes in the tails, one would expect to see a positive relation between Δy_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.[75] This variable captures a potential timing mismatch between shocks received by firms and those received by workers. Relative to profits, the response of wages to aggregate shocks is more sluggish, generating cyclical variation in overall profitability. If profits and wages are cointegrated, pw_{t-1} can be interpreted as an error-correction term. Thus, when profits are high relative to wages, it is likely that firm recently experienced a series of favorable shocks. Future wages are likely to be higher and firms are more likely to be hiring than firing, causing the right tail of the income growth distribution to expand and the left tail to contract.

Additional motivation for Δy_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

[75] I filter the series to eliminate very low frequency movements in this ratio, which could be related to changes in the composition of the private sector relative to the economy as a whole over time. As such, I use a smoothing parameter of 12,800, 8 times higher than the standard quarterly choice of 1600. Similar results obtain if the series is detrended by using a 10-year backward-looking moving average.

firms encounter financial distress, dissolving completely if the firm goes bankrupt. Workers whose contracts are terminated experience sudden, large declines in wages–i.e. idiosyncratic "disaster risk" arises as an equilibrium outcome of the model.[76]

Berk et al. (2010) is a partial equilibrium model, lacking any sources of aggregate risk. However, if one takes the structure of their optimal contract as given and applies it to a world where aggregate productivity is time-varying, the implications for the cross-sectional distribution of income growth rates are relatively clear. When productivity increases, firms' average profit margins increase, making it easier for firms to insure workers against bad shocks in the future. When profitability increases, average wages also increase. Conversely, firms have lower risk-bearing capacity when overall profitability is low and wages are falling, increasing the risk that workers experience large negative shocks and making the income growth rate distribution more negatively skewed.[77]

The last variable, Δc_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 Δy_t. One would expect the distribution of income growth rates to be more negatively skewed when household consumption is falling. However, compared with Δy_t, there is more scope for a timing mismatch between Δc_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 Δc_t to lead the cross-sectional moments.

Results

As a precursor to our regressions, Figure 8 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-

[76]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."

[77]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.

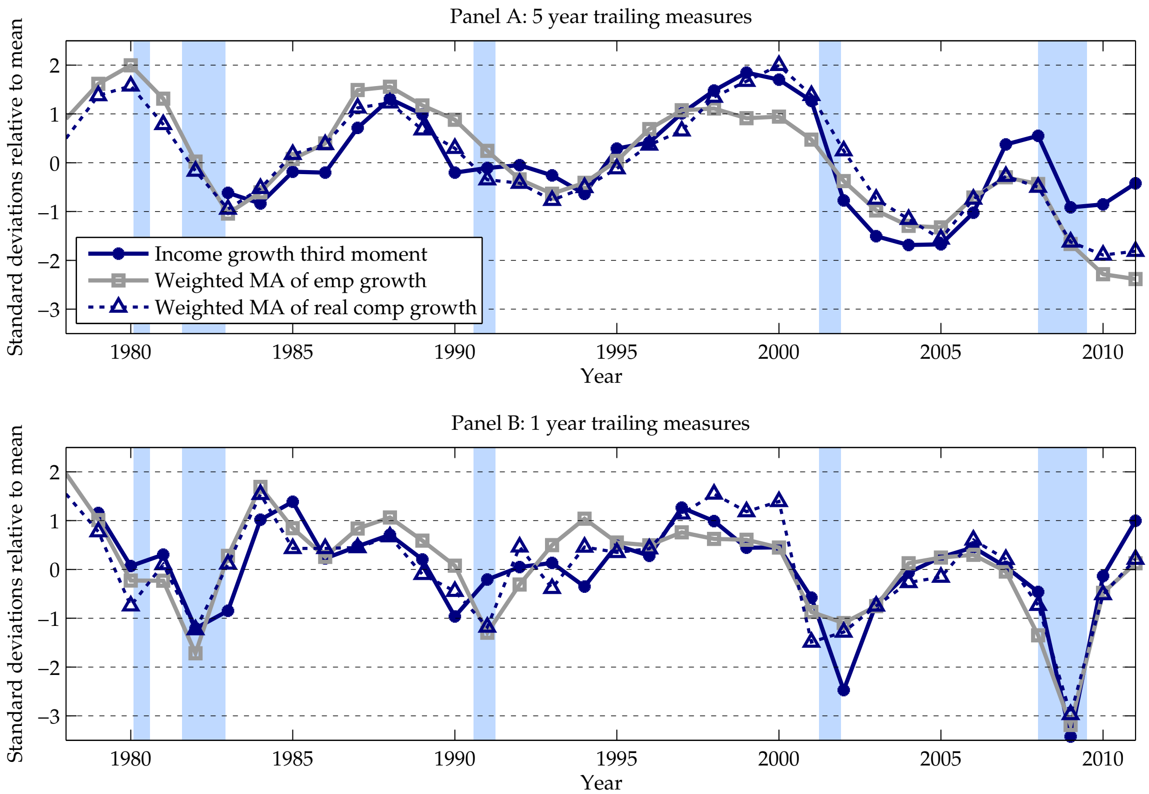

Figure 8: Co-movement of aggregate variables with third central moment of idiosyncratic income growth rates

Panel A plots the co-movement of 5-year idiosyncratic third central moments from GOS with weighted moving averages of logarithmic employment growth and real compensation growth. Panel B repeats the analysis for a 1 year measures. Series are standardized to have mean zero and unit variance. The weights are the lag polynomials $\phi_{20}(L;0)$ and $\phi_4(L;0)$ for 5 year and 1 year changes, respectively, which are defined in equation (42).

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.[78] 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.

[78] See Table 14, Panel A.

Table 14 estimates the vector b in (4) 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: $\rho(L) = \rho \cdot [1 + L]$.[79] In this latter specification, the partial derivative of y_{it} with respect to η_{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 ρ%. All estimates are obtained by minimizing the sum of squared residuals, which is an OLS regression when ρ is held fixed, and nonlinear least squares otherwise.[80]

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 only. When estimating these univariate regressions in Panel B, we fix the value of ρ at its estimated value from the bivariate model.[81]

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 Δy_t and Δc_t are both generally significant but enter with opposite signs, are somewhat less intuitive. However, the time series of quarterly skewness measures from this model track those from the other, more intuitive models relatively closely.

Panel A of Figure 9 plots our estimates of quarterly conditional third central moments, $\hat{b}' z_t$, from the pooled GMM estimates of models 5-7 from Panel B of Table 14. 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.

[79] Similar results obtain with different lag lengths.

[80] 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.

[81] Thus, the associated standard errors are best interpreted as conditional on $\hat{\rho}$.

		Panel A: Specifications with $\rho(L) = 0$						
		Pooled GMM			5 year only		1 year only	
Model	Variable	Coefficient	$\hat{\rho}$	R_5^2, R_1^2	Coefficient	R^2	Coefficient	R^2
1	Δemp_t	1.7959***	0	0.599	1.5640***	0.612	6.1042***	0.677
		(0.415)		0.340	(0.424)		(0.865)	
2	Δy_t	1.1325***	0	0.708	0.9944***	0.722	3.1307***	0.666
		(0.143)		0.395	(0.152)		(0.434)	
3	pw_{t-1}	0.0438***	0	0.610	0.0402***	0.615	0.1303***	0.366
		(0.006)		0.205	(0.007)		(0.020)	
4	Δc_t	1.5459***	0	0.233	1.3163***	0.241	6.0070***	0.348
		(0.383)		0.156	(0.388)		(1.691)	
5	Δemp_t	1.2245***	0	0.766	0.9908***	0.780	5.1872***	0.715
		(0.236)		0.353	(0.214)		(0.991)	
	pw_{t-1}	0.0257***			0.0256***		0.0497***	
		(0.006)			(0.006)		(0.018)	
6	Δy_t	0.8456***	0	0.808	0.6895***	0.823	2.7027***	0.689
		(0.129)		0.391	(0.093)		(0.666)	
	pw_{t-1}	0.0202***			0.0210***		0.0407	
		(0.007)			(0.006)		(0.026)	
7	Δy_t	1.8210***	0	0.853	1.6517***	0.866	3.5460***	0.673
		(0.123)		0.432	(0.104)		(0.584)	
	Δc_t	-1.9281***			-1.8194***		-1.4021	
		(0.230)			(0.249)		(1.840)	

		Panel B: Specifications with $\rho(L) = \rho \cdot [1 + L]$						
		Pooled GMM			5 year only		1 year only	
Model	Variable	Coefficient	$\hat{\rho}$	R_5^2, R_1^2	Coefficient	R^2	Coefficient	R^2
1	Δemp_t	1.5676***	0.5763	0.545	1.3332***	0.562	2.9610***	0.629
		(0.473)	(0.371)	0.489	(0.371)		(0.455)	
2	Δy_t	0.9565***	0.6002**	0.676	0.8390***	0.690	1.4730***	0.651
		(0.160)	(0.237)	0.571	(0.152)		(0.196)	
3	pw_{t-1}	0.0433***	-1.3590***	0.614	0.0393***	0.620	0.0891***	0.374
		(0.007)	(0.285)	0.275	(0.008)		(0.016)	
4	Δc_t	1.2088***	0.8957*	0.239	1.0426***	0.246	1.7950***	0.278
		(0.376)	(0.448)	0.248	(0.369)		(0.650)	
5	Δemp_t	1.1851***	0.4639	0.702	0.9456***	0.720	2.8498***	0.676
		(0.261)	(0.313)	0.495	(0.208)		(0.617)	
	pw_{t-1}	0.0221***			0.0217***		0.0295***	
		(0.007)			(0.005)		(0.010)	
6	Δy_t	0.7952***	0.5240**	0.752	0.6568***	0.768	1.3848***	0.678
		(0.125)	(0.243)	0.566	(0.116)		(0.326)	
	pw_{t-1}	0.0154*			0.0158***		0.0224*	
		(0.008)			(0.005)		(0.012)	
7	Δy_t	1.5525***	0.4392**	0.810	1.4137***	0.822	2.1247***	0.670
		(0.140)	(0.217)	0.574	(0.115)		(0.372)	
	Δc_t	-1.5139***			-1.4530***		-1.1432	
		(0.305)			(0.246)		(1.124)	

Table 14: Regressions of third central moment of income growth on aggregate variables

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

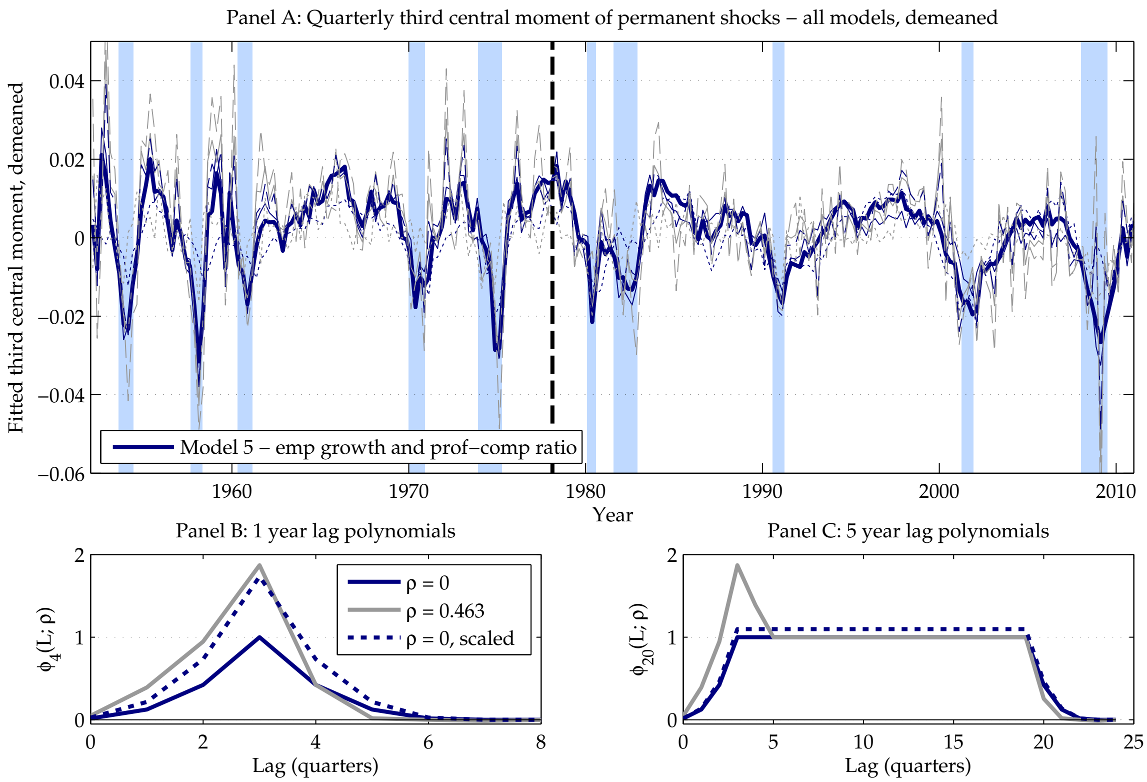

Figure 9: Key features from estimated regression models

Panel A plots pooled GMM estimates of quarterly conditional third central moments from the estimated specifications in Panel B of Table 14, i.e. $\hat{b}' z_t$. A dashed vertical line indicates the beginning of the sample period used to estimate the skewness measures. Panel B plots the coefficients of the lag polynomial $\phi_4(L;\rho)$ in the regression equation (4) for 1 year skewness measures. The first series imposes $\rho(L) = 0$, while the second corresponds with the estimated $\rho(L) = \hat{\rho}(1 + L)$ from Model 5. The third line rescales the first line so that the sum of the weights is the same as that from the estimated specification. Panel C repeats the analysis in Panel B for 5 year measures.

Moreover, a quarterly NBER recession indicator has almost no explanatory power.

With the exception of Δc_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 Δc_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 impose 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 9 offer an explanation for such a result. Panel B plots the coefficients of the lag polynomial $\phi_4(L; \rho)$ in the regression equation (4) for 1 year skewness measures. First, 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 9, Panel C shows the corresponding weights in the lag polynomials for 5 year measures. Once again, the specification with $\rho > 0$ puts higher weight on recent lags. However, as 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 ρ, and the weighting functions have essentially identical shapes after the 5^{th} lag. Accordingly, changes in ρ will have much larger effects on the

1-year moving averages relative to the 5-year moving averages.

Looking at Panel B of Table 14, 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 Δy_t and Δemp_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.

[Volatility Tests]

A.3.2 Implementation Details - 3PRF Approach

Table 15 lists the variables which I use to construct my skewness index. Following Wu and Xia (2014), 97 of the variables are obtained from Global Insight, which is the subset of 120 series from Bernanke et al. (2005) which are available through the present. I augment these time series with 12 variables from the literature on return predictability, which are taken from Ivo Welch. The second column provides the Global Insight or Goyal and Welch (2008) mnemonic for each series, and the final column indicates which transformation, if any, I perform to the original time series.

[Estimation details - under construction]

A.3.3 Quasi-analytic calculation of conditional quantiles

[Characteristic functions]

[Fourier inversion]

[Approximation accuracy]

	Mneumonic	Description	Transform
		REAL OUTPUT AND INCOME	
1	IPS11.M	INDUSTRIAL PRODUCTION INDEX - PRODUCTS, TOTAL	Δ log
2	IPS299.M	INDUSTRIAL PRODUCTION INDEX - FINAL PRODUCTS	Δ log
3	IPS12.M	INDUSTRIAL PRODUCTION INDEX - CONSUMER GOODS	Δ log
4	IPS13.M	INDUSTRIAL PRODUCTION INDEX - DURABLE CONSUMER GOODS	Δ log
5	IPS18.M	INDUSTRIAL PRODUCTION INDEX - NONDURABLE CONSUMER GOODS	Δ log
6	IPS25.M	INDUSTRIAL PRODUCTION INDEX - BUSINESS EQUIPMENT	Δ log
7	IPS32.M	INDUSTRIAL PRODUCTION INDEX - MATERIALS	Δ log
8	IPS34.M	INDUSTRIAL PRODUCTION INDEX - DURABLE GOODS MATERIALS	Δ log
9	IPS38.M	INDUSTRIAL PRODUCTION INDEX - NONDURABLE GOODS MATERIALS	Δ log
10	IPS43.M	INDUSTRIAL PRODUCTION INDEX - MANUFACTURING (SIC)	Δ log
11	IPS311.M	INDUSTRIAL PRODUCTION INDEX - OIL & GAS WELL DRILLING & MANUFACTURED HOMES	Δ log
12	IPS307.M	INDUSTRIAL PRODUCTION INDEX - RESIDENTIAL UTILITIES	Δ log
13	IPS10.M	INDUSTRIAL PRODUCTION INDEX - TOTAL INDEX	Δ log
14	UTL11.M	CAPACITY UTILIZATION - MANUFACTURING (SIC)	
15	PMI.M	PURCHASING MANAGERS' INDEX (SA)	
16	PMP.M	NAPM PRODUCTION INDEX (PERCENT)	
17	PI001.M	PERSONAL INCOME, BIL$, SAAR	Δ log
18	A0M051.M	PERS INCOME LESS TRSF PMT (AR BIL. CHAIN 2009 $), SA-US	Δ log
		EMPLOYMENT AND HOURS	
19	LHEM.M	CIVILIAN LABOR FORCE: EMPLOYED, TOTAL (THOUS.,SA)	Δ log
20	LHNAG.M	CIVILIAN LABOR FORCE: EMPLOYED, NONAGRIC.INDUSTRIES (THOUS.,SA)	Δ log
21	LHUR.M	UNEMPLOYMENT RATE: ALL WORKERS, 16 YEARS and OVER (%,SA)	
22	LHU680.M	UNEMPLOY.BY DURATION: AVERAGE(MEAN)DURATION IN WEEKS (SA)	
23	LHU5.M	UNEMPLOY.BY DURATION: PERSONS UNEMPL.LESS THAN 5 WKS (THOUS.,SA)	
24	LHU14.M	UNEMPLOY.BY DURATION: PERSONS UNEMPL.5 TO 14 WKS (THOUS.,SA)	
25	LHU15.M	UNEMPLOY.BY DURATION: PERSONS UNEMPL.15 WKS + (THOUS.,SA)	
26	LHU26.M	UNEMPLOY.BY DURATION: PERSONS UNEMPL.15 TO 26 WKS (THOUS.,SA)	
27	CES001.M	EMPLOYEES, NONFARM - TOTAL NONFARM	Δ log
28	CES002.M	EMPLOYEES, NONFARM - TOTAL PRIVATE	Δ log
29	CES003.M	EMPLOYEES, NONFARM - GOODS-PRODUCING	Δ log
30	CES006.M	EMPLOYEES, NONFARM - MINING	Δ log
31	CES011.M	EMPLOYEES, NONFARM - CONSTRUCTION	Δ log
32	CES015.M	EMPLOYEES, NONFARM - MFG	Δ log
33	CES017.M	EMPLOYEES, NONFARM - DURABLE GOODS	Δ log
34	CES033.M	EMPLOYEES, NONFARM - NONDURABLE GOODS	Δ log
35	CES046.M	EMPLOYEES, NONFARM - SERVICE-PROVIDING	Δ log
36	CES048.M	EMPLOYEES, NONFARM - TRADE, TRANSPORT, UTILITIES	Δ log
37	CES049.M	EMPLOYEES, NONFARM - WHOLESALE TRADE	Δ log
38	CES053.M	EMPLOYEES, NONFARM - RETAIL TRADE	Δ log
39	CES140.M	EMPLOYEES, NONFARM - GOVERNMENT	Δ log
40	CES154.M	AVG WKLY HOURS, PROD WRKRS, NONFARM - MFG	Δ log
41	CES155.M	AVG WKLY OVERTIME HOURS, PROD WRKRS, NONFARM - MFG	Δ log
42	PMEMP.M	NAPM EMPLOYMENT INDEX (PERCENT)	

Table 15: List of Variables Included in Calculation of Skewness Index (cont.)

	Mneumonic	Description	Transform
		CONSUMPTION	
43	PI031.M	PERSONAL CONSUMPTION EXPENDITURES, BIL$, SAAR	Δ log
44	PI032.M	PERSONAL CONSUMPTION EXPENDITURES - DURABLE GOODS, BIL$, SAAR	Δ log
45	PI033.M	PERSONAL CONSUMPTION EXPENDITURES - NONDURABLE GOODS, BIL$, SAAR	Δ log
46	PI034.M	PERSONAL CONSUMPTION EXPENDITURES - SERVICES, BIL$, SAAR	Δ log
		HOUSING STARTS AND SALES	
47	HSFR.M	HOUSING STARTS:TOTAL FARM&NONFARM (THOUS.,SA)	log
48	HSNE.M	HOUSING STARTS:NORTHEAST (THOUS.U.)S.A.	log
49	HSMW.M	HOUSING STARTS:MIDWEST(THOUS.U.)S.A.	log
50	HSSOU.M	HOUSING STARTS:SOUTH (THOUS.U.)S.A.	log
51	HSWST.M	HOUSING STARTS:WEST (THOUS.U.)S.A.	log
52	HS6BR.M	HOUSING AUTHORIZED: TOTAL NEW PRIV HOUSING UNITS (THOUS.,NSA)	log
53	HMOB.M	MOBILE HOMES: MANUFACTURERS' SHIPMENTS (THOUS.OF UNITS,SAAR)	log
		REAL INVENTORIES, ORDERS, AND UNFILLED ORDERS	
54	PMNV.M	NAPM INVENTORIES INDEX (PERCENT)	
55	PMNO.M	NAPM NEW ORDERS INDEX (PERCENT)	
56	PMDEL.M	NAPM VENDOR DELIVERIES INDEX (PERCENT)	
57	MOCMQ.M	NEW ORDERS (NET) - CONSUMER GOODS and MATERIALS, 1996 $ (BCI)	Δ log
58	MSONDQ.M	NEW ORDERS, NONDEFENSE CAPITAL GOODS, IN 1996 $ (BCI)	Δ log
		STOCK RETURNS AND PREDICTABILITY STATE VARIABLES	
59	FSPCOM.M	S&P'S COMMON STOCK PRICE INDEX: COMPOSITE (1941-43=10)	Δ log
60	FSPIN.M	S&P'S COMMON STOCK PRICE INDEX: INDUSTRIALS (1941-43=10)	Δ log
61	DY (GW)	DIVIDEND YIELD	
62	EP (GW)	EARNINGS-PRICE RATIO	
63	BM (GW)	BOOK-TO-MARKET RATIO	
64	SVAR (GW)	STOCK MARKET REALIZED VARIANCE	
65	TBL (GW)	3 MONTH T-BILL RATE	
66	TMS (GW)	TERM SPREAD	
67	DFY (GW)	DEFAULT YIELD: BAA - AAA SPREAD	
68	LTY (GW)	LONG-TERM YIELD	
69	NTIS (GW)	NET ISSUANCE	
70	INFL (GW)	INFLATION	
71	DFR (GW)	CORPORATE - GOVT BOND RETURN	
72	LTR (GW)	LONG-TERM BOND RETURN	
		EXCHANGE RATES	
73	EXRUK.M	FOREIGN EXCHANGE RATE: UNITED KINGDOM (CENTS PER POUND)	Δ log
74	EXRCAN.M	FOREIGN EXCHANGE RATE: CANADA (CANADIAN $ PER U.S.$)	Δ log

Table 15: List of Variables Included in Calculation of Skewness Index (cont.)

	Mneumonic	Description	Transform
		INTEREST RATES	
75	FYFF.M	NTEREST RATE: FEDERAL FUNDS (EFFECTIVE) (% PER ANNUM,NSA)	
76	FYGM3.M	INTEREST RATE: U.S.TREASURY BILLS,SEC MKT,3-MO.(% PER ANN,NSA)	
77	FYGM6.M	INTEREST RATE: U.S.TREASURY BILLS,SEC MKT,6-MO.(% PER ANN,NSA)	
78	FYGT1.M	INTEREST RATE: U.S.TREASURY CONST MATURITIES,1-YR.(% PER ANN,NSA)	
79	FYGT5.M	INTEREST RATE: U.S.TREASURY CONST MATURITIES,5-YR.(% PER ANN,NSA)	
80	FYGT10.M	INTEREST RATE: U.S.TREASURY CONST MATURITIES,10-YR.(% PER ANN,NSA)	
81	FYGM3.M-FYFF.M	SPREAD: FYGM3.M-FYFF.M	
82	FYGM6.M-FYFF.M	SPREAD: FYGM6.M-FYFF.M	
83	FYGT1.M-FYFF.M	SPREAD: FYGT1.M-FYFF.M	
84	FYGT5.M-FYFF.M	SPREAD: FYGT5.M-FYFF.M	
85	FYGT10.M-FYFF.M	SPREAD: FYGT10.M-FYFF.M	
		MONEY AND CREDIT QUANTITY AGGREGATES	
86	ALCIBL00.M	COML&IND LOANS OUTST IN 2009 $,SA-U	Δ log
87	CCINRV.M	CONSUMER CREDIT OUTSTANDING - NONREVOLVING(G19)	Δ log
88	FM1.M	MONEY STOCK: M1(CURR,TRAV.CKS,DEM DEP,OTHER CK'ABLE DEP)(BIL$,SA)	Δ log
89	FM2.M	MONEY STOCK:M2(M1+O'NITE RPS,EURO$,G/P&B/D MMMFS&SAV&SM TIME DEP(BIL$,SA)	Δ log
90	MBASE.M	REVISED MONETARY BASE-ADJUSTED-(FED RESERVE BANK-SAINT LOUIS),SA-US	Δ log
91	MNY2.M	M2 - MONEY SUPPLY - M1 + SAVINGS DEPOSITS, SMALL TIME DEPOSITS, & MMMFS [H6],SA-US	Δ log
		PRICE INDICES	
92	PMCP.M	APM COMMODITY PRICES INDEX (PERCENT)	
93	PWFSA.M	PRODUCER PRICE INDEX: FINISHED GOODS (82=100,SA)	Δ log
94	PWFCSA.M	PRODUCER PRICE INDEX:FINISHED CONSUMER GOODS (82=100,SA)	Δ log
95	PWIMSA.M	PRODUCER PRICE INDEX:INTERMED MAT.SUPPLIES & COMPONENTS(82=100,SA)	Δ log
96	PWCMSA.M	PRODUCER PRICE INDEX:CRUDE MATERIALS (82=100,SA)	Δ log
97	PUNEW.M	CPI-U: ALL ITEMS (82-84=100,SA)	Δ log
98	PU83.M	CPI-U: APPAREL & UPKEEP (82-84=100,SA)	Δ log
99	PU84.M	CPI-U: TRANSPORTATION (82-84=100,SA)	Δ log
100	PU85.M	CPI-U: MEDICAL CARE (82-84=100,SA)	Δ log
101	PUC.M	CPI-U: COMMODITIES (82-84=100,SA)	Δ log
102	PUCD.M	CPI-U: DURABLES (82-84=100,SA)	Δ log
103	PUS.M	CPI-U: SERVICES (82-84=100,SA)	Δ log
104	PUXF.M	CPI-U: ALL ITEMS LESS FOOD (82-84=100,SA)	Δ log
105	PUXHS.M	CPI-U: ALL ITEMS LESS SHELTER (82-84=100,SA)	Δ log
106	PUXM.M	CPI-U: ALL ITEMS LESS MIDICAL CARE (82-84=100,SA)	Δ log
		AVERAGE HOURLY EARNINGS	
107	CES277.M	AVG HRLY EARNINGS, PROD WRKRS, NONFARM - CONSTRUCTION	Δ log
108	CES278.M	AVG HRLY EARNINGS, PROD WRKRS, NONFARM - MFG	Δ log
		CONSUMER EXPECTATIONS	
109	U0M083.M	BUSINESS CYCLE INDICATORS,CONSUMER EXPECTATIONS,NS	

Table 15: List of Variables Included in Calculation of Skewness Index (cont.)

B Theoretical Model Appendix

B.1 The Term Structure of Risk Premia

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

- Dividend strips / "Zero-coupon" equity: $D_{k,t+h}$, a single dividend payment from one of the risky assets,
- Non-defaultable bonds: a real or nominal risk-free payment at time $t+h$, or
- Consumption strips: C^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 (π_t) equals $S'_\pi y_{t+1}$, 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 $-S'_\pi y_{t+1}$, the prices of its "dividends" are proportional to nominal bond prices.

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

$$R_{k,t+1} = \frac{P_{k,t+1} + D_{k,t+1}}{P_{k,t}} = \sum_{h=0}^{\infty} \frac{P^h_{k,t+1}}{P_{k,t}} = \sum_{h=0}^{\infty} \frac{P^{h+1}_{k,t}}{P_{k,t}} \cdot \frac{P^h_{k,t+1}}{P^{h+1}_{k,t}} = \sum_{h=1}^{\infty} \frac{P^h_{k,t}}{P_{k,t}} \cdot R^h_{k,t+1}, \tag{44}$$

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

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

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

(i) $\log(P_{c,t}^h/C_t) \equiv wc_t^h = A_0^h + y_t' A^h$,

(ii) $\log(P_{k,t}^h/D_{k,t}) \equiv pd_{k,t}^h = A_{0,k}^h + y_t' A_k^h$, *for* $k = 1, \ldots, K$.

for all t *and* $h \geq 0$, *where* $A_0^h, A_{0,1}^h, \ldots, A_{0,K}^h$ *are scalars and* $A^h, A_1^h, \ldots, A_K^h \in \mathbb{R}^K$.

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.

B.2 Proofs of Propositions

B.2.1 Proof of Proposition 1 (Aggregation Restrictions)

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

$$\begin{aligned}
E[\exp(\eta_{t+1}^i)|y_{t+1}, y_t] &= \exp(1_M' \mu_\eta + 1_M' F_\eta y_t + 1/2(h_{\eta 0} + h_{\eta 1}' y_t)) E[\exp(1_M' J_{\eta,t+1}^i)|y_{t+1}, y_t] \\
&= \exp[1_M' \mu_\eta + 1_M' F_\eta y_t + 1/2(h_{\eta 0} + h_{\eta 1}' y_t)] \\
&\times \exp(1_M' (l_{\eta 0}(\Psi_\eta(1_M) - 1_M) + l_{\eta 1} \odot [(\Psi_\eta(1_M) - 1_M) \otimes 1_L'] y_t)),
\end{aligned}$$

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 μ_η and F_η yields zero, so the restriction holds.

B.2.2 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, η^i_{t+1} and y_{t+1}, respectively.

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

$$E[\exp(u \cdot \eta^i_{t+1})|y_{t+1}] = \exp[\beta_0(u) + \beta'(u)y_{t+1}], \tag{45}$$

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

(i) $\beta_0(u) = \mu'_\eta 1_M u + \frac{1}{2}u^2 h_{\eta 0} + l'_{\eta 0}(\Psi_\eta(u1_M) - 1_M)$ *and*

(ii) $\beta(u) = F'_\eta 1_M u + \frac{1}{2}u^2 H_{\eta 1} + l'_{\eta 1}(\Psi_\eta(u1_M) - 1_M)$.

Proof. By definition, $\eta^i_{t+1} = 1_M \tilde{\eta}^i_{t+1}$. By the conditional independence of $z^i_{\eta,t+1}$ and $J^i_{\eta,t+1}$,

$$\begin{aligned}
\log E[\exp(u \cdot \eta^i_{t+1})|y_{t+1}] &= u1'_M[\mu_\eta + F_\eta y_{t+1}] + \log E[\exp(u1'_M G_{\eta,t+1} z^i_{\eta,t+1})|y_{t+1}] \\
&+ \log E[\exp(u1'_M J^i_{\eta,t+1}])|y_{t+1}] \\
\log E[\exp(u1'_M G_{\eta,t+1} z^i_{\eta,t+1})|y_{t+1}] &= \tfrac{1}{2}u^2 1'_M G_{\eta,t+1} G'_{\eta,t+1} 1_M = \tfrac{1}{2}u^2[h_{\eta,0} + H'_{\eta 1} y_{t+1}] \\
\log E[\exp(u1'_M J^i_{\eta,t+1}])|y_{t+1}] &= \lambda'_{\eta,t+1}[\Psi_\eta(u1_M) - 1_M] = [l_{\eta 0} + l_{\eta 1} y_{t+1}]'[\Psi_\eta(u1_M) - 1_M],
\end{aligned}$$

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,*

$$E_t[\exp(u' y_{t+1})] = \exp[f(u) + g(u)' y_t], \tag{46}$$

where $f(u)\colon \mathbb{R}^L \to \mathbb{R}$ *and* $g(u)\colon \mathbb{R}^L \to \mathbb{R}^L$ *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

$$\begin{aligned}
\log E_t[\exp(u'y_{t+1})] &= u'[\mu_y + F_y y_t] + \log E_t[\exp(u'G_{y,t+1}z_{y,t+1})] + \log E_t[\exp(u'J_{y,t+1}])] \\
\log E_t[\exp(u'G_{y,t}z_{y,t+1})] &= \tfrac{1}{2}u'G_{y,t}G'_{y,t}u = \tfrac{1}{2}u'h_y u + \tfrac{1}{2}\sum_{j=1}^{L} u'H_{y,j}u \cdot y_{j,t} \\
\log E_t[\exp(u'J_{y,t+1}])] &= \lambda'_{y,t}[\Psi_y(u) - 1_L] = [l_{y0} + l_{y1}y_t]'[\Psi_\eta(u) - 1_L].
\end{aligned}$$

As before, we use moment generating functions to evaluate expectations, plug in the functional forms from Assumptions 3.i-ii, then collect terms. See also DY section A.1. □

We will assume that the wealth-consumption ratio $wc_t = A_0 + A'y_t$. By Assumptions 1-2, we can write consumption growth in vector notation as $\Delta c_t^i = S_c' y_t + \eta_t^i$. Combining (20) with our assumption, the log-linearized Euler equation for the consumption claim is

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

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

$$\begin{aligned}
\theta(\log\delta + \kappa_c + (\rho_c - 1)A_0) - \theta A'y_t &= -f((1-\gamma)S_c + \theta\rho_c A + \beta(1-\gamma)) + \beta_0(1-\gamma) \\
&\quad - g((1-\gamma)S_c + \theta\rho_c A + \beta(1-\gamma))'y_t.
\end{aligned}$$

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

$$\begin{aligned}
f((1-\gamma)S_c + \theta\rho_c A + \beta(1-\gamma)) + \beta_0(1-\gamma) &= -\theta(\log\delta + \kappa_c + (\rho_c - 1)A_0) \qquad (47) \\
g((1-\gamma)S_c + \theta\rho_c A + \beta(1-\gamma)) &= A\theta, \qquad (48)
\end{aligned}$$

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 (47-48) also depends on the log-linearization constants κ_c and ρ_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 $\overline{wc}$ so that

$$\log(\rho_c) - \log(1-\rho_c) = \overline{wc} = E(wc_t) = A_0 + A'E(y_t), \tag{49}$$

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

$$\kappa_c + (\rho_c - 1)A_0 = -\log\rho_c - (\rho_c - 1)A'E(y_t). \tag{50}$$

We can then substitute (50) into (47), yielding

$$f((1-\gamma)S_c + \theta\rho_c A + \beta(1-\gamma)) + \beta_0(1-\gamma) = -\theta(\log\delta - \log\rho_c - (\rho_c-1)A'E(y_t)), \tag{51}$$

leaving (48) and (51), an exactly identified system of equations in A and ρ_c. Then, given these solutions, we can use the expressions above to derive A_0, κ_c, and κ.

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 (22). Plugging in the projected kernel, the log-linearized Euler equation for the k^{th} dividend claim is

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

As before, using Lemma 2 to evaluate the expectation and taking logs yields the $(L+1)$-dimensional system of equations

$$f(-\Lambda + S_k + \rho_k A_k) = -[\kappa - (1-\theta)(\rho_c - 1)A_0 + \kappa_k + (\rho_k - 1)A_{0,k}] \tag{52}$$

$$g(-\Lambda + S_k + \rho_k A_k) = A_k - (1-\theta)A. \tag{53}$$

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 κ_k and ρ_k to differ across assets. This amounts to replacing equation (52) with

$$f(-\Lambda + S_k + \rho_k A_k) = -[\kappa - (1-\theta)(\rho_c - 1)A_0 - \log\rho_k - (\rho_k - 1)A_k'E(y_t)]. \tag{54}$$

B.2.3 Proof of Proposition 4 (ICAPM)

In this section, I derive an ICAPM representation for my incomplete markets economy. Many of the steps of the derivation follow Campbell et al. (2012), so I highlight the incremental effects of adding incomplete markets. From Proposition 2, the log of the average return on the consumption claim is an affine function of the state vector, y_t. Following Campbell (1993), we substitute out Δc_{t+1} using the identity

$$\Delta c_{t+1} \approx r_{c,t+1} - \kappa_c - \rho_c wc_{t+1} + wc_t. \tag{55}$$

Plugging (55) into the log-linearized pricing kernel (20), one obtains

$$m_{i,t+1} \quad = \quad \text{const.} - \gamma(r_{c,t+1} + \eta_{i,t+1}) + \tfrac{\theta}{\psi}(\rho_c wc_{t+1} - wc_t). \tag{56}$$

Plugging (56) into the Euler equation for $r^i_{c,t+1}$ and projecting out $\eta_{i,t+1}$ yields

$$1 \quad = \quad E_t\left[\exp\left(\text{const.} + (1-\gamma)(r_{c,t+1} + \nu^*_{t+1}) + \tfrac{\theta}{\psi}(\rho_c wc_{t+1} - wc_t)\right)\right], \tag{57}$$

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

$$wc_t = \text{const.} + (\psi - 1)[E_t r_{c,t+1} + E_t\nu^*_{t+1}] + \rho_c E_t wc_{t+1} + \tfrac{1}{2}\tfrac{\psi}{\theta}\vartheta_t, \tag{58}$$

where ϑ_t is a Jensen's inequality term. In the absence of jump risk, this term equals $Var_t[m^i_{t+1} + r^i_{c,t+1}]$, i.e. the risk-neutral variance of the consumption claim. When jumps are present, there is an analogous term capturing Gaussian volatility and jump risk.

Iterating forward on (58) and assuming that $\lim_{j\to\infty}\rho^j wc_{t+1} = 0$, one obtains

$$\begin{aligned}
wc_t \quad &= \quad \text{const.} + E_t\sum_{j=0}^{\infty}\rho_c^j[(\psi-1)(r_{c,t+1+j} + \nu^*_{t+1+j}) + \tfrac{1}{2}\tfrac{\psi}{\theta}\vartheta_{t+j}] && (59)\\
\rho_c[wc_{t+1} - E_t wc_{t+1}] \quad &= \quad [E_{t+1} - E_t]\sum_{j=1}^{\infty}\rho_c^j[(\psi-1)(r_{t+1+j} + \nu^*_{t+1+j}) + \tfrac{1}{2}\tfrac{\psi}{\theta}\vartheta_{t+1+j}] && (60)\\
&\equiv \quad (\psi-1)[N_{DR,t+1} + N_{FIR,t+1}] + \tfrac{1}{2}\tfrac{\psi}{\theta}N_{UNC,t+1}, && (61)
\end{aligned}$$

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

$$r_{c,t+1} - E_t r_{c,t+1} = [E_{t+1} - E_t]\sum_{j=0}^{\infty}\rho_c^j[\Delta c_{t+1+j} - \rho\cdot r_{c,t+2+j}] \equiv N_{CF,t+1} - N_{DR,t+1}. \tag{62}$$

The key difference with respect to the representative agent model is the second term, $N_{FIR,t+1}$. The subscript FIR is shorthand for *future idiosyncratic risk*, which captures news about the higher moments of idiosyncratic shocks.[82] Equation (61) 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 ν^*_{t+1}. The third term $N_{UNC,t+1}$ captures news about uncertainty, i.e. the higher moments of future aggregate shocks. Plugging (61-62) into the projected pricing kernel (21), (26) obtains.

[Still need to add moment condition for ϑ_t here.]

B.2.4 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

$$\begin{aligned} m_{t+1} &= \underbrace{\kappa - (1-\theta)(\rho_c - 1)A_0 + (1-\theta)A'y_t - \Lambda'(\mu_y + F_y y_t + G_{y,t}z_{y,t+1})}_{\equiv m^g_{t+1}} + \underbrace{-\Lambda' J_{y,t+1}}_{\equiv m^J_{t+1}} \\ r_{k,t+1} &= \underbrace{\kappa_k + (\rho_k - 1)A_{0k} - A'_k y_t + B'_k(\mu_y + F_y y_t + G_{y,t}z_{y,t+1})}_{\equiv r^g_{k,t+1}} + \underbrace{B'_k J_{y,t+1}}_{\equiv r^J_{k,t+1}} \end{aligned}$$

DY show that the risk premium may be decomposed as

$$\begin{aligned} \log(E_t[R_{k,t+1}]) - rf_{t+1} &= -\mathrm{cov}_t(m^g_{t+1}, r^g_{k,t+1}) + \log(E_t[\exp(r^J_{k,t+1})]) \\ &+ \log(E_t[\exp(m^J_{t+1})]) - \log(E_t[\exp(r^J_{k,t+1} + m^J_{t+1})]) \\ &= B'_k G_{y,t} G'_{y,t}\Lambda + \lambda'_{y,t}[\Psi_y(B_k) - 1_L] - \lambda'_{y,t}[\Psi_y(B_k - \Lambda) - \Psi_y(-\Lambda)]. \end{aligned}$$

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

B.2.5 Proof of Proposition 5 (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 $pd^0_{k,t} = A^0_{0,k} + (A^0_k)'y_t$. By our no

[82] 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.

arbitrage restriction that $D_{k,t} = P^0_{k,t}$, implying that $pd^0_{k,t} = 0$ and thus that $A^0_{0,k} = 0$ and $A^0_k = 0$. Next, 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

$$\begin{aligned} pd^{h+1}_{k,t} &= \log E_t[\exp\{m_{t+1} + \Delta d_{k,t+1} + pd^h_{k,t+1}\}] \\ A^{h+1}_{0,k} + (A^{h+1}_k)' y_t &= \log E_t[\exp\{m_{t+1} + A^h_{0,k} + (S_k + A^h_k)' y_{t+1}\}] \\ &= \log E_t[\exp\{m_0 + (1-\theta)A' y_t + A^h_{0,k} + (S_k + A^h_k - \Lambda)' y_{t+1}\}] \\ &= m_0 + A^h_{0,k} + f(S_k + A^h_k - \Lambda) + [g(S_k + A^h_k - \Lambda) + (1-\theta)A]' y_t\}] \end{aligned}$$

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

$$\begin{aligned} A^{h+1}_{0,k} &= m_0 + A^h_{0,k} + f(S_k + A^h_k - \Lambda) \qquad (63) \\ A^{h+1}_k &= g(S_k + A^h_k - \Lambda) + (1-\theta)A, \qquad (64) \end{aligned}$$

which establishes the claim. One obtains coefficients for the real risk-free asset, by setting $S_k = 0$ in (63-64). 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 η^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 $wc^{h+1}_t = A^{h+1}_0 + (A^{h+1})' y_t$:

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

where $\tilde{m_0} \equiv \theta \log\delta - (1-\theta)(\kappa_c + (\rho_c - 1)A_0) + \beta_0(1-\gamma)$ and $\tilde{\Lambda} \equiv \gamma S_c + (1-\theta)\rho_c A + \beta(1-\gamma)$. The third equality uses the law of iterated expectations and Lemma 1, and the fourth equality follows from Lemma 2. Matching coefficients yields the recursion

$$
\begin{aligned}
A_0^{h+1} &= \tilde{m}_0 + A_{0,k}^h + f(S_k + A_k^h - \tilde{\Lambda}) && (65)\\
A^{h+1} &= g(S_k + A_k^h - \tilde{\Lambda}) + (1-\theta)A, && (66)
\end{aligned}
$$

so the only substantive difference between (63-64) and (65-66) comes from the definitions of $\tilde{m}_0$ and $\tilde{\Lambda}$. Further note that $\tilde{m}_0 = m_0 - \beta_0(-\gamma) + \beta_0(1-\gamma)$ and $\tilde{\Lambda} = \Lambda - \beta(-\gamma) + \beta(1-\gamma)$, so the difference between the recursions comes entirely from the projection terms.

B.2.6 Risk-free rate

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

$$
\begin{aligned}
rf_{t+1} &= -\log\left(E_t[\exp(m_{t+1})]\right)\\
&= -\left[\kappa - (1-\theta)(\rho_c - 1)A_0 + (1-\theta)A'y_t + \log\left(E_t[\exp(-\Lambda y_{t+1})]\right)\right]\\
&= rf_0 - [g(-\Lambda) + (1-\theta)A']y_t,
\end{aligned}
$$

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