NORTHWESTERN UNIVERSITY

THREE ESSAYS IN MACROECONOMICS

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

For the degree

DOCTOR OF PHILOSOPHY

Field of Economics

By

Joshua Mark Davis

EVANSTON, ILLINOIS

JUNE, 2008

ABSTRACT

Three Essays in Macroeconomics

Joshua Mark Davis

Chapter one investigates the impact of agents' expectations about future fundamental economic disturbances (news) on macroeconomic dynamics. Several intuitive tests provide insight into the information content of the yield curve and its' ability to identify these 'news' disturbances. Bayesian estimation of a dynamic stochastic general equilibrium (DSGE) model using conventional macroeconomic aggregates and term structure data suggests that news shocks are important for understanding economic fluctuations.

Chapter 2 presents a 'hybrid' model of the yield curve that systematically incorporates the cross-equation restrictions of a structural dynamic stochastic general equilibrium model into an affine macro-factor model of interest rates. News , factors identified by interest rates that help to forecast future macroeconomic aggregates, are introduced into the modeling framework. Bayesian model comparison and classical likelihood ratio tests confirm the presence of news in the yield curve. Variance decompositions reveal that news shocks are responsible for a considerable amount of variation in yield curve factors.

The Bayesian methodology provides a natural identification scheme for the various fundamental economic shocks and reveals the dimensions on which the structural model is misspecified. Interestingly estimated risk premia are found to vary much less over time when news shocks are included in the estimation.

Using `business cycle accounting' (BCA), Chari, Kehoe and McGrattan (2006) (CKM) conclude that models of financial frictions which create a wedge in the intertemporal Euler equation are not promising avenues for modeling business cycle dynamics. There are two reasons that this conclusion is not warranted. First, small changes in the implementation of BCA overturn CKM's conclusions. Second, one way that shocks to the intertemporal wedge impact on the economy is by their spillover effects onto other wedges. This potentially important mechanism for the transmission of intertemporal wedge shocks is not identified under BCA. CKM potentially understate the importance of these shocks by adopting the extreme position that spillover effects are zero.

Acknowledgements

A smooth sea never made a skilled sailor

English Proverb

I would like to express my deep and sincere gratitude to my father, Mark D. Davis. He has been, and continues to be, a great role model whom I hold in the highest regard. This thesis is largely a product of his support and dedication.

I wish to express my warm and sincere thanks to my advisor and friend, Lawrence J. Christiano, Ph.D., the Alfred W. Chase Chair of Business institutions at Northwestern University. His vast knowledge, patience and undying curiosity provided much of the catalyst that brought these ideas together. Teachers like him are rare and cherished.

I also wish to express thanks to Martin Eichenbaum, Arvind Krishnamurthy and Giorgio Primiceri for encouraging me to pursue the ideas contained herein

Table of Contents

List of Figures

CHAPTER 1

News and the Term Structure in General Equilibrium

The effects and quantitative significance of news shocks has been the subject of a recent strand of literature whose roots lie in the classic work of Arthur Pigou (1926). There it is proposed that business cycle fluctuations are the result of agents' inability to properly forecast future economic activity. Agents dynamically respond to expected future disturbances and the failure of these shocks to materialize can lead to fluctuations similar to those characterizing modern economies. Economic frictions that result in sluggish transitions can magnify the effects of these shocks[1]. The intertemporal adjustment process is likely to complicate the identification of news shocks which are thought to arise on the supply-side but resemble demand shocks via a wealth effect channel[2]. Allocation of resources in response to these types of disturbances, even when ex-ante efficient, may have severe welfare repercussions if such shocks fail to materialize. History has produced countless examples of individuals' whose expectations are met with dissappointment, as seen in the recent era of 'irrational exuberance[3].'

Most of the economic research to date examines various successes and shortcoming of existing structural models in the presence of news shocks. This has led some researchers

[1]This is a point made by Christiano, Ilut, Motto and Rostagno (2006)

[2]Beaudry and Portier (2006) present a nice discussion of this point. It may be difficult to distinguish between the economic response to current/past fundemental economic shocks and expected future shocks due to the sluggish nature of aggregate dynamics.

[3]A micro-founded literature has emerged to explain some of this behavior. Examples include theories of rational herding and information cascades. See, for instance, Banerjee(1992), Caplin and Leahy (1993) or Zeira (1994).

to propose new frameworks whose properties are consistent with the presence of these disturbances. Interestingly, little has been done to document the quantitative significance of news shocks in business cycles. Beaudry and Portier (2006) quantify the significance of one type of news shock in a VAR framework. It is reasonable to believe that there are other types of news shocks that lie outside of their analysis whose economic significance can be identified by a structural economic model. This paper makes progress on this dimension by fitting such a model to data on macroeconomic and financial aggregates. Bayesian model comparison determines the types of news shocks which have the most promising implications. The main empirical finding is that news shocks are critical components of aggregate fluctuations.

In order to identify the importance of news shocks the structural economic model is fit to conventional macroeconomic variables, such as output growth and inflation, as well as data on the term structure of interest rates. Beaudry and Portier (2006) argue that stock prices contain information about these disturbances. The results presented herein contribute to this line of research by documenting the 'news' properties of the term structure. Fundamentally, this is due to the Expectations Hypothesis (EH) which, loosely stated, implies that the interest rate payed on a long-term zero coupon bond is equal to the expected return on a strategy that repeatedly rolls over short-term bonds. While there is some evidence of time variation in term premiums (see, for instance, Cochrane and Piazzesi (2006)), most agree that the EH should account for most of the variation in the long rates (see Bekaert, Cho and Moreno (2007) for a discussion).

The main reason for using interest rates to identify news shocks, as opposed to stock returns, is that they appear to be more closely related to broader economic fundamentals.

Viewing the term structure through the lense of the EH provides insight into agents' expectations about future monetary policy. The reason for this is due to the credibility of modern central banks in reacting to inflation. In the presence of nominal frictions, news about future disturbances are reflected in current and expected future prices. Agents, recognizing the central bank to be credible in its quest to quell inflation using monetary policy, immediately reveal their expectations through the path of short rates. Under the EH this path can be backed out from the yield curve, providing useful information to agents and policymakers alike.

To clarify this idea it is useful to walk through a simple fictional example[4]. Imagine an economy currently resting at steady state and news unfolds about a technology which, ceteris paribus, is expected increase aggregate productivity in the future. These agents anticipate that the technology will make them much wealthier and, consistent with consumption smoothing, expect to start spending a portion of this wealth prior to the actual realization of the impact of the technology. Thus, the agents expect demand for consumption goods to increase and prices to rise. The central bank, a reputable inflation targetting institution, is thus expected to increase interest rates in the face of this rising inflation. In response to all of these expectations yields on bonds rise. The magnitude and timing of the changes in the implied short-rates from these bond yields reveals much about the technology and its expected impact on the economy. The structural model examined here uses this sort of logic to examine the importance of news shocks for explaining aggregate fluctuations.

[4]This fictional example is further supported by a simple structural model in Section 3

This paper also contributes to the term structure literature. In this literature it is common to have latent factors drive most of the dynamics of the yield curve. Three factors in particular, commonly referred to as slope, level and curvature, drive most of the variation in interest rates. The model presented here relates the yield curve to a set of latent factors that have a structural interpretation in the context of the underlying economic model. This allows for an investigation of the significance of shocks to agents' expectations for determining the shape of the yield curve in a large scale structural model.

The remainder of the paper is organized as follows. Section 2 gives a brief overview of the related literature. Section 3 presents evidence on news shocks in the term structure of interest rates. Section 4 presents a detailed economic model. Section 5 details how the model is estimated. Section 6 derives the bond prices implied by the structural model. Section 7 presents the results and Section 8 concludes.

1.1. Related Literature

Business cycle researchers are skeptical of conventional supply and demand shocks being the driving force behind aggregate fluctuations[5]. Thus, researchers have expanded their hunt for a set of economic impulses capable of reproducing the empirical facts, among which are shocks expected arrive at future dates. There is ample evidence for the presence of news shocks in macroeconomic quantities[6]. Diffusion of new technologies can take time, the impact of which is often anticipated by agents[7]. Events such as political transition, fiscal spending and seasonality are known to play a role in the individuals'

[5]See for example Cochrane (1994)

[6]Such aggregates are given by the leading indicators. See Stock and Watson (1999) for a discussion.

[7]Consider for example railroads and computers. Rotemberg (2003) and Alexopoulos (2004) explore the impact of technology diffusion on aggregate fluctuations.

decision process[8]. Beaudry and Portier (2006) present a statistical procedure for quantifying the importance of these types of disturbances. Using stock prices and measures of productivity they argue that up to 50% of business cycle fluctuations are caused by shocks best interpreted as news about future production opportunities. This paper claims that interest rates contain the same type of information about such disturbances and perhaps have better signal quality due to their close relationship with economic fundamentals.

Only recently have researchers begun to explore the implications of news shocks in structural economic models. In response to the lackluster evidence for technological regress[9], Beaudry and Portier (2004) claim that Pigou's theory of the business cycle may be the mechanism underlying economic recessions. These authors show that a DSGE model with news shocks can match the observed frequency and depth of downturns which characterize developed economies. They also conjecture that the Asian financial crisis of the late 1990's could have been trigerred by revisions to long term expectations. Jaimovich and Rebelo (2006) expand on these results by presenting a framework for understanding these behavioral effects in a real business cycle model modified to include several real frictions. In their analysis, overconfidence increases business cycle volatility which can have significant welfare-reducing effects. Christiano, Ilut, Motto and Rostagno (2006) offer the theory as an explanation for boom/bust cycles which are magnified by an inflation targeting monetary policy. Their results highlight the potential consequences that can arise from policy that ignore the presence of news shocks. Lastly, Engel and

[8]Bussie and Mulder (2000) document the effects of political instability on resource allocation

[9]See Kydland and Prescott (1982) and King, Plosser and Rebelo(1987) for models where economic downturns are related to technological regress.

Rogers (2006) argue that the US current account deficit may be a rational outcome if expectations about future US growth are sufficiently greater than the rest of the world. This paper contributes to this macroeconometric literature by documenting the quantitative significance of news shocks in a large scale DSGE model. I find that these disturbances are extremely important components of aggregate fluctuations and should be a primary focus of future economic research.

A sizeable literature on the term structure has identified macro-factors as important for understanding the relationship between interest rates of various maturities[10]. The yield spread, the difference between the yield on long-term and short-term bonds, is widely considered to be a leading economic indicator, forecasting 5 of the last 6 economic recessions. Evidence suggests that interest rates contain information about future activity that may be useful for policymakers. The significance of yield curve information in predicting other economic aggregates is pursued by Estrella and Mishkin (1998) and Estrella and Hardouvelis (1991). These authors conclude that the yield spread can have a useful role in macroeconomic prediction, particularly in the case of longer lead times. Further evidence for the information content of the term structure can be seen in statistical representations of the yield curve, where latent variables are found to be necessary for fitting observed yield curve shapes[11]. Evans and Marshall (1998), in a VAR framework, determine that shocks to preferences for consumption induce large, persistent and statistically significant shifts in the yield curve. Beaudry and Portier (2004) argue that news shocks and preference shocks have similar implications for the contemporaneous responses

[10]Prominent examples include Ang and Piazzesi (2003), Arouba, Diebold and Rudebusch (2004), Bikbov and Chernov (2005) and Evans and Marshall (1998)

[11]One interpretation of the results of Ang and Piazzesi (2003) is that the information content of macroeconomic variables is different from that of the yield curve.

of macroeconomic aggregates. This perspective suggests that the results of Evans and Marshall may be hidden evidence for the significance of news shocks in economic dynamics. This paper fits into the term structure literature by providing a new perspective on the types of disturbances that determine equilibrium interest rates. In doing so, I provide a decomposition of the movemements in conventional yield curve factors due to various fundamental economic disturbances and show that news shocks are a dominant source.

The approach used in this paper most closely relates to that of Bekaert, Cho and Moreno(2006). These authors construct a New-Keynesian model and fit it to data on the macroeconomic aggregates and the term structure under the EH. Their approach differs in many respects from the one presented here, namely that they use a small-scale reduced form model, include a limited amount of macroeconomic information when fitting the model and don't incorporate news shocks into their framework. They find that the inclusion of the yield curve data in the estimation helps to identify large and significant estimates of the Phillips curve and interest rate response parameters. They also examine the effects of different economic shocks on the yield curve factors.

1.2. News Shocks and the Term Structure

The goal of this section is to document the informational properties of the term structure of interest rates and its ability to identify the economic significance of news shocks. Evidence suggests that information contained in macroeconomic quantities cannot span the various yield curve shapes and time-dynamics of interest rates. In response the literature has determined that a set of 'latent' factors, backed out from the cross-section of bond yields, provide the degrees of freedom necessary for explaining the properties of interest rates. The importance of these factors for modelling the yield curve is evidence that interest rates contain information that is either: (1) not encompassed by that contained in macroeconomic quantities or (2) difficult to extract. The material presented in this section supports, both theoretically and statistically, this conjecture by showing that the yield curve contains information relevant for macroeconomic prediction. Interest rates, in a very literal sense, contain news that cannot be gleaned from conventional aggregates and can thus be useful for identifying the economic significance of news shocks.

To provide motivation for the promising identification properties of the term structure of interest rates I will first present a simple economic model. Suppose that a representative agent maximizes

$$E_0 \sum_{t=0}^{\infty} \beta^t \ln c_t$$

subject to the following budget constraint and exogenous stochastic processes for technology

$$\begin{aligned} c_t + k_{t+1} &\leq z_t k_t^{\alpha} \equiv y_t \\ \ln z_{t+1} &= \rho_z \ln z_t + \zeta_{z,t+1}^{t} + \varepsilon_{z,t+1} \end{aligned}$$

where $\zeta_{z,t+1}^{t}$ is interpreted to be a 'news' shock, an innovation to the stochastic process for technology expected at time t to arrive at time $t+1$. The disturbance $\varepsilon_{z,t+1}$ is orthogonal to time t information, consistent with the traditional interpretation of a technology shock.

It is straightforward to show that the efficient allocations in the model satisfy

$$\begin{aligned} y_t &= z_t k_t^{\alpha} \\ k_{t+1} &= \alpha\beta y_t \\ c_t &= (1-\alpha\beta)\, y_t \end{aligned}$$

The econometrician who observes data on these allocations would not be able to identify news shocks[12]. This provides motivation for using the yield curve to identify news shocks. The 1-period interest rate in this model is

$$R_t = \frac{1}{\beta c_t} E_t \left[\frac{1}{c_{t+1}}\right]^{-1} = \frac{f(\alpha, \beta, \rho_z)\, z_t^{\rho_z} y_t^{\alpha-1}}{E_t\left[e^{-\xi^t_{z,t+1}-\varepsilon_{z,t+1}}\right]}$$

where f is a function of the parameters α, β and ρ_z. It is straightforward to see that the news shock appears in the interest rate, a result of the fact that interest rates carry information about agents' expectations.

Using this simple model as motivation I now document additional evidence for the information content of the term structure. The slope of the yield curve, the difference between yields on long and short maturity bonds, may be the most referenced leading indicator of the business cycle. Under the expectations hypothesis the implied interest rate on a long term bond is simply equal to the average expected rate payed by short term bonds over that same period. If one were to assume that the hypothesis held a large amount of information about future economic activity could be simply read off of the yield

[12]This model implies that news shocks are completely insignificant as well. The idea behind this example is to highlight the theoretical identification properties of interest rates, not the economic significance of the news shocks

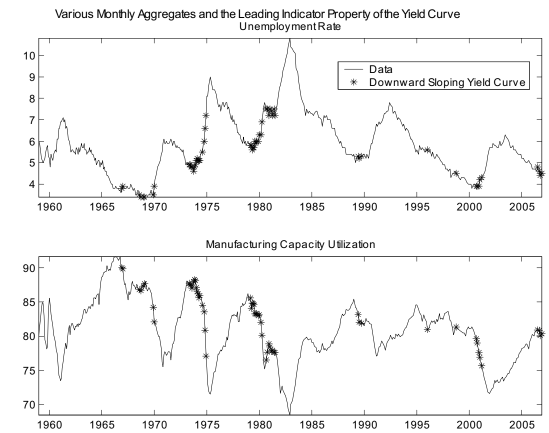


Figure 1.1. Inverted yield curve recession signals

curve. A downward sloping curve indicates that the short term interest rate will drop in the future, consistent with a loose monetary policy run during a recession. Furthermore, the term structure also contains information about the timing and magnitude of these expected interest rate movements, a signal of agents' expectations.

Figure 1 plots several monthly macroeconomic time series covering a broad portion of the economy. It is straightforward to see why the term structure is thought to contain information about future economic activity. Manufacturing capacity utilization, a measure of the intensity at which capital is run in the manufacturing sector, displays a very distinct business cycle pattern. The inverted yield curve indicator begins to occur roughly at each peak of this series and continues halfway through the decline, when capital is used less intensely since demand for goods decreases during recessionary episodes. The indicator also occurs at every trough in the unemployment rate when the economy is at the peak of an expansionary period and short term interest rates are relatively high.

Figure 1 displays leading indicator property of the slope of the yield curve for predicting future economic activity. To see this note that the yield curve slopes downward prior to economic slowdowns, ie. times when unermployement is increasing and capacity utilization decreasing. It remains to be shown that the yield curve contains information that is either difficult to extract from conventional macroeconomic quantities. To do this I will present a simple statistical model of agents' expectations based on a Bayesian Vector-Autoregressions (BVAR) and compare the informational properties of this model with those of the yield curve. BVAR's are typically used because of their superior forecasting performance (see, for instance, Kadiyala and Karlsson (1997)). The variables used in the BVAR are real per-capita GDP growth, an indicator of the current level of real activity, the log difference in the GDP deflator, an indicator of the current level of inflation, and the Federal Funds Rate, a proxy for the 3 month risk-free rate. These variables are similar to those used by Ang and Piazzesi (2003) who present a statistical factor model of the yield curve. The appendix contains a discussion of the prior as well as the estimation methodology.

I construct the yield on a zero coupon bond under the expectations hypothesis, referred to as the 'EH-BVAR Yield,' by taking the average short rate forecasted by the BVAR. The artificial yield from the BVAR does reasonably well at approximating the market yield at the 1 year horizon, lying an average of 11.89 basis points below the market rate with the standard deviation of the difference being 78.4 basis points. During the 'great moderation,' the post-1985 period , this standard deviation is cut in half at 43.56 basis points. These statistics suggest that the information used in the markets' determination of interest rates is similar to that contained in the BVAR variables at the one year horizon. Furthermore, agents' anticipation of disturbances expected to arrive within a year have identifiable contemporaneous effects on macroeconomic quantities, to the extent that the BVAR implied rate doesn't differ significantly from the market rate.

Difference between Market Yield and EH-BVAR Yield				
	1 Year Horizon		5 Year Horizon	
Time Period	Mean	Stdev	Mean	Stdev
1967 Q3 : 2005 Q4	11.77	74.66	64.72	129.21
1967 Q3 : 1984 Q4	13.38	101.74	74.75	175.62
1985 Q1 : 2005 Q4	10.40	39.63	56.15	68.49

Figure 2 shows that the 5 year horizon artificial yield is noticeably different from that in the market. Prior to the great moderation the deviations are concentrated in volatile, short lived spikes. These episodes mainly occur during economic transition and result in an overreaction of the artificial yield relative to that of the market. One explanation for these episodes, which are much less predominant at shorter horizons, is that the conditional expectations derived from the BVAR and those in the market are inconsistent. This interpretation would imply that the market uses a much broader information set than that encompassed by inflation and output growth. In particular, the forward looking nature of interest rates may be evidence of news shocks which haven't yet presented themself in the broader economic aggregates[13].

Pursuing this idea further consider the oil crisis episode around 1974. One structural interpretation of skyrocketing oil prices is that of a downward exogenous shock to productivity. The market expects this is a temporary phenomena likely to have a quick

[13]This may also be evidence of risk premia. This appears to be an unlikely explanation since risk premia are often documented to be small and positive. See Cochrane and Piazzesi (2006) for references and a discussion

political resolution, corresponding to a small movement in the long term rate. However, the BVAR, lacking this information from the yield curve, may be overestimating the persistence of the shock. In this case the BVAR should predict a persistent rise in inflation which accompanies a negative productivity shock, resulting in a sharp increase in the short term interest rate by the central bank.

There is a very muted difference between the two yields during the great moderation, most notably in the periods of 1987-1988 and 1994-1995 when the yield curve is slightly higher than the rate constructed using the expectation hypothesis. One interpretation of the similarity between the market yield and the artificial yield following the high inflation period of the early 1980's is that of credibility on the part of policymakers in targetting inflation, driving the price of risk associated with inflation downward and resulting in better fit under the EH (see Cochrane and Piazzesi (2006) for a discussion of the relationship between inflation risk and interest rates).

Another way to examine the information content of the term structure relative to the aggregates used in the BVAR is to compare the leading indicator properties of the slope of the artificial yield curve to that of the market. It has been shown above that a negative sloping market curve signals a business cycle peak and tends to forewarn of an upcoming recession. The second and third plots of figure 2 show the dates of these negative slopes using both the EH-VAR yield and the market yield as well as the dates of NBER defined recessions. The time series used in these figures is HP-filtered log per-capita real GDP with a smoothing parameter of 1600. Of 60 potential quarters where the artificial yield has a negative slope, 17 of them occur at times where a recession did not occur in the near future, a false signal percentage of 28.33%. The market yield curve slopes downward a

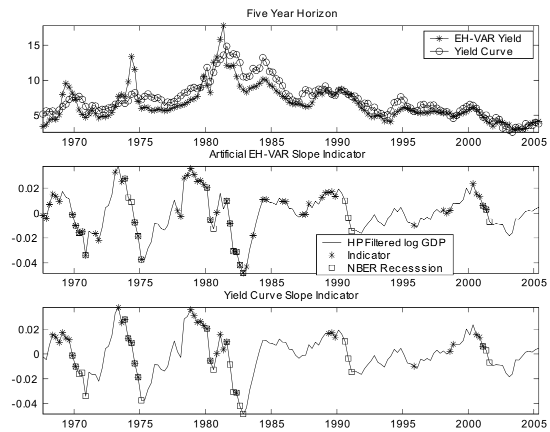


Figure 1.2. Evidence of news in the yield curve

total of 36 times, of which 3 do not appear in the vicinity of an upcoming recession. In this sample the EH-VAR artificial yield is more than 3 times as likely to produce a false signal than that implied by the market's curve. This ad hoc analysis shows that there is valuable information in the term structure of interest rates which is not contained in conventional macroeconomic quantities.

Beaudry and Portier(2006) provide evidence for the importance of news shocks in macroeconomic dynamics. Their method is based on comparing the implications for shocks identified using particular restrictions in a VAR. The approach is meant for news shocks which have no contemporaneous impact on TFP, yet, after being realized, lead to a permanent change in TFP[14]. Expected disturbances to the supply-side, such as future investment-specific or neutral technology shock, that are represented by a unit root stochastic process are consistent with such a framework. As discussed above, news shocks are initially captured in forward looking variables and not TFP, motivating BP to

[14]Shocks not having permanent long-run impact may be explored in the context of medium horizon restrictions. See Uhlig (2004) for a discussion of this type of identification.

include stock prices in their VAR analysis. As a robustness check on the evidence for the information content of the yield curve presented above, I modify BP's bivariate VAR to include term structure data. BP's results should remain intact if bond and stock prices contain similar macroeconomic information. Furthermore, the statistical implications of the exercise serve as a benchmark for structural representations to match.

Now I quickly describe the identification scheme proposed by BP. Denote $\Delta(y_t/h_t)$ as the change in labor productivity and p_t^b as the log of the 5 year bond price. A VAR can be written in lag operator form as

$$(I - B(L))\begin{bmatrix}\Delta(y_t/h_t) \\ p_t^b\end{bmatrix} = \begin{bmatrix}u_{1,t} \\ u_{2,t}\end{bmatrix}$$

The VAR residuals, u, are linear combinations of the fundamental economic disturbances; ε. The problem of the econometrician is to recover the $\varepsilon' s$, the variables of economic interest, from the VAR residuals. This amounts to finding a matrix Γ which can be used to untangle the VAR residuals. Estimation of the VAR identifies a whole family of $\Gamma' s$ indirectly by identifying the variance covariance matrix $V = \Gamma\Gamma'$. Isolating a partifular Γ requires additional restrictions. These additional restrictions are provided by economic assumptions about the affects of news shocks at different time horizons. BP assume that news shocks have a permanent long-run impact on labor productivity. This assumption identifies a particular matrix $\widetilde{\Gamma}$ from which the fundamental economic disturbances $\{\widetilde{\varepsilon}_t\}_{t=1}^{T}$ can be recovered[15]. In this context $\widetilde{\varepsilon}_1$ is interpreted to be a news shock.

$$\begin{bmatrix} \Delta\left(y_t/h_t\right) \\ p_t^b \end{bmatrix} = \widetilde{\Gamma} \begin{bmatrix} \widetilde{\varepsilon}_{1,t} \\ \widetilde{\varepsilon}_{2,t} \end{bmatrix}$$

In order to distinguish news shocks from other productivity shocks BP use another set of restrictions. These restrictions correspond to the assumption that news shocks have no contemporaneous impact on labor productivity, identifying a matrix $\widehat{\Gamma}$ and corresponding set economic disturbances $\{\widehat{\varepsilon}_t\}_{t=1}^{T}$ [16]. In this context $\widehat{\varepsilon}_2$ is interpreted to be a news shock.

$$\begin{bmatrix} \Delta\left(y_t/h_t\right) \\ p_t^b \end{bmatrix} = \widehat{\Gamma} \begin{bmatrix} \widehat{\varepsilon}_{1,t} \\ \widehat{\varepsilon}_{2,t} \end{bmatrix}$$

BP argue that the remarkable similarities between $\widetilde{\varepsilon}_1$ and $\widehat{\varepsilon}_2$ is evidence of news shocks in the data. Figure 4 plots the impulse responses of labor productivity and bond prices to these disturbances[17]. The response of the long-term bond price to a news shock is essentially the same for both identification schemes. Furthermore, the response is very persistent and positive, corresponding to a sustained decline in the long-term yield.

---

[15]Long run identification is achieved using an estimate of the spectral density at frequency zero. Let $S_0 = (I - B(1))^{-1} V (I - B(1)')^{-1}$ where $B(1) = \sum_{k=1}^{p} B_k$ and $V = E[u_t u_t']$. Let D be the lower triangular cholesky factorization of the spectral density at frequency zero. Then the long run restriction in this 2 variable VAR amounts to

$$\widetilde{\Gamma} = (I - B(1))\, D$$

[16]The short run restriction corresponds to a lower triangular cholesky factorization of the variance-covariance matrix of VAR residuals

$$\widehat{\Gamma} = chol(V)$$

[17]The VAR is estimated with 2 lags, the optimal choice according to both the Aikike and Hannon Quinn information criteria. The results are aslo robust to using 3 lags

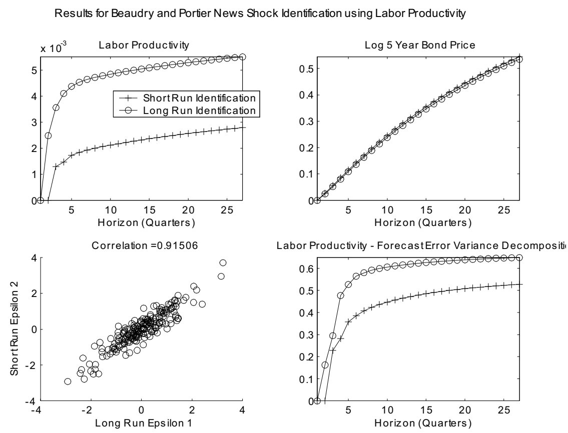


Figure 1.3. Beaudry and Portier news decomposition of bond prices

This story is consistent with the supply-side interpretation of news shocks: Anticipated increases in supply induce a deflation to which an inflation-targeting central bank would drop the short-term interest rate. The impulse response of labor productivity to the two disturbances also exhibit similar persistence and direction though are quite different in magnitude. The short-run restriction used to identify the news shock can be seen in the delayed response of labor productivity. Further evidence for the news disturbances in the data are given by the high correlation between $\widetilde{\varepsilon}_1$ and $\widehat{\varepsilon}_2$, as seen in the lower-left plot of Figure 4.

Decomposing the variance in labor productivity for each identification method reveals the significance of news shocks. Similar to BP's results using stock prices, these disturbances explain up to 50% of the volatility of labor productivity. If news shocks are as important as this exercise reveals then they should be incorporated in structural economic models. The appendix contains contains a figure where the analysis is carried out using a measure of total factor productivity instead of labor productivity.

1.3. Model

The model is closely related to Christiano, Eichenbaum and Evans (2005), a benchmark in the modern macroeconometric literature. It was originally designed to explain the dynamic response of macroeconomic aggregates to monetary policy shocks. The CEE model may lack several features important for understanding the intertemporal adjustment process associated with news shocks[18]. Nonetheless, the many desirable feature of the model documented by CEE as well as other authors[19] make it a sensible place to start an analysis of Pigou's theory of the business cycle[20].

The real frictions underlying the CEE model include internal habit formation in household preferences, costs to adjusting the level of investment, capacity utilization and monopolistic competition in the production and labor markets. Nominal frictions include sticky wages and prices modeled using the framework of Calvo (1983) and a cash in advance constraint on firms' wage bill. The model presented here contains 6 non-news shocks: permanent labor augmenting technology shock, investment specific technology shock, consumption preference shock, labor preference shock, government spending shock and a monetary policy shock. The unit root technology shock requires the model be made stationary before being solved, accomplished by scaling the equilibrium conditions. The model also includes some flexibility relative to CEE by allowing for partial indexing of prices and wages and partial financing of the wage bill. I will derive the model in

[18]Den Haan and Kaltenbrunner make a case for incorporating search frictions into the labor market specification when news shocks are present. This is due to the intertemporal type of adjustment cost associated with search frictions.

[19]see Smets and Wouters (2003) and Del Negro, Schorfheide, Smets and Wouters (2006)

[20]Christiano, Ilut, Motto and Rostagno (2006) use a similar model to explore news shocks. Their analysis focuses on boom/bust cycles as a consequence of an inflation targeting monetary policy.

recursive form and following the notation presented in Schmitt-Grohe and Uribe[21] (SGU, (2004)). Their methodology differs from CEE in their specification of price-setting, having first order effects in economic dynamics whereas in CEE it is of second-order. Though not pursued in this paper, nonlinear solutions (via perturbation methods) can be simply obtained when the model is cast in recursive form. These solutions are of economic interest in the context of the yield curve because they have nontrivial implications for risk premia (ie. time-varying instead of constant risk premia).

1.3.1. Households

.The model consists of a continuum of agents whose preferences are modeled using a representative family construct. The household's preferences are defined over per capita consumption, c_t, per capita hours worked, h_t^d, and per capita holdings of real money balances, m_t^h. The following utility function defines an ordering over these various bundles and is given by

$$E_0 \sum_{t=0}^{\infty} \beta^t \phi_{\beta,t} \left[\ln\left(c_t - bc_{t-1}\right) - \frac{\phi_{0,t}}{2}\left(h_t^d\right)^2 + \frac{\phi_1 \left(m_t^h\right)^{1-\sigma}}{1-\sigma} \right] \tag{1.1}$$

The value b captures internal habit formation. The preference shock, $\phi_{\beta,t}$, is modeled to be stationary exogenous stochastic process, implying that variations in agents' rate of time preference are temporary.

$$\ln \phi_{\beta,t} = \rho_\beta \ln \phi_{\beta,t-1} + \varepsilon_t^\beta \tag{1.2}$$

[21]SGU provide additional discussion and different interpretations of the economic mechanisms underlying the CEE model.

$\phi_{0,t}$ represents a shock to preferences for leisure. It evolves according to the following stationary exogenous stochastic process

$$\ln \phi_{0,t} = \left(1 - \rho_\phi\right) \ln \phi_0 + \rho_\phi \ln \phi_{0,t-1} + \varepsilon_t^\phi \tag{1.3}$$

Capital, k_t , depreciates at a rate δ but can be replenished by investment, i_t, which is subject to an adjustment cost

$$k_{t+1} = (1-\delta)k_t + \tau_t^k i_t \left(1 - \frac{\kappa}{2}\left(\frac{i_t}{i_{t-1}} - 1\right)^2\right) \tag{1.4}$$

τ_t^k represents an investment specific technology shock[22]. It evolves according to the following stationary exogenous stochastic process

$$\ln \tau_t^k = (1-\rho_k) \ln \tau^k + \rho_k \ln \tau_{t-1}^k + \varepsilon_t^k \tag{1.5}$$

The household budget constraint, expressed in real terms, is given by

$$b_t + c_t + i_t + m_t^h + a(u_t)k_t = \frac{b_{t-1}}{\pi_t} R_{t-1} + h_t^d \frac{w_t}{\bar{\mu}_t} + r_t^k u_t k_t + \varphi_t - \tau_t + \frac{m_{t-1}^h}{\pi_t} \tag{1.6}$$

Here b_t represents agents' purchases of 90 day t-bills, nominally state-contingent securities which pay interest at the rate R_t. Agents must decide how intensely capital should be operated which is captured by the capacity variable u_t. They rent effective capital, $u_t k_t$, to firms at a rate r_t, but entail a cost of $a(u_t)k_t$ consumption goods for operating the capital at the scale u_t. The function $a(\cdot)$ is assumed to satisfy $a(1) = 0$, $a'(1) > 0$ and $a''(1) > 0$. φ_t represents dividends payed by firms for time t profits. τ_t represents

[22]Fisher [**29**] argues that these disturbances are important in economic fluctuations

a lump sum tax levied by the government. $\frac{w_t}{\bar{\mu}_t}$ represents the effective real hourly wage agents earn from their work efforts. The term $\bar{\mu}_t$ can be interpreted as a wedge between agents' disutility of leisure and the prevailing average wage in the economy. Lastly, π_t represents inflation.

Agents maximize 1.1 choosing c_t, i_t, k_{t+1}, h_t^d, m_t^h, u_t and b_t subject to the constraints 1.4 and 1.6. Letting $\lambda_t q_t$ and λ_t be the Lagrange multipliers on 1.4 and 1.6 the first order conditions of the problem are:

$$\frac{1}{c_t - bc_{t-1}} - E_t\left[\frac{b}{c_{t+1} - bc_t}\left(\beta\frac{\phi_{\beta,t+1}}{\phi_{\beta,t}}\right)\right] = \lambda_t \tag{1.7}$$

$$\phi_0 h_t = \frac{\lambda_t w_t}{\bar{\mu}_t} \tag{1.8}$$

$$\lambda_t q_t = E_t\left[\left(\beta\frac{\phi_{\beta,t+1}}{\phi_{\beta,t}}\right)\lambda_{t+1}\left(r_{t+1}^k u_{t+1} - a(u_{t+1}) + q_{t+1}(1-\delta)\right)\right] \tag{1.9}$$

$$\begin{aligned} \lambda_t \quad &= \quad \tau_t^k \lambda_t q_t \left[1 - \frac{\kappa}{2}\left(\frac{i_t}{i_{t-1}} - 1\right)^2 - \kappa\frac{i_t}{i_{t-1}}\left(\frac{i_t}{i_{t-1}} - 1\right)\right] ... \\ &\quad ... + E_t\left[\kappa\left(\beta\frac{\phi_{\beta,t+1}}{\phi_{\beta,t}}\right)\lambda_{t+1}\tau_{t+1}^k q_{t+1}\left(\frac{i_{t+1}}{i_t}\right)^2\left(\frac{i_{t+1}}{i_t} - 1\right)\right] \end{aligned} \tag{1.10}$$

$$r_t^k = a'(u_t) \tag{1.11}$$

$$\lambda_t = \phi_1 \left(m_t^h\right)^{-\sigma} + E_t \left[\left(\beta \frac{\phi_{\beta,t+1}}{\phi_{\beta,t}} \right) \frac{\lambda_{t+1}}{\pi_{t+1}} \right] \tag{1.12}$$

$$\frac{\lambda_t}{R_t} = E_t \left[\left(\beta \frac{\phi_{\beta,t+1}}{\phi_{\beta,t}} \right) \frac{\lambda_{t+1}}{\pi_{t+1}} \right] \tag{1.13}$$

1.3.2. Wage Setting by Households

Based on the labor market model of Erceg, Henderson and Levin (2000), CEE suppose that the household is a monopoly supplier of differentiated labor service, h_t^j. It sells this service to a representative, competitive firm that transforms it into an aggregate labor input h_t, using the following technology

$$h_t^d = \left(\int_0^1 \left(h_t^j\right)^{\frac{\overline{\eta}-1}{\overline{\eta}}} dj \right)^{\frac{\overline{\eta}}{\overline{\eta}-1}} \tag{1.14}$$

Profit maximization yields a demand curve for h_t^j given by

$$h_t^j = \left(\frac{W_t}{W_t^j} \right)^{\overline{\eta}} h_t^d \tag{1.15}$$

Note that $\overline{\eta}$ is the wage elasticity of demand for a specific labor variety. Also, w_t is the real wage corresponding to the aggregate labor input h_t^d and $W_t^j = p_t w_t^j$ is the nominal wage corresponding to labor variety j.

Households can optimally set wages in a fraction $1 - \overline{\alpha}$ of randomly chosen labor markets at time t. Markets in which the households cannot set the wage optimally use the following indexing rule

$$W_t^j = W_{t-1}^j \pi_{t-1}^{\overline{\chi}} \tag{1.16}$$

Wages set optimally, $\widetilde{w}_t$, are a markup over the the marginal cost of labor that would prevail in the absence of wage stickiness

$$\begin{aligned} 0 \quad = \quad & E_t \sum_{s=0}^{\infty} \overline{\alpha}^s \beta^s \frac{\phi_{\beta,s}}{\phi_{\beta,0}} \lambda_{t+s} \left(\frac{\widetilde{w}_t}{w_{t+s}} \right)^{-\overline{\eta}} h_{t+s}^d \prod_{k=1}^{s} \left(\frac{\pi_{t+k}}{\pi_{t+k-1}^{\overline{\chi}}} \right)^{\overline{\eta}} \times ... \\ & ... \times \left[\frac{\overline{\eta} - 1}{\overline{\eta}} \frac{\widetilde{w}_t}{\prod_{k=1}^{s} \left(\frac{\pi_{t+k}}{\pi_{t+k-1}^{\overline{\chi}}} \right)} - \frac{w_{t+s}}{\overline{\mu}_{t+s}} \right] \end{aligned} \tag{1.17}$$

$\overline{\mu}_{t+s}$ is simply the wedge between the disutility of labor and the average real wage prevailing in the economy. Note in an economy without labor frictions this wedge is always zero. This can be seen by recalling 1.8

$$\bar{\mu}_t = \frac{\lambda_t w_t}{\phi_0 h_t} \tag{1.18}$$

It is useful to rewrite 1.17 in recursive form. The following equations derived by SGU accomplish this

$$f_t^1 = \lambda_t \left(\frac{w_t}{\bar{w}_t}\right)^{\bar{\eta}} h_t^d + \bar{\alpha}\left(\beta\frac{\phi_{\beta,t+1}}{\phi_{\beta,t}}\right)\left(\frac{\pi_{t+1}}{\pi_t^{\bar{\chi}}}\right)^{\bar{\eta}-1}\left(\frac{\bar{w}_{t+1}}{\bar{w}_t}\right)^{\bar{\eta}} f_{t+1}^1 \tag{1.19}$$

$$f_t^2 = \frac{\lambda_t}{\bar{\mu}_t} w_t \left(\frac{w_t}{\bar{w}_t}\right)^{\bar{\eta}} h_t^d + \bar{\alpha}\left(\beta\frac{\phi_{\beta,t+1}}{\phi_{\beta,t}}\right) E_t \left(\frac{\pi_{t+1}}{\pi_t^{\bar{\chi}}}\right)^{\bar{\eta}}\left(\frac{\bar{w}_{t+1}}{\bar{w}_t}\right)^{\bar{\eta}} f_{t+1}^2 \tag{1.20}$$

$$0 = (1-\bar{\eta})\bar{w}_t f_t^1 + \bar{\eta} f_t^2 \tag{1.21}$$

1.3.3. Final Goods Firms

A final consumption good, y_t, is produced by a perfectly competitive, representative firm. The firm produces this final good by combining a continuum of intermediate good varieties, indexed by $i \in [0,1]$, using the following technology:

$$y_t = \left(\int_0^1 \left(y_t^i\right)^{\frac{\eta-1}{\eta}} di\right)^{\frac{\eta}{\eta-1}} \tag{1.22}$$

The firm takes its output price, p_t, and input prices, p_t^j, as given. Profit maximization by this firm implies the following demand for goods of variety j

$$y_t^i = \left(\frac{p_t}{p_t^i}\right)^{\eta} y_t \tag{1.23}$$

An expression for the aggregate price level can be obtained by integrating 1.23 over the set of firms while imposing 1.22. This yields the following expression

$$p_t = \left[\int_0^1 \left(p_t^i\right)^{1-\eta} di \right]^{\frac{1}{1-\eta}} \tag{1.24}$$

1.3.4. Intermediate Goods Firms

Intermediate goods of variety $i \in [0, 1]$ are produced by a monopolist using capital, rented out by the household sector, and labor, also provided by the household sector.

$$y_t^i = \left(k_t^i\right)^{\theta} \left(z_t h_t^i\right)^{1-\theta} - z_t \psi \tag{1.25}$$

z_t represents a labor-augmenting technology shock, modelled as a unit root which is consistent with the identification assumptions used in the VAR analysis of news shocks. $z_t \psi$ represents the fixed cost of production, a cost that scales with the level of technology operating in the economy. $\chi_t = z_t / z_{t-1}$ is assumed to evolve according to the following stochastic process

$$\ln \chi_t = (1 - \rho_z) \ln g_z + \rho_z \ln \chi_{t-1} + \varepsilon_t^z \tag{1.26}$$

where g_z is the steady state growth rate of technology in the economy.

Firms must finance a portion of their wage bill prior to engaging in production. They pay interest on loans equal to the nominal interest rate prevailing in the economy. The effective hourly wage payed by the firm including interest costs is given by

$$w_t \left[1 + v \frac{R_t - 1}{R_t} \right] \tag{1.27}$$

Here the term v can be interpreted as the fraction of funds that must be borrowed in advance. Letting m_t^i denote loans to the firm, the constraint is simply:

$$m_t^i \geq v w_t h_t^i$$

Dividends of the firm, paid out to households, can be written as

$$\varphi_t^i = p_t^i y_t^i - r_t^k k_t^i - w_t \left[1 + v \frac{R_t - 1}{R_t}\right] h_t^i \tag{1.28}$$

Price setting by intermediate goods firms is modeled analagous to that of wage setting. It is assumed that a fraction $1 - \alpha$ of firms are randomly chosen to reoptimize their price in each period. Those firms not chosen to reoptimize update their prices according to the following indexing rule

$$p_t^i = p_{t-1}^i \pi_{t-1}^{\chi} \tag{1.29}$$

Firms chosen to set prices optimally do so by maximizing the present discounted value of dividends payed out to households. In valuing these dividends, firms use the households' marginal rate of intertemporal substitution. Cash flows valued in period t, expected to arrive in period $t+s$, are thus discounted by the stochastic nominal discount factor

$$\Lambda_{t,t+s} = \left(\beta^s \frac{\phi_{\beta,t+s}}{\phi_{\beta,t}}\right) \frac{\lambda_{t+s}}{\lambda_t} p_{t+s} \tag{1.30}$$

Using 1.23 and 1.28 and the constraint that firms must satisfy demand at posted prices, the problem of the firm can be expressed using the following Lagrangian:

$$
\begin{aligned}
L \;=\; & E_t \sum_{s=0}^{\infty} \Lambda_{t,t+s} \alpha^s \{ \left(\frac{\widetilde{p}_t}{p_t} \right)^{1-\eta} \prod_{k=1}^{s} \left(\frac{\pi_{t+k-1}^{\chi}}{\pi_{t+k}} \right)^{1-\eta} y_{t+s} - r_{t+s}^k k_{t+s}^i \\
& ... - w_{t+s} h_{t+s}^i \left[1 + v \left(1 - R_{t+s}^{-1} \right) \right] ... \\
& ... - r_{t+s}^k k_{t+s}^i - w_{t+s} h_{t+s}^i \left[1 + v \left(1 - R_{t+s}^{-1} \right) \right] \\
& ... + mc_{t+s}^i \left[\left(k_t^i \right)^{\theta} \left(z_t h_t^i \right)^{1-\theta} - z_t \psi - \left(\frac{\widetilde{p}_t}{p_t} \right)^{-\eta} \prod_{k=1}^{s} \left(\frac{\pi_{t+k-1}^{\chi}}{\pi_{t+k}} \right)^{-\eta} y_{t+s} \right] \}
\end{aligned}
$$

The mulitplier on the constraint admits an expression for the marginal cost of producing an additional unit of output. $\widetilde{p}_t$ represents the optimal price set by a firm at time t. Note that the choice of capital and labor by firm i are intratemporal decisions. The optimal price set by firm i at time t is given by the following first order condition

$$
\begin{aligned}
0 \;=\; & E_t \sum_{s=0}^{\infty} \Lambda_{t,t+s} \alpha^s \left(\frac{\bar{p}_t}{p_t} \right)^{-\eta} \prod_{k=1}^{s} \left(\frac{\pi_{t+k-1}^{\chi}}{\pi_{t+k}} \right)^{-\eta} y_{t+s} \times ... \\
& ... \times \left[\frac{\eta - 1}{\eta} \left(\frac{\widetilde{p}_t}{p_t} \right) \prod_{k=1}^{s} \left(\frac{\pi_{t+k-1}^{\chi}}{\pi_{t+k}} \right) - mc_{t+s}^i \right]
\end{aligned} \tag{1.31}
$$

Similarly, the optimal choice of capital and labor are given by the following first order conditions

$$
r_t^k = \theta mc_t^i \left(\frac{k_t^i}{z_t h_t^i} \right)^{\theta - 1} \tag{1.32}
$$

$$
w_t \left[1 + v \left(1 - R_t^{-1} \right) \right] = (1 - \theta) mc_t^i \left(\frac{k_t^i}{z_t h_t^i} \right)^{\theta} \tag{1.33}
$$

Note that these can be rearranged to express marginal cost as a function of the relevant factor costs

$$mc_t^i = \left(\frac{1}{1-\theta}\right)^{1-\theta}\left(\frac{1}{\theta}\right)^{\theta}\left(r_t^k\right)^{\theta}\left(w_t\left[1+v\left(1-R_t^{-1}\right)\right]\right)^{1-\theta}$$

Condition 1.31 can be expressed in recursive form. Defining $\bar{p}_t = \widetilde{p}_t/p_t$. this is accomplished by the following set of equations

$$x_t^1 = y_t mc_t \bar{p}_t^{-\eta-1} + \alpha E_t\left[\left(\beta\frac{\phi_{\beta,t+1}}{\phi_{\beta,t}}\right)\frac{\lambda_{t+1}}{\lambda_t}(\bar{p}_t/\bar{p}_{t+1})^{-\eta-1}(\frac{\pi_t^{\chi}}{\pi_{t+1}})^{-\eta}x_{t+1}^1\right] \tag{1.34}$$

$$x_t^2 = y_t \bar{p}_t^{-\eta} + \alpha E_t\left[\left(\beta\frac{\phi_{\beta,t+1}}{\phi_{\beta,t}}\right)\frac{\lambda_{t+1}}{\lambda_t}\left(\frac{\pi_t^{\chi}}{\pi_{t+1}}\right)^{1-\eta}(\frac{\bar{p}_t}{\bar{p}_{t+1}})^{-\eta}x_{t+1}^2\right] \tag{1.35}$$

$$\eta x_t^1 = (\eta-1)x_t^2 \tag{1.36}$$

1.3.5. Government and Monetary Authority

Government expenditure is financed via lump-sum taxes and seignorage

$$g_t = \tau_t + m_t - \frac{m_{t-1}}{\pi_t}$$

where m_t is total money holding of the household and firm sectors combined. It is assumed that government expeniduture evolves according to the following exogenous stochastic process

$$\ln \widetilde{g}_t = \left(1 - \rho_g\right) \ln g + \rho_g \ln \widetilde{g}_{t-1} + \varepsilon_t^g$$

where $\widetilde{g}_t = g_t / z_t$ is the stationary representation consistent with the unit root specification of the technology process.

Monetary policy is given by the following taylor rule for the short rate

$$\widehat{R}_t = \alpha_R \widehat{R}_{t-1} + (1 - \alpha_R) \left(\alpha_\pi E_t \left[\widehat{\pi}_{t+1}\right] + \alpha_y \widehat{y}_t\right) + mp_t$$

where a hat over the variable represents denotes the log-deviation of that variable from its deterministic steady state level. The policy rule is nearly the same as that analyzed in Clarida, Gali and Gertler (2000). It assumes that the monetary authority sets the interest rate using information about expected inflation and a measure of the output gap. Furthermore, the policy rule incorporates the lagged interest rate to capture the persistent nature of interest rate time-series. The monetary policy shock evolves according to the following exogenous stochastic process

$$\ln mp_t = \rho_{mp} \ln mp_{t-1} + \varepsilon_t^{mp} \tag{1.37}$$

1.3.6. Aggregation and Equilibrium

1.3.6.1. Market Clearing in the Production Sector. Using 1.24 the aggregate price index can be written as

$$p_t^{1-\eta} = \alpha \left(p_{t-1} \pi_{t-1}^{\chi}\right)^{1-\eta} + (1 - \alpha) \widetilde{p}_t^{1-\eta}$$

Dividing through by the aggregate price index yields the following stationary equilibrium equation

$$1 = \alpha \left(\pi_t^{-1}\pi_{t-1}^{\chi}\right)^{1-\eta} + (1-\alpha)\bar{p}_t^{1-\eta} \tag{1.38}$$

Using 1.23 and 1.25, market clearing in the final goods market yields the following equation

$$\left(k_t^i\right)^{\theta}\left(z_t h_t^i\right)^{1-\theta} - z_t\psi = \left(c_t + i_t + g_t + a(u_t)k_t\right)\left(\frac{p_t^i}{p_t}\right)^{-\eta} \tag{1.39}$$

and market clearing in the labor market requires

$$h_t^d = \int_0^1 h_t^i di$$

and in the capital market requires

$$u_t k_t = \int_0^1 k_t^i di$$

Integrating 1.39 over all firms yields

$$\left(u_t k_t\right)^{\theta}\left(z_t h_t^d\right)^{1-\theta} - z_t\psi = \left(c_t + i_t + g_t z_t + a(u_t)k_t\right)s_t \tag{1.40}$$

where

$$s_t = \int_0^1 \left(\frac{p_t^i}{p_t}\right)^{-\eta} di$$

The appearance of $s_t$in the resource constraint captures the output loss due to price dispersion. SGU show that the model has the property that s_t is bounded below by 1, its value in steady state. This term admits the following recursive representation

$$s_t = (1-\alpha)\bar{p}_t^{-\eta} + \alpha \left(\frac{\pi_t}{\pi_{t-1}^{\chi}}\right)^{\eta} s_{t-1} \tag{1.41}$$

1.3.6.2. Market Clearing in the Labor Market. Recall that the aggregate demand for labor of variety j is given by

$$h_t^j = \left(\frac{W_t^j}{W_t}\right)^{-\overline{\eta}} h_t^d$$

where $h_t^d = \int_0^1 h_t^i di$ denotes the aggregate demand for the composite labor input but firms. Labor demand in markets where wages change optimally are

$$\widetilde{h}_t = \left(\frac{\widetilde{w}_t}{w_t}\right)^{-\overline{\eta}} h_t^d$$

Wage dispersion, as in the case of price dispersion, creates an inefficiency in the demand for labor. A fraction of labor markets, $1-\overline{\alpha}$, use a wage indexing rule and the aggregate demand for labor for these markets can be written

$$h_t = (1-\overline{\alpha})\, h_t^d \overline{s}_t \tag{1.42}$$

where $\overline{s}_t$ is a measure of wage dispersion in these markets. Noting the timing assumptions, the average wage in these markets can be found by properly indexing past optimal wages

$$\overline{s}_t = \sum_{s=0}^{\infty} \overline{\alpha}^s \left(\frac{\widetilde{W}_{t-s} \prod_{k=1}^{s} \pi_{t+k-s-1}^{\overline{\chi}}}{W_t}\right)^{-\overline{\eta}}$$

$\overline{s}_t$ can be written in recursive form as follows

$$\overline{s}_t = (1-\overline{\alpha})\left(\frac{\widetilde{w}_t}{w_t}\right)^{-\overline{\eta}} + \overline{\alpha}\left(\frac{w_{t-1}}{w_t}\right)\left(\frac{\pi_t}{\pi_{t-1}^{\overline{\chi}}}\right)^{\overline{\eta}} \overline{s}_{t-1} \tag{1.43}$$

The aggregate real wage can be written in an analagous manner to the aggregate price level

$$w_t^{1-\overline{\eta}} = (1-\overline{\alpha})\widetilde{w}_t^{1-\overline{\eta}} + \overline{\alpha} w_{t-1}^{1-\overline{\eta}} \left(\frac{\pi_{t-1}^{\overline{\chi}}}{\pi_t}\right)^{1-\overline{\eta}} \tag{1.44}$$

1.4. Introduction of News Shocks

Future fundamental economic disturbances can be incorporated into any of the exogenous stochastic processes of the model. To show how this is done consider the a simple $ar(1)$ stochastic process z_t embedded in the following canononical form:

$$\begin{bmatrix} z_t \\ \zeta_{t+p}^t \\ \zeta_{t+p-1}^t \\ \vdots \\ \zeta_{t+1}^t \end{bmatrix} = \begin{bmatrix} \rho & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ & & & & \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} z_{t-1} \\ \zeta_{t+p}^{t-1} \\ \zeta_{t+p-1}^{t-1} \\ \vdots \\ \zeta_t^{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_t^t \\ \zeta_{t+p}^t \\ \epsilon_t^{t+p-1} \\ \vdots \\ \epsilon_t^{t+1} \end{bmatrix}$$

where ζ_{t+j}^t is interpreted as a news shock: a fundemental economic impulse expected to arrive in period $t+j$. The values ϵ_t^{t+j} capture the time t revisions in expectations, updating expectations from ζ_{t+j}^{t-1} to ζ_{t+j-1}^t, assumed to be mean zero and uncorrelated over time and across horizons.

Note that the canonical form above implies the following expectations

$$E_t z_{t+j} = \rho^j z_{t+j} + \sum_{k=1}^{j} \rho^{j-k} \zeta_{t+k}^t$$

where ζ_{t+1}^t is an impulse first anticipated p periods ago,updated each period since its arrival. In this context it is easy to see the effects of news shocks on agents' expectations about the future path of the stochastic process.

1.5. Solution and Estimation

The presence of a unit root technology shock requires that the model be scaled before being solved. The appendix defines the stationary competitive equilibrium associated with the model. The model is solved by log-linearizing around the nonstochastic steady state, resulting in the following state-space representation

$$\begin{aligned} S_t &= FS_{t-1} + Q\varepsilon_t \\ Obs_t &= Hs_t + v_t \end{aligned}$$

where S_t represents the state and Obs_t the observation vector. ε_t corresponds to the fundamental economic impulses and v_t is a vector of measurement error, both assumed to be distributed gaussian. The variance of the measurement error for the yield variables is discussed in the next section. The following observation vector is used in the estimation of the model[23]

$$Obs_t = \left[\Delta y_t, \quad \pi_t, \quad u_t, \quad h_t, \quad \left(\frac{c_t}{y_t}\right), \quad \left(\frac{i_t}{y_t}\right), \quad yields_t'\right]' \tag{1.45}$$

where $yields_t$ is a vector of 5 yields at time t on zero coupon bonds maturing in 1-5 years. The following section describes the way bonds are priced in the model and the appendix contains a description of the data. In the appendix the results are presented for an alternative observation equation where the model's return on capital is made to correspond to the real return on the S&P 500,

$$R_t^k = \frac{r_t^k + q_t(1-\delta)}{q_{t-1}}$$

Stock market data is incorporated into the observation equation because it is likely to contain information about new shocks, as shown in Beaudry and Portier (2006). Other authors have used stock market data in the estimation of DSGE's (see, for instance, Christiano, Ilut, Motto and Rostagno (2007)). The results presented herein are robust to the inclusion/exclusion of stock price data, suggesting that this series is not critical for identifying the role of news shocks in the data.

I estimate the model using Bayesian methods to fit the vector of observables defined in 1.45. Many recent papers have employed these Bayesian methods to estimate DSGE models of the economy[24]. Smets and Wouters (2003) show that a variant of the CEE model estimated using Bayesian techniques has similar fit to that of an a-theoretic Bayesian VAR.

[23]The data is demeaned prior to estimation. Thus, the econometric procedure fits deviations from mean. The model parameters that determine the levels (or means) are fixed in the estimation and made to coincide with the mean of from the data.
[24]Examples include the work of Dejong, Ingram and Whiteman who employ importance sampling (2000) and Otrok (2001) who uses a Metropolis-Hastings algorithm.

Bayesian methods also have an appealing feature as illustrated by Fernandez-Villaverde and Rubio Ramirez (2005). These authors show that Bayesian point estimates converge to their psuedo-true values. Furthermore, the model obtaining the highest posterior probability in Bayesian model comparison is also the one which fits the best according to the Kullback-Liebler entropy measure.

The priors used in the analysis are given in the following table. The standard deviation of the news shocks are estimated assuming a uniform prior over the region 0 to 0.5. The priors are motivated by those used in similar Bayesian analyses of macroecnomic models (see, for instance, Del Negro, Schorfheide, Smets and Wouters (2006)). Due to the large dimension of the structural model I fix many of the parameters prior to estimation. The fixed parameters determine the steady-state values of the economy and are chosen to match the historical means of the observables used in the estimation. Since these are made to coincide ex-ante I estimate the model using demeaned data.

The measurement error prior is uniform with mass on regions of the parameter space where measurement error can explain no more than 5% of the variance of the observables (30% in the case of the yields). This choice of prior corresponds to the requirement that the shocks explain the majority of the variation in the observer equation and can be viewed as a prior on the set of shocks important for understanding economic dynamics. For this reason the 6 'traditional' contemporaneous shocks are chosen to represent the economic disturbances the literature has found to be important for explaining the bulk of aggregate fluctuations.

Description	Parameter	Distribution	Mean	Sdev
Mean Growth Rate	g_z	Fixed	1.0037	0
Discount Rate	β	Fixed	0.9926	0
Capital Share	θ	Fixed	0.35	0
Fixed Cost	ψ	Fixed	0.2961	0
Depreciation	δ	Fixed	0.025	0
Firm Borrowing Constraint	ν	Uniform(0,1)	0.5	0.29
Price Elasticity of Demand	η	Fixed	6	0
Wage Elasticity of Demand	$\overline{\eta}$	Fixed	21	0
Price Stickiness	α	Beta	0.6	0.1
Wage Stickiness	$\overline{\alpha}$	Beta	0.64	0.1
Internal Habit	b	Fixed	0.70	0
Money Demand Param.	σ	Fixed	10.62	0
Investment Adj. Cost	κ	Normal	5	0.5
Price Indexing	χ	Uniform(0,1)	0.5	0.29
Wage Indexing	$\overline{\chi}$	Uniform(0,1)	0.5	0.29
Capacity Utlilization Param.	γ_2	Beta	0.2	0.1
Inflation Target (Quarterly)	π^*	Fixed	1.0074	0
Government Spending	g	Fixed	0.53	0
Labor Pref. Param.	ϕ_0	Fixed	1.11	0

Money Demand Param.	ϕ_1	Fixed	0.5393	0
Capital Accumulation Param	τ^k	Fixed	1	0
Inflation Feedback	α_π	Normal	2	0.25
Output Gap Feedback	α_y	Beta	0.08	0.03
Interest Rate Feedback	α_R	Beta	0.8	0.1
Autoreg. Tech Shock	ρ_z	Beta	0.2	0.15
Autoreg. Gov. Shock	ρ_g	Beta	0.9	0.1
Autoreg. Pref. Shock	ρ_β	Beta	0.9	0.1
Autoreg. Labor Pref. Shock	ρ_{ϕ_0}	Beta	0.9	0.1
Autoreg. Inv. Specific Shock	ρ_k	Beta	0.9	0.1
Autoreg. Mon. Pol. Shock	ρ_{mp}	Normal	0	0.025
Sdev Tech Shock	σ_z	Inv Gamma	0.02	0.05
Sdev. Gov. Shock	σ_g	Inv Gamma	0.01	0.05
Sdev. Pref. Shock	σ_β	Inv Gamma	0.05	0.05
Sdev. Labor Pref. Shock	σ_{ϕ_0}	Inv Gamma	0.04	0.05
Sdev. Inv. Specific Shock	σ_k	Inv Gamma	0.05	0.05
Sdev. Mon. Pol. Shock	σ_{mp}	Inv Gamma	0.0015	0.0015

1.6. Affine Bond Prices

The pricing kernel process in the model is given by

$$M_{t+1} = \beta \frac{\phi_{\beta,t+1}}{\phi_{\beta,t}} \frac{\lambda_{t+1}}{\lambda_t} \frac{1}{\pi_{t+1}} \tag{1.46}$$

Define $m_t \equiv \log M_t$ and the steady state of the pricing kernel as M (the log of the steady state defined analagously as m). Log-linearizing 1.46 results in the following approximation

$$\widehat{m}_{t+1} \approx \widehat{\phi}_{\beta,t+1} - \widehat{\phi}_{\beta,t} + \widehat{\lambda}_{t+1} - \widehat{\lambda}_t - \widehat{\pi}_{t+1} \tag{1.47}$$

which I will be using to price zero-coupon bonds. The pricing kernel prices all securities subject to the relation

$$1 = E_t\left[M_{t+1}R_{t+1}\right] \tag{1.48}$$

where R_{t+1} is the time $t+1$ return on a financial security. In particular, for an n-period bond, $R_{t+1} = \frac{P_{t+1}^{n-1}}{P_t^n}$, with P_t^n the time t price of an n-period zero coupon bond. If $M_{t+1} > 0$ for all t, the resulting returns satisfy the no arbitrage conditions (Harrison and Kreps (1979)). The approximation provided by 1.47 maps into a more general affine pricing framework. In this setting the log of the pricing kernel is a conditionally linear process, which can be written in the following general form:

$$m_{t+1} \equiv \log(M_{t+1}) = -r_t - \frac{1}{2}\Lambda_t' D \Lambda_t - \Lambda_t' \varepsilon_{t+1}$$

where $\Lambda_t = \Lambda_0 + \Lambda_1 S_t$ and r_t represents the interest rate on a 3 month zero coupon bond. Dai and Singleton (1996) examine the case where $D_t = D$, the Gussian price of risk model, and claim it accounts for the deviations of the Expectations Hypothesis observed in US term structure data. Several other popular bond pricing models, such as that of Cox, Ingersoll and Ross (1985), can be framed in this general structure. The DSGE model presented in this exposition corresponds to a particular model withing this class where Λ_0 is a function of the model parameters, $\Lambda_1 = 0$ and $D_t = D$. This model is homoskedastic with constant prices of risk, implying that the expectations hypothesis holds.

In a log-normal model, like that presented here, pricing a 1 period bond implies

$$E_t\left[m_{t+1}\right] + .5V_t\left[m_{t+1}\right] = -r_t$$

Noting 1.47, the log-linear solution of the model implies that the log-deviation of the pricing kernel from its steady-state is an affine function of the state vector:

$$\begin{aligned}
\widehat{m}_{t+1} &= \Psi' S_{t+1} \\
&= \Psi'\left(FS_t + Q\varepsilon_{t+1}\right)
\end{aligned}$$

for some vector Ψ. Now note that

$$E_t\left[m_{t+1}\right] = \Psi' F S_t + m$$

and the conditional variance can be expressed as

$$\begin{aligned}
V_t\left[m_{t+1}\right] &= V_t\left[\widehat{m}_{t+1}\right] \\
&= \Psi' Q Q' \Psi
\end{aligned}$$

Thus,

$$\underbrace{\Psi' F S_t + m}_{E_t[m_{t+1}]} + \frac{1}{2}\underbrace{\Psi' Q Q' \Psi}_{V_t[m_{t+1}]} = -r_t$$

and rearranging

$$\begin{aligned} m_{t+1} &= -r_t - \frac{1}{2}\Psi' Q Q' \Psi + \Psi' Q \varepsilon_{t+1} \\ m_{t+1} &= -r_t - \frac{1}{2}\Lambda'\Lambda - \Lambda' \varepsilon_{t+1} \end{aligned}$$

where $\Lambda' \equiv -\Psi' Q$ correspond to the prices of risk for the innovations ε_{t+1}. In this model the bond pricing equation is affine

$$p_t^n = a_n + b_n' S_t$$

where, by log-normality,

$$p_t^n = E_t\left[m_{t+1} + p_{t+1}^{n-1}\right] + \frac{1}{2}V_t\left[m_{t+1} + p_{t+1}^{n-1}\right]$$

Using an induction argument (see, for instance, the appendix of Cochrane and Piazzesi (2005)) the bond-price coefficients satisfy the following recursion

$$\begin{aligned} A_{n+1} &= A_n - \Lambda' Q' B_n + \frac{1}{2} B_n' Q Q' B_n \\ B_{n+1}' &= -H_r' + B_n' F \end{aligned}$$

where H_r is the vector mapping the state vector into the short rate, defined implicitly by $\widehat{r}_t = H_r S_t$. In equilibrium the yields can be expressed as a simple affine function of the state vector

$$y_{n,t} = -\frac{A_n}{n} - \frac{B_n'}{n} S_t$$

where $-\frac{A_n}{n}$ represents the constant risk premium and $-\frac{B_n'}{n}$ the loadings on the state vector.

1.7. Results

Several versions of the model are estimated using Bayesian techniques. In order to understand and quantify the effects of news shocks I consider several different types and time horizons for their arrival. BP [**7**] focus their analysis on the potential for revisions in agents' expectations about future permanent changes in technology to be the source of economic downturns. The baseline news shock model considered in this analysis is similar in spirit to the ideas pursued by BP, allowing only for news shocks to arrive via the unit root technology process with a maximum horizon of 3 years. Since news shocks are likely to arise in many different sectors of the economy I will also consider several variants of the model that allow other types of news disturbances.

The first type of alternative news shock considered in this analysis allows for agents to anticipate future deviations of government spending, consistent with the evidence on fiscal policy expenditures being known well in advance. The second type of news shock considered occurs via the stochastic process for investment specific technology. Several authors [25] have highlighted the quantitative importance of such disturbances in economic fluctuations. Furthermore, these types of economic impulses are thought to be a reduced form proxy for financial market disturbances. It is well known that financial markets are

[25] Example includes Fisher [**29**]

leading indicators of future economic activity. Consistent with this evidence I expand the set of news disturbances to include expected future exogenous innovations to the capital accumulation process.

The 'baseline' model used in the analysis, to which all news shock variants will be compared, only allows for conventional contemporaneous fundamental economic impulses. Bayesian estimation allows for model comparison using the marginal likelihood metric, given by the following formula

$$P(Data|Model_i) = \int L(\theta^{(i)}|Data, Model_i)p(\theta^{(i)}|Model_i)d\theta^{(i)}$$

Geweke's [**30**] modified harmonic mean estimator is used to evaluate the above integral[26]. The appendix contains a formula for approximating this integral based on monte-carlo methods.

The following table summarizes the various models considered in the analysis. In Model #5, the model containing all three types of news shocks, I allow for investment specific news shocks to occur at a 12 quarter horizon instead of 6. This is because I have found that in preliminary work the data tends to identify these shocks at longer horizons.

[26]This is discussed in the appendix. The marginal likelihood includes a penalty for the number of parameters.
The appendix also derives a methodology for increasing the computational precision of the marginal likelihood calculation

Model #	# Technology News Shocks	# Gov Spending News Shocks	# Inv. Specific News Shocks
1	0	0	0
2	12	0	0
3	0	12	0
4	0	0	12
5	6	6	12

The table below contains information about the marginal likelihood of each model. Perhaps the most striking result is the superior fit of model's incorporating news shocks relative to Model #1 which omits these type of economic disturbances. Recall that the marginal likelihood is a measure that corrects for the dimension of the parameter space. Thus, the positive values of the log marginal likelihood ratio imply that news shocks provide model's the degrees of freedom necessary to explain the dynamics in the data. News shocks, regardless of the channel through which they arrive, improve upon the model incorporating only conventional exogenous stochastic processes for the fundamental economic disturbances. This result is robust across information sets [27] suggesting that news shocks are statistically significant components of aggregate economic data.

The table shows that the model incorporating all three types of news shocks (here denoted Model #5) does the best among the models considered. Thus, all three types of news shocks are statistically significant components of the data according the the marginal likelihood measure. Previously only news shocks arriving via the stochastic process for techonology have been proposed as important contributors to aggregate fluctuations

[27] The appendix contains information about the marginal likelihood of these models for different data vectors

(see Beaudry and Portier, 2006). The marginal likelihood values imply that news shocks arriving via the stochastic process for investment specific technology are the most significant amongst the single news shock versions of the model. This result suggests that future research on the importance of news in economic dynamics should explore these additional channels.

Unconditional Variance Decomposition of Output Growth						
Observations include yields but no stock returns						
Model		1	2	3	4	5
log ratio Marg Lik[28]		0	474.79	388.74	505.01	530.04
Neutral tech.	χ_t	1.08	0.13	0.09	0.14	0.19
Government	g_t	17.58	21.24	32.10	8.53	8.37
Demand	$\varphi_{b,t}$	7.91	6.46	0.96	5.32	4.75
Labor supply	$\varphi_{0,t}$	66.89	29.50	11.50	19.18	18.57
Investment	$\tau_{k,t}$	2.41	11.18	0.53	3.67	3.69
Monetary policy	mp_t	4.14	14.35	3.62	18.39	14.72
Tech. news	ξ_χ	-	17.14	-	-	0.46
Gov . news	ξ_{mp}	-	-	51.20	-	0
Inv. news	ξ_{τ_k}	-	-	-	44.78	49.25

In order to identify the economic significance of news shocks in aggregate fluctuations I examine the proportion of variation in output growth due to different exogenous disturbances. The variance decompositions presented below corresponds to an orthogonal

[28] This is defined as $\ln\left(\frac{P(data|Model_i)}{P(data|Model_1)}\right)$

decomposition of the unconditional variance of the model[29]. Model #1, the 'conventional' model that omits the presence of news shocks, admits results broadly in line with the literature. Here the labor supply shock dominates the low-frequency properties of the data, the frequency range emphasized by the unconditional variance decomposition. To a much lesser extent the government spending and demand disturbances also appear as important components of aggregate fluctuations. These results change dramatically with the introduction of news shocks.

Model #2, the model incoporating only news shocks arriving via the stochastic process for neutral technology, attributes much less variation in output growth at the lower frequency to labor supply shocks. Instead, the model finds that investment-specific technology and monetary policy shocks drive a larger proportion of the variation in the data. Furthermore, the neutral technology news shocks are also deemed to have quantitatively large impact on the output growth, though found to be much less than the 50% found by Beaudry and Portier[30]. Model #3, the model incorporating only news shocks arriving via the stochastic process for government spending, is found to have the worst fit among the various news shock specifications as shown by the marginal likelihood. Nonetheless, this version attributes 50% of the variation in output to government spending news shocks.

Model #4, the model incorporating only news shocks arriving via the stochastic process for investment-specific technology, exhibits some of the most interesting results. This version attributes 45% of the variation in output to news shocks, which, in combination with the superior fit shown by the marginal likelihood, suggests that this type of

[29]The appendix contains results for the 1-step ahead variance decomposition of output growth.

[30]Beaudry and Portier use a different methodology and stock returns to determine that importance of these types of news shocks in aggregate fluctuations

news shock may provide the most promise for future research. The results for Model #5 support this claim by showing that though all news shocks appear to be statistically significant it is the investment-specific news shocks that are economically significant.

The table reveals another interesting observation that relates to optimal monetary policy. Note that monetary policy shocks in Model #1 are found to be relatively unimportant for explaining the low frequency movements in output growth. In conventional models optimal monetary policy is often characterized by an inflation-targetting Taylor rule. These models, when fit to the data, often imply that monetary policy shocks have small but significant effects on output growth. The table suggests that in the presence of news shocks monetary policy disturbances may have a much larger impact on economic dynamics. Application of these types of monetary policy rules in the presence of news shocks may in fact be sub-optimal and lead to variations in output growth that are larger than desired. Christiano, Ilut, Motto and Rostagno (2006) examine optimal monetary policy in the presence of news shocks and find that inflation-targetting may in fact be sub-optimal. The results presented here show that their work is very relevant since the data appears to contain these types of distortions.

Figure #4 provides a different perspective on output growth for Model #5. This plot decomposes the time dimension of output growth into contemporaneous and news disturbances. The top plot in figure 4 shows Model #5's smoothed estimate of actual output growth obtained via the 2-sided Kalman filter is close to observed output growth. The middle plot shows that contemporaneous disturbances, the 'conventional' disturbances considered in economic models, are responsible for a good proportion of the variation in output growth. In particular, Model #5 does not attribute the 1990 and 2000 recessions

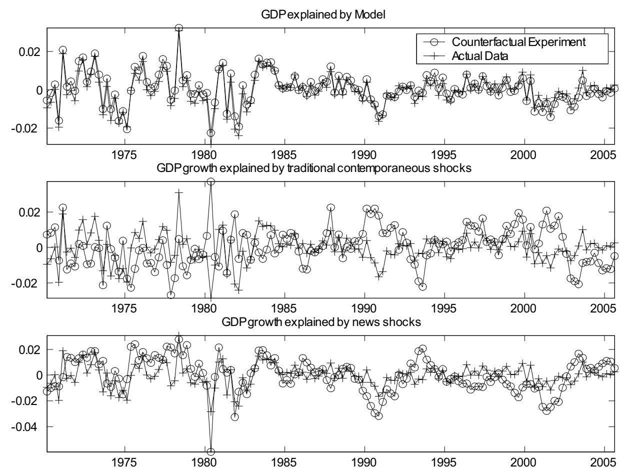


Figure 1.4. Orthogonal decomposition of output growth according to Model #5, the model that includes all three types of news shocks.

to contemporaneous shocks. Instead, as shown in the bottom plot, the decline in output in these periods are attributed to news shocks. This finding is in line with the work of Beaudry and Portier (2003) who advocate Pigou's theory of the business cycle as a theory of recessions. Furthermore, Christiano et al. (2006) provide support for this idea in their examination of the late 1990's boom/bust cycle but emphasize the importance of neutral-technology news shocks for understanding this episode. Model #5 suggests that investment-specific technology shocks are behind the downturns in question. The appendix contains plots that further decompose the 'contemporaneous' and 'news' shock series into their constituent disturbances.

Another intersting exersize explores the model's implications for the term structure of interest rates. Three 'latent' factors characterize much of the yield curve literature and are commonly referred to as slope, level and curvature. These factors have been shown to explain the cross-sectional behavior of bond prices and are closely related to macroeconomic activity[31]. To examine the relationship between these factors and the model's fundamental economic shocks I present another table of variance decompositions. I take the average of the short-rate, 1-year rate and 5-year rate as a proxy for the level factor, the difference between the 5-year rate and the short-rate as a proxy for the slope factor, and the difference between the sum of the short-rate and 5-year rate and the 2-year rate as a proxy for curvature.

Unconditional Variance Decomposition of Yield Curve Factors

Model #5, Observation vector excludes stock returns

Yield Curve Factor		Level	Slope	Curvature
Neutral tech.	χ_t	0.31	0.19	0.29
Government	g_t	0.04	0.26	0.03
Demand	$\varphi_{b,t}$	71.02	10.79	70.91
Labor supply	$\varphi_{0,t}$	1.59	16.57	1.25
Investment	$\tau_{k,t}$	7.66	3.53	7.87
Monetary policy	mp_t	0.83	28.50	1.52
Tech. news	ξ_χ	0.07	0.23	0.06
Gov. news	ξ_{mp}	0	0	0
Inv. news	ξ_{τ_k}	18.46	39.94	18.04

Examination of the table shows that the low-frequency variation in the level factor is primarily driven by demand shocks, a finding consistent with that of Evans and Marshall (1998). Investment-specific news shocks are found to be important components of the variation in all three yield curve factors. While only explaining about a fifth of the

[31] see for example Ang and Piazzesi 2003

variation in the level factor these shocks explain around 40% of the movement in the slope factor.

CHAPTER 2

On the Presence of News in a Hybrid Model of the Yield Curve

This paper presents a 'hybrid' model of the yield curve that incorporates the cross-equation restrictions of a structural dynamic stochastic general equilibrium model (DSGE) into an affine factor model of interest rates. The approach assumes that yield curve factor dynamics are described by a vector autoregression (VAR). Factor models are often used in asset pricing because they impose no-arbitrage conditions without having to satisfy the other restrictive conditions of an equlibrium. The lack of equilibrium information means that these models are often highly parameterized and estimation may result in regions of the parameter space that fit well statistically but are far from those implied by economics. A naturual solution to this problem is to incorporate information from an equilibrium model in the estimation of the factor model. Following the work of Del Negro and Schorfheide (2006, DS) I construct a prior for the VAR parameters from a DSGE model by making a VAR approximation to the DSGE model's dynamics. These authors show that the cross equation restriction implied by DSGE models can improve the forecasting performance of VAR's.

The goal of this paper is twofold, to modify an affine yield curve factor model estimation to include equilibrium information as well as to test for the presence of 'news' in interest rates. The effects and quantitative significance of news shocks has been the subject of a recent strand of literature whose roots lie in the classic work of Arthur Pigou (1926). There it is proposed that business cycle fluctuations are the result of agents' inability to properly forecast future economic activity. Agents dynamically respond to expected future disturbances and the failure of these shocks to materialize can lead to fluctuations similar to those characterizing modern economies.

Most of the economic research to date examines various successes and shortcoming of existing structural models in the presence of news shocks. This has led some researchers to propose new frameworks whose properties are consistent with the presence of these disturbances. Beaudry and Portier (2006) quantify the significance of news shocks that arrive via a unit root stochastic process for neutral technology identified using certain restrictions in a VAR setting. Davis (2007) expands on these results by testing for alternative types of news shocks using a large-scale New Keynesian DSGE model. He finds that they are quantitatively and economically significant, the most important being news about future innovations to investment-specific technology. Davis (2007) claims that interest rates are likely to carry information about news shocks expected to arrive in the future, which he uses in his estimation under the assumption that the expectations hypothesis holds[1]. A large literature exists that rejects the expectations hypothesis suggesting that risk premia are time-varying. The model presented here allows for time-varying risk premia in testing for news shocks. Bayesian model comparison determines the types of news shocks which have the most promising implications. The main empirical findings are that news shocks are critical components of aggregate fluctuations and they are identified by interest rates.

[1]The linear approximation to the solution of the DSGE model used in Davis (2007) implies the expectations hypothesis

Beaudry and Portier (2006) argue that stock prices contain information about these disturbances. The results presented herein contribute to this line of research by documenting the 'news' properties of the term structure in an affine factor model framework. Fundamentally, this is due to the Expectations Hypothesis (EH) which, loosely stated, implies that the interest rate payed on a long-term zero coupon bond is equal to the expected return on a strategy that repeatedly rolls over short-term bonds. While there is some evidence of time variation in term premiums, most agree that the EH should account for most of the variation in the long rates (see Bekaert, Cho and Moreno (2007) for a discussion). In modelling interest rates I work within the no-arbitrage affine macro-factor model framework that allows for time-varying risk premia. In this vein the approach is most closely related to the work of Cochrane and Piazzesi (2006) and Ang and Piazzesi (2003)[2].

To provide motivation for the promising identification properties of news shocks in the term structure of interest rates I will first present a simple economic model based on the work of Long and Plosser (1981)[3]. Suppose that a representative agent has preferences of streams of consumption given by:

$$E_0 \sum_{t=0}^{\infty} \beta^t \ln c_t \tag{2.1}$$

[2] The use of macro-factors is motivated by the work of Ang and Piazzesi (2003). These authors document the close relationship between interest rates and macroeconomic variables. The model of risk premia is based on the findings of Cochrane and Piazzesi (2006).

[3] Davis (2007) uses this example to motivate the promising identification properties of interest rates

where c_t is a representative basket of consumption goods. The agent maximizes 2.1 subject to the following budget constraint and exogenous stochastic processes for technology:

$$\begin{aligned} c_t + k_{t+1} &\leq z_t k_t^{\alpha} \equiv y_t \\ \ln z_{t+1} &= \rho \ln z_t + \eta_t^{t+1} + \varepsilon_{t+1} \end{aligned}$$

where η_t^{t+1} is interpreted to be a 'news' shock, an innovation to the stochastic process for technology expected at time t to arrive at time $t+1$. The disturbance ε_{t+1} is orthogonal to time t information, consistent with the traditional interpretation of a technology shock. It is straightforward to show that the efficient allocations in the model satisfy

$$\begin{aligned} y_t &= z_t k_t^{\alpha} \\ k_{t+1} &= \alpha\beta y_t \\ c_t &= (1-\alpha\beta)\, y_t \end{aligned}$$

The econometrician who observes data on these allocations would not be able to identify news shocks[4]. Now note that 1-period interest rate does identify the news shocks via the term c_{t+1}:

$$R_t = \frac{1}{\beta c_t} E_t \left[\frac{1}{c_{t+1}} \right]^{-1}$$

This stylized example shows that if news shocks are present in macroeconomic aggregates they are likely to be identified using interest rates, an obvious consequence their information for agents' expectations.

[4]This model implies that news shocks are completely insignificant as well. The idea behind this example is to highlight the theoretical identification properties of interest rates, not the economic significance of the news shocks

Empirical evidence for the news content of the interest rates may be found in the well documented leading indicator property of the slope of the yield curve[5]. Under the expectations hypothesis the implied interest rate on a long term bond is simply equal to the average expected rate payed by short term bonds over that same period plus a constant risk premium. If one were to assume that the hypothesis held a large amount of information about future economic activity could be simply read off of the yield curve. A downward sloping curve indicates that the short term interest rate will drop in the future, consistent with a loose monetary policy run during a recession. Furthermore, the term structure also contains information about the timing and magnitude of these expected interest rate movements, a signal of agents' broader economic expectations.

2.1. An Affine Factor Model of Interest Rates

Before setting up the model it is useful to define the notation. Denote log prices by

$$p_t^{(n)} = \text{log of price of } n - year \text{ discount bond}$$

The log yield is defined as

$$y_t^{(n)} = -\frac{1}{n} p_t^{(n)}$$

The log forward rate at time t for loans between time $t + n - 1$ and $t + n$ is given by

$$f_t^{(n)} = p_t^{(n-1)} - p_t^{(n)}$$

[5]See, for example, Stock and Watson (????)

and the log holding period return from buying an n-year bond at time t and selling it as an $n-1$ year bond at time $t+1$ is

$$r_{t+1}^{(n)} = p_{t+1}^{(n-1)} - p_t^{(n)}$$

The excess log return is defined as

$$rx_{t+1}^{(n)} = r_{t+1}^{(n)} - y_t^{(1)}$$

In constructing the model I follow the discrete-time homoskedastic exponential-affine structure from Cochrane and Piazzesi (2006). Let X_t denote the vector of factors used in the model at time t. The factor dynamics are assumed to be described by a general VAR process expressed (in companion form) as

$$X_{t+1} = \mu + \phi X_t + v_{t+1}$$

Note that this VAR specification encompasses the case where lags are included in the factor dynamics. VAR's have a long and rich history in macroeconomics and are often used to describe the dynamics of yield curve factor in the discrete-time affine setting. The log nominal discount factor is a linear function of these factors,

$$M_{t+1} = \exp\left(-\delta_0 - \delta_1' X_t - \frac{1}{2}\lambda_t' V \lambda_t - \lambda_t' v_{t+1}\right)$$

and the short rate process and market prices of risk are assumed to be affine in the factors[6]

$$\begin{aligned} r_t &= \delta_0 + \delta_1' X_t = -p_t^{(1)} \\ \lambda_t &= \lambda_0 + \lambda_1 X_t \end{aligned}$$

Following the empirical evidence provided in Cochrane and Piazzesi (2006) the vector λ_0 is restricted to be zero except in the row corresponding to the level factor. The matrix λ_1 contains one nonzero entry corresponding to the row of the level factor and the column of the excess return factor. These restrictions greatly reduce the number of parameters that need to be estimated in the affine model.

Bond prices can be found by taking expectations

$$p_t^{(n)} = \log E_t \left(M_{t+1} M_{t+2} ... M_{t+n} \right)$$

These expectations can be calculated recursively as follows:

$$\begin{aligned} p_t^{(1)} &= \log E_t \left(M_{t+1} \right) = -\delta_0 - \delta_1' X_t \\ p_t^{(n)} &= \log E_t \left(M_{t+1} \exp \left(p_{t+1}^{(n-1)} \right) \right) \end{aligned}$$

It can be shown[7] that the log bond prices are affine functions of the state variable,

$$p_t^{(n)} = A_n + B_n' X_t$$

[6] In the case where X_t includes lagged variables the vector δ_1 is restricted to load only on current factors

[7] see the web appendix to Cochrane and Piazzesi (2005) for details

The coefficients A_n and B_n can be computed recursively:

$$
\begin{aligned}
A_0 &= 0, \quad B_0 = 0 \\
B_{n+1'} &= -\delta_1' + B_n' \phi^* \\
A_{n+1} &= -\delta_0 + A_n + B_n' \mu^* + \frac{1}{2} B_n' V B_n
\end{aligned}
$$

where ϕ^* and μ^* are defined as

$$
\begin{aligned}
\phi^* &\equiv \phi - V\lambda_1 \\
\mu^* &\equiv \mu - V\lambda_0
\end{aligned}
$$

The coefficients for the forward rates are simply given by

$$
\begin{aligned}
f_t^{(n)} &= A_n^f + B_n^{f\prime} X_t \\
&= (A_{n-1} - A_n) + \left(B_{n-1}' - B_n'\right) X_t
\end{aligned} \tag{2.2}
$$

2.2. Structural Prior and a Framework for 'News'

A recent vein of macroeconomic research has conjectured that agents' inability to accurately forecast future events may contribute to business cycle fluctuations. Interest rates are closely related to agents' expectations about future macroeconomic activity and may be useful for asessing the validity of this hypothesis. Here I describe a model of 'news,' shocks expected to arrive in the future though not yet realized, into the VAR framework described above.

2.2.1. A Stochastic Representation of News

As an introduction to news shocks consider the general $AR(1)$ process augmented to include the disturbances η:

$$s_t = \rho s_{t-1} + \varepsilon_t + \eta^t_{t-1} + \eta^t_{t-2} + \cdots + \eta^t_{t-p}$$

When $\eta^t_{t-1} = \eta^t_{t-2} = \cdots = 0$ this is the familiar $AR(1)$ process. Here it is assumed that the vector η^t_{t-j} is realized at time $t-j$ and represents 'news' about s_t. The superscipt on the variable indicates the date of s_t that the news is relevant for, while the subscript indicates the date that the news is realized. Consider, the pure moving average representation of s_t

$$\begin{aligned} s_t \;\; = \;\; & \left[\varepsilon_t + \eta^t_{t-1} + \eta^t_{t-2} + \cdots + \eta^t_{t-p}\right] \\ & +\rho\left[\varepsilon_{t-1} + \eta^{t-1}_{t-2} + \eta^{t-1}_{t-3} + \cdots + \eta^{t-1}_{t-p-1}\right] \\ & +\rho^2\left[\varepsilon_{t-2} + \eta^{t-2}_{t-3} + \eta^{t-2}_{t-4} + \cdots + \eta^{t-2}_{t-p-2}\right] \\ & +\cdots \end{aligned}$$

Note that the objects in square brackets at different dates are not correlated with each other, but each has the same variance. As a result, s_t has a first order autoregressive representation regardless of the variances of the news shocks. The number of signals in s_t is not identified from observations on s_t alone. However, the cross equation restriction implied by the no-arbitrage pricing framework can deliver identification.

We now set this process up in state-space/observer form. I begin with the simple case of $p = 2$

$$s_t = \rho s_{t-1} + \varepsilon_t + \eta^t_{t-1} + \eta^t_{t-2} \tag{2.3}$$

It is useful to set up some auxiliary variables, u^{t+1}_{t-1} and u^t_{t-2}:

$$\begin{bmatrix} s_t \\ u^{t+2}_t \\ u^{t+1}_t \end{bmatrix} = \begin{bmatrix} \rho & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} s_{t-1} \\ u^{t+1}_{t-1} \\ u^t_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ \eta^{t+2}_t \\ \eta^{t+1}_t \end{bmatrix} \tag{2.4}$$

This auxillary form for the news process will be used in estimating the model. It is easy to confirm that this is the same as 2.3. Write the first equation:

$$s_t = \rho s_{t-1} + u^t_{t-1} + \varepsilon_t$$

To determine u^t_{t-1} evaluate 2.4 at the previous date

$$\begin{aligned} u^{t+1}_{t-1} &= \eta^{t+1}_{t-1} \\ u^t_{t-1} &= u^t_{t-2} + \eta^t_{t-1} \end{aligned}$$

and the second of the above two expressions indicates that we must evaluate 2.4 at an even earlier date

$$u^t_{t-2} = \eta^t_{t-2}$$

and putting these together 2.3 is confirmed. The case for $p > 2$ follows analagously.

2.2.2. A DSGE Model with News

I now embed the 'news' framework described above into a simple economic model. The model is based on that of Bekaert, Cho and Moreno (2005)[8] and is described in further detail in the appendix. Optimization on behalf of the agents in the model results in four log-linear equations , which, along with the stochastic processes for the exogenous shocks, characterize the equilibrium of the model. The optimality conditions are:

$$\pi_t = \delta E_t\left[\pi_{t+1}\right] + (1-\delta)\pi_{t-1} + \kappa(y_t - y_t^n) + \xi_t^{as} \tag{2.5}$$

$$y_t = \alpha_{is} + \mu E_t\left[y_{t+1}\right] + (1-\mu)y_{t-1} - \phi\left(r_t - E_t\left[\pi_{t+1}\right]\right) + \xi_t^{is} \tag{2.6}$$

$$r_t = \alpha_{mp} + \rho r_{t-1} + (1-\rho)\left[\beta\left(E_t\left[\pi_{t+1}\right] - \pi_t^*\right) + \gamma\left(y_t - y_t^n\right)\right] + \varepsilon_t^{mp} \tag{2.7}$$

$$\pi_t^* = \varphi_1 \pi_{t+1}^* + \varphi_2 \pi_{t-1}^* + \varphi_3 \pi_t + \varepsilon_t^{\pi^*} \tag{2.8}$$

And the stochastic processes for the shocks are:

$$y_t^n = \alpha_{y^n} + \lambda y_{t-1}^n + \varepsilon_t^{y^n} + \eta_t^{y^n} \tag{2.9}$$

$$\xi_t^{as} = \alpha_{is} + \rho^{as}\xi_{t-1}^{as} + \varepsilon_t^{as} + \eta_t^{as} \tag{2.10}$$

$$\xi_t^{is} = \alpha_{is} + \rho^{is}\xi_{t-1}^{is} + \varepsilon_t^{is} + \eta_t^{is} \tag{2.11}$$

[8] These authors argue that this simple model does a good job at replicating the yield curve under the expectations hypothesis

Note that the stochastic process for y_t^n includes the news structure described above. News shocks, η_t, can be incorporated along any of the stochastic processes. Here it should be understood that the terms η^{y^n}, η_t^{as} and η_t^{is} correspond to a sum of news shocks, each corresponding to a different horizon:

$$\eta_t^L = \eta_{t-1}^t + \eta_{t-2}^t + \cdots + \eta_{t-p}^t; \; L = y^n, as \text{ or } is$$

In this context the news shocks mean that agents' forecast innovations to future values of y_t^n, ξ_t^{as} or ξ_t^{is} before they actually occur. Note that this paper presents results where news shocks arrive via only one stochastic process at a time. Thus, when estimating news shocks that arrive via the stochastic process for y^n, the terms η_t^{as} and η_t^{is} are set to be zero.

Linearly approximating the dynamic rational expectations solution[9] results in the following state-space system

$$\begin{aligned} S_t &= F(\theta) S_{t-1} + Q(\theta) \varepsilon_t \\ F_t &= H(\theta) S_t \end{aligned} \tag{2.12}$$

where F_t, here the observables, correspond to the factors used in the affine model. The mappings from S_t to F_t are provided in the appendix. Note that the linearity of the solution makes computation of moments straightforward.

[9]This can be done using the methods of Klein (2000)

2.2.3. Prior for VAR

The prior used for the VAR in the interest rate model is based on the simple DSGE model described above and stems from the work of Del Negro and Schorfheide (2003). To fix ideas let F_t represent the time t factors in the affine model: the Cochrane and Piazzesi risk factor, a level factor (the average of the 2, 5 and 10 year yields), a slope factor (the difference between the 10 and 2 year yields), year-on-year log real GDP growth and the difference between the level factor and the year-on-year inflation as measured by the Personal Consumption Expenditure deflator[10]. The VAR describing the factor dynamics is

$$\begin{aligned} F_t &= \mu + \phi_1 F_{t-1} + \cdots + \phi_k F_{t-k} + v_t \\ &\quad + \gamma_1 u_{t-1}^{t} + \gamma_2 u_{t-1}^{t+1} + \cdots + \gamma_{p-1} u_{t-1}^{t+p-2} + \eta_{t-1}^{t+p-1} \end{aligned}$$

where u_{t-1}^{t} represents the accumulation of news shocks as discussed above, v_t represents the current innovation to the factors that is independent of past factors and news, assumed here to be distributed Gaussian. Note that the news shocks, in the structural model occuring via the stochastic process for y_t^n, arrive in the future but, due to wealth and substitution effects, may be correlated with the factors dated prior to their arrival. For this reason the vectors $\gamma_2, ..., \gamma_{p-1}$ may be nonzero in the factor VAR.

[10]I follow the 'difference spread' specification of Cochrane and Piazzesi (2006). In doing so the only factor that contains information about the level of interest rates is the 'level' factor. Since inflation contains similar informaiton I instead use the difference between inflation and the 'level' factor as a factor.

Here X_t, the vector used in describing the VAR process in companion form, can be written as

$$X_t' = \left[F_t', \;\; F_{t-1}', \;\; \cdots \;\; F_{t-k+1}', \;\; \eta_t^{t+p\prime}, \;\; u_t^{t+p-1\prime}, \;\; \ldots, \;\; u_t^{t+1\prime} \right]' \tag{2.13}$$

In estimating the model the news shocks are 'latent' variables recovered using the Kalman Filter. Note that the VAR can thus be written as

$$\widehat{F}_t = \widehat{X}_{t-1}' \Phi + v_t$$

where the 'hats' represent deviations from the steady state.

To motivate the prior distribution for the VAR parameters it is useful to treat these latent variables as though they are observed. Letting $\Sigma = E\left[v_t v_t'\right]$, the conditional probability of observing v_t is proportional to

$$p(\widehat{F}_t|\Phi, \Sigma) \propto |\Sigma_\eta|^{-1/2} \exp\{-\frac{1}{2}\left(\widehat{F}_t' - \widehat{X}_{t-1}'\Phi\right)' \Sigma^{-1} \left(\widehat{F}_t' - \widehat{X}_{t-1}'\Phi\right)\}$$

The application of the trace operator to the conditional probability above implies

$$tr\left(\left(\widehat{F}_t' - \widehat{X}_{t-1}'\Phi\right)' \Sigma^{-1} \left(\widehat{F}_t' - \widehat{X}_{t-1}'\Phi\right)\right) = tr\left(\Sigma_\eta^{-1}\left(\widehat{F}_t' - \widehat{X}_{t-1}'\Phi\right)'\left(\widehat{F}_t' - \widehat{X}_{t-1}'\Phi\right)\right)$$

To see how the prior is constructed it is useful to write the likelihood using a slightly different notation

$$F = \begin{bmatrix} \widehat{F}_1' \\ \vdots \\ \widehat{F}_T' \end{bmatrix}, \quad X = \begin{bmatrix} \widehat{X}_0' \\ \vdots \\ \widehat{X}_{T-1}' \end{bmatrix}$$

Then the likelihood of this model, again assuming the news shocks are observed, can be written as

$$
\begin{aligned}
p(F|\Phi,\Sigma,F_0) &= \prod_{t=1}^{T} p(\widehat{F}_t|\widehat{F}^{t-1},F_0,\Phi,\Sigma) \\
&\propto |\Sigma|^{-T/2}\exp\left\{-\frac{1}{2}\sum_{t=1}^{T} tr\left(\Sigma^{-1}\left(\widehat{F}_t'-\widehat{X}_{t-1}'\Phi\right)\prime\left(\widehat{F}_t'-\widehat{X}_{t-1}'\Phi\right)\right)\right\} \\
&\propto |\Sigma|^{-T/2}\exp\left\{-\frac{1}{2}tr\left(\sum_{t=1}^{T}\Sigma^{-1}\left(\widehat{F}_t'-\widehat{X}_{t-1}'\Phi\right)'\left(\widehat{F}_t'-\widehat{X}_{t-1}'\Phi\right)\right)\right\} \\
&\propto |\Sigma|^{-T/2}\exp\left\{-\frac{1}{2}tr\left(\Sigma_\eta^{-1}\left(F-X\Phi\right)\prime\left(F-X\Phi\right)\right)\right\} \\
p(F|\Phi,\Sigma,F_0) &\propto |\Sigma|^{-T/2}\exp\left\{-\frac{1}{2}tr\left(\Sigma^{-1}\left(F'F-\Phi'X'F-F'X\Phi+\Phi'F'F\Phi\right)\right)\right\}
\end{aligned}
$$

Define the following moments the DSGE model

$$
\begin{aligned}
\Gamma_{FF}(\theta) &= E_\theta\left[\widehat{F}_t\widehat{F}_t'\right] \\
\Gamma_{XX}(\theta) &= E_\theta\left[\widehat{X}_t\widehat{X}_t'\right] \\
\Gamma_{XF}(\theta) &= E_\theta\left[\widehat{X}_{t-1}\widehat{F}_t'\right]
\end{aligned}
$$

these moments can be computed quickly using the linear state-space system 2.12. From these moments we can find the 'best' p-lag VAR approximation to the DSGE model

$$
\begin{aligned}
\Phi^*(\theta) &= \Gamma_{XX}^{-1}(\theta)\,\Gamma_{XF}(\theta) \\
\Sigma^*(\theta) &= \Gamma_{FF}(\theta)-\Gamma_{XF}'(\theta)\,\Gamma_{XX}^{-1}(\theta)\,\Gamma_{XF}(\theta)
\end{aligned}
$$

Following Del Negro and Schorfheide I will derive a prior that is proportional to the expected likelihood ratio under the assumption that the latent variables are observable, expectations being taken with respect to the model. Suppose that we generate a sample of λT observations from the DSGE model, collected in matrices F_* and X_*. Write $\Phi^{\Delta} = \Phi^*(\theta) - \Phi$ for the discrepancy between the yield curve and the DSGE model's factor dynamics. Then the likelihood ratio is

$$\ln\left[\frac{L\left(\Phi^*(\theta)+\Phi^{\Delta},\Sigma^*(\theta)|F_*,X_*\right)}{L\left(\Phi^*(\theta),\Sigma^*(\theta)|F_*,X_*\right)}\right] = -\frac{1}{2}tr[\Sigma_{\eta}^{*-1}(\Phi^{\Delta}X_*'X_*\Phi^{\Delta} + 2\Phi^{*\prime}X_*'X_*\Phi^{\Delta} - 2\left(\Phi^* - \Phi^{\Delta}\right)X_*'F_* + 2\Phi^{*\prime}X_*'F_*]$$

and taking expectations with respect to the model gives (note that $E_\theta\left[\Phi^{\Delta}\right] = 0$)

$$E_\theta\left[\ln\left[\frac{L\left(\Phi^*(\theta)+\Phi^{\Delta},\Sigma^*(\theta)|F_*,X_*\right)}{L\left(\Phi^*(\theta),\Sigma^*(\theta)|F_*,X_*\right)}\right]\right] = \frac{1}{2}tr\left[\Sigma_{\eta}^{*-1}(\Phi^{\Delta}\lambda T\Gamma_{XX}\Phi^{\Delta})\right]$$

The prior density is designed to be proportional to the expected log likelihood ratio

$$P\left(\Phi^{\Delta}|\Sigma^*(\theta)\right) \propto \exp\left\{-\frac{1}{2}tr\left[\Sigma^{*-1}(\Phi^{\Delta}\lambda T\Gamma_{XX}\Phi^{\Delta})\right]\right\}$$

where the hyperparameter λ is a number between 0 and ∞. This parameter controls the tightness of the prior, ie. the importance of the prior in the estimation. A value of 0 corresponds to a flat prior or 'uninformative' prior consistent with classical estimation. A value of ∞ corresponds to the VAR generated by the DSGE model. Conditional on θ, a prior for the VAR parameters that accomplishes this is

$$\Sigma|\theta,\lambda \sim IW\left(\lambda T\Sigma^*(\theta), \lambda T - k\right)$$

$$\Phi|\Sigma,\theta,\lambda \sim N\left(\Phi^*(\theta), \frac{1}{\lambda T}\left[\Sigma^{-1}\otimes\Gamma_{XX}(\theta)\right]^{-1}\right)$$

The appendix further describes these priors in detail as well as the computational implementation.

2.3. Estimation

Now I describe the estimation of the affine factor model of interest rates. The model is estimated using data on the 5 factors as well as continuously compounded forward rates on bonds of maturing in one to ten years. Again, these 5 factors are: the Cochrane and Piazzesi risk factor (defined in the appendix), a level factor (the average of the 2, 5 and 10 year yields), a slope factor (the difference between the 10 and 2 year yields), year-on-year log real GDP growth and the difference between the level factor and the year-on-year inflation as measured by the Personal Consumption Expenditure deflator. Collectively the factors and forward rates form 15 time-series on which the model is estiamted using the Kalman Filter applied to the following state-space system[11]:

$$\begin{aligned} X_t &= \mu + \varphi X_{t-1} + v_t \\ \begin{bmatrix} F_t \\ f_t \end{bmatrix} &= \begin{bmatrix} 0 \\ A_f \end{bmatrix} + \begin{bmatrix} H_F \\ B_f \end{bmatrix} X_t + \begin{bmatrix} 0 \\ \psi_t \end{bmatrix} \end{aligned} \tag{2.14}$$

where F_t represents a vector of the 5 time t factors used in the affine model, X_t is the vector given in 2.13, f_t is a vector of the 10 forward rates at time t, A_f and B_f are the mappings from the factors to the forward rates obtained from 2.2, H_F is a choice matrix that picks off only the F_t vector embedded in X_t and ψ_t is a vector of measurement error fixed to have variance equal to 5% of the historical variance of the forward rate. Recall that the matrices A_f and B_f include the market prices of risk which are restricted according to

[11]see Hamilton (1994) chapter 13 for details

the specification of Cochrane and Piazzesi (2006). Furthermore, I fix the measurement error at a small number in order to maintain parsimony and focus the estimation on economically interesting regions of the parameter space[12].

Maximum Likelihood estimation can be achieved by using the Kalman Filter. Evaluating the parameters at the prior and multiplying this by the likelihood results in a Bayesian Posterior distribution. Both of these estimation techniques will be utilized to assess the fit of various versions of the affine model presented herein. The priors are placed over the VAR parameters, as discussed in the previous section, as well as over the parameters of the underlying DSGE model.

[12]Chen and Scott (????) and Ang and Piazzesi (200?) use measurement error in their estimation. Fixing the variance of the measurement error at a small number guides the estimation toward economically interesting regions of the parameter space where the factors explain the majority of the variance of the forward rates.

Parameter	Description	Distribution(Mean,Stdev)
y	Steady state log output	Fixed, =0
r	Steady state short rate	Fixed, =.05/4
π	Steady state Inflation	Fixed, =.03/4
δ	Phillips coeff: $E_t\pi_{t+1}$	Beta(0.6,0.1)
κ	Phillips coeff: $(y_t - y_t^n)$	Unif(0,0.3)
σ	Risk Aversion	Normal(2,1)[13]
η	Habit Formation	Normal(3,1)
ρ	I.R. Persistence	Beta(0.8,0.1)[14]
β	Taylor Rule Infl	Normal(1.5,0.2)[15]
γ	Taylor Rule Out. Gap	Beta(0.08,0.03)
λ	Persistence Nat. Rate	Beta(0.9,0.1)
w	Infl Target Param	Beta(0.85,0.1)
d	Long Run Infl Exp Param	Beta(0.85,0.1)
ρ_{as}	AS shock Persistence	Unif(0,1)
ρ_{is}	IS shock Persistence	Unif(0,1)
σ_{as}	Stdev AS shock	IG(0.0125,0.005)
σ_{is}	Stdev IS shock	IG(0.0067,0.005)
σ_{mp}	Stdev MP shock	IG(0.02,0.005)
σ_{y^n}	Stdev Nat Rate shock	IG(0.015,0.005)
σ_{π^*}	Stdev Infl Target shock	IG(0.0075,0.005)

[13] This Normal distribution is truncated to only have mass above 1

[14] Truncated to have mass only on positive numbers

[15] This prior is truncated to have mass above 1

2.4. Results

Before proceeding to the results from the Bayesian estimation it is useful to conduct a likelihood ratio test for the presence of news shocks from maximum likelihood estimation. The news shock version of the yield curve model, estimated under the assumption that news shocks arrive up to 4 periods in advance, has an additional 48 paramaters that need to be estimated[16]. Maximizing the log-likelihood of the model[17] results in a value of 11,624.45 while the restricted version of the model, where news shock parameters are all zero, results in a log-likelihood value of 11,084.61. The 1% cutoff value for a $\chi^2(48)$ variable is 73.69. The relevant value of this statistic for the hypothesis that the news shocks are not present is $2 \times (11,624.45 - 11,084.61) = 1,079.68$. Thus the hypothesis that news shocks are all zero is rejected at the 1% significance level.

The likelihood ratio test cannot reject the presence of news shocks, which, in the model presented here, are simply unobserved latent variables that follow a particular type of stochastic process. This result is consistent with the literature on the term-structure of interest rates where latent variables are needed to explain various yield curve shapes[18]. Previous results have emphasized the importance of three 'unobservable' factors commonly referred to as slope, level and curvature. Here the latent variables are factors extracted from the yield curve that are not encompassed by conventional slope and level factors, as proxies for these are included directly as observable factors in the estimation. Instead, the dynamics of the VAR are restricted so that the latent variables can loosely be interpreted

[16]There are 4 additional parameters in δ_1 for corresponding to each of the news shocks, 24 additional paramaters in the variance-covariance matrix (4 news x 5 observed factors = 20 covariances) as well as 4 variances for each of the news innovaitons, and lastly 4 news x 5 observed factors = 20 parameters added to the VAR process

[17]Note that the VAR for the observable factors includes 4 lags

[18]see, for example, Ang and Piazzesi (2003)

as information in interest rates that help to forecast the observable VAR factors. A useful exercise to verify this conjecture, though not pursued here due to computational burden, is to compare out-of-sample forecasting performance of a news and non-news versions of the model.

While this classical maximum likelihood estimation is useful for validating the presence of the news shock structure presented above it does not shed light on the type of news important for understanding interest rates and macroeconomic dynamics. The reason is that the different types of news shocks have different implications in both the VAR and residual variance-covariance matrix. For example, news that the natural rate of output is going to increase in the future likely has a much different effect on macroeconomic allocations from news that the aggregate demand will shift at some point in the future. The Bayesian version of the model distinguishes between these various types of news shocks by incorporating information about these disturbances from an underlying structural model. The classical estimates presented above doesn't have these implications and thus the 'news' that it identifies doesn't have a clear economic interpretation.

Bayesian estimation of the affine term structure model provides insight into the economic importance of news shocks. One measure commonly used to distinguish between competing Bayesian models is the marginal likelihood. The marginal likelihood has a simple interpretation as the probability of observing the data conditional on a particular model. It can be computed by integrating over the posterior distribution as follows:

$$P(Data|Model_i) = \int L(\theta^{(i)}|Data, Model_i)p(\theta^{(i)}|Model_i)d\theta^{(i)}$$

where $L(\theta^{(i)}|Data, Model_i)$ is the likelihood, produced here using the Kalman filter, and $p(\theta^{(i)}|Model_i)$ is the obtained by evaluating the prior at the parameters in the vector $\theta^{(i)}$. In practice the integral above is approximated using samples taken from the posterior distribution that are obtained from a Markov-chain-monte-carlo, Gibbs sampling, or similar type of algorithm. The integral is then approximated using Geweke's [**30**] modified harmonic mean estimator[19].

The hyperparameter λ, discussed in the derivation of the prior for the VAR parameters using the DSGE model, scales the importance of the prior distribution in the estimation. Theoretically a value of $\lambda = 0$ results in a flat prior where information from the DSGE model plays no role in the estimation. The prior is well defined for values of $\lambda > \frac{k+n}{T}$, where k equals the dimension of the vector X_t in 2.14 and n is equal to the sum of the number of news shocks and the number of observable factors[20]. A value of $\lambda = \infty$ places all of the weight on the prior distribution and forces the estimation to recover the same VAR parameters generated by the DSGE. Thus, the prior belief that the DSGE accurately describes the factor dynamics corresponds to a high value of λ and the belief that it does a poor job corresponds to a low value of λ. In speaking about the prior it is useful to convert the hyperparameter λ to a weight between 0 and 1, 0 interpreted as

---

[19]This is discussed in the appendix. The marginal likelihood includes a penalty for the number of parameters.
The appendix also includes a detailed description of the computational implementation
[20]For example when 4 news shocks are present $n = 4 + 5 = 9$.
In estimating the model I only allow news shocks to arrive up to 4 periods in advance, have 5 observable factors and consider 4 lags in the VAR. The minimum value of λ that can be considered so that the prior is well defined is: $\lambda_{\min} = \frac{\overbrace{(5*4+4)}^{k}+\overbrace{(5+4)}^{n}}{\underbrace{136}_{T}}$

being 0% weight on the prior and 1 being 100% weight on the prior. A conversion that accomplishes this is $prior\ wt = \frac{\lambda}{1+\lambda}$.

I estimate 4 versions of the model, allowing for no news shocks and news shocks along the stochastic processes for y^n, ξ^{as} and ξ^{is}. The 4 versions are then estimated using 25%, 50% and 75% weights on the VAR parameters generated by the DSGE model. This results in 12 different models which the following table summarizes succinctly.

Log Marginal Likelihood Ratio[21]

DSGE Weight. \ News:	None	y^n	ξ^{as}	ξ^{is}
25%	0	1220	1156	1138
50%	176	1261	1254	1259
75%	-29	**1412**	1146	1346

The largest value for the log marginal likelihood ratio is 1412 corresponding to the model that places 75% weight on the prior and allows for news shocks to arrive up to 4 periods in advance along the stochastic process for y^n. The fact that the largest marginal likelihood value corresponds to the model with the largest prior weight suggests that the DSGE does a good job at describing the dynamics of the factors when news is present. Interestingly the Bayesian estimation strongly prefers models with news over those without news, regardless of the stochastic process by which it arrives. This is similar to the results in Davis (2007) who finds that models incorporating news have much larger marginal likelihood values.

The estimation suggests news shocks arriving along the stochastic process for y^n are the most likely regardless of the weight placed on the prior. I will now describe a

[21] The log marginal likelihood ratio presented here is $\ln\left(\frac{P(Data|Model(wt,news))}{P(Data|Model(25\%,none))}\right)$. Thus, all numbers are normalized by the marginal likelihood of the model with no news and the least weight on the prior.

method for determining the economic significance of news shocks, as well as conventional contemporaneous shocks, that utilizes the identifying properties of the structural DSGE model. To begin recall the state-space system corresponding to the linear solution of the DSGE

$$\begin{aligned} S_t &= F(\theta) S_{t-1} + Q(\theta) \varepsilon_t \\ F_t &= H(\theta) S_t \end{aligned}$$

Here the matrix $Q(\theta)$ identifies the effect of an economic shock at time t on the state S_t. The matrix $H(\theta)$ in turn translates this effect from the state to the observed factor F_t. Thus, the DSGE can be used to identify the effects of fundamental shocks on the factors. Unfortunately the VAR doesn't contain such nice identification properties which has led economists to propose a variety of methods to identify different economic shocks.

The DSGE provides a natural identification scheme for untangling the effects of economic shocks in the VAR. Note that the 1 period forecast error variance-covariance matrix of the factors F_t in the DSGE is given by $V_{dsge}^1 = H(\theta) Q(\theta) Q(\theta)' H(\theta)'$. Define the (unique) lower triangular Cholesky factorization of the matrix V_{dsge}^1 as $\widehat{\Sigma}_{dsge}$ so that $V_{dsge}^1 = \widehat{\Sigma}_{dsge} \widehat{\Sigma}'_{dsge}$. It is straightforward to see that $H(\theta) Q(\theta)$ is simply a particular rotation of $\widehat{\Sigma}_{dsge}$. Simply stated, the orthonormal matrix $R(\theta)$ solves

$$H(\theta) Q(\theta) = \widehat{\Sigma}_{dsge} R(\theta)$$

and $R(\theta) R(\theta)' = I$. This rotation matrix $R(\theta)$ provides identification by showing how to rotate the (unique) Cholesky factor of the 1-period forecast error variance-covariance matrix in such a way as to obtain the same identification as the DSGE. A natural way to achieve identification of the economic shocks in the VAR is to simply apply this rotation in an analagous fashion. To be explicit recall that the VAR for the factors can be expressed in companion form as

$$X_t = \mu + \varphi X_{t-1} + v_t$$

Letting $\widehat{\Sigma}$ be the cholesky factorization of the 1-period forecast error variance-covariance matrix $V^1 = E\left[v_t v_t'\right]$, the matrix used to identify the initial response of the factors to economic impulses is $Q \equiv \widehat{\Sigma}_{var} R(\theta)$. Thus, the first column of Q reveals how a 1-unit change in the first shock of ε_t, here corresponding to a 1-unit change in $\varepsilon_t^{y^n}$, would initially displace the factors in the VAR. For the sake of completeness the identification described here allows the VAR to be expressed as

$$X_t = \mu + \varphi X_{t-1} + Q\varepsilon_t$$

Using this identification method the overall variance of year-on-year output growth, a factor in the VAR, can be decomposed to reveal which economic shocks are most important. The following table provides such a breakdown across prior weights for the 'preferred' model that incorporates news along the stochastic process for y^n

Variance Decomposition of Output Growth			
Shock \ DSGE Weight:	25%	50%	75%[22]
$\varepsilon_t^{y^n}$ (Nat. Rate)	4.61	4.15	**5.62**
ε_t^{as} (AS)	8.07	10.75	**11.65**
ε_t^{is} (IS)	6.67	6.76	**2.25**
ε_t^{mp}(Mon. Pol.)	28.36	49.17	**41.59**
$\varepsilon_t^{\pi^*}$(Infl. Target)	9.89	13.00	**32.10**
$\eta_t^{y^n}$(News)	42.39	16.15	**6.78**

Three types of shocks stand out as driving the majority of the variance of output growth: ε_t^{mp}, $\varepsilon_t^{\pi^*}$ and $\eta_t^{y^n}$. Interestingly the importance of the news shocks declines and the inflation target shock rises as the prior weight increases. The monetary policy shock is a significant component of output growth variation across all DSGE weight specifications. In the preferred model news shocks are not economically significant in the sense that they aren't responsible for a large portion of the variance in output growth. The appendix contains similar variance decompositions of the level, slope and CP risk factors for the case of the preferred model. There it can be seen that news shocks drive a significant amount of the variation in the slope, a result consistent with what Davis (2007).

Risk premia cannot be directly measured but instead must be estimated using various techniques, making them an issue of debate in the modern bond-pricing literature. The models presented here have implications for volatility of these premia and the extent to which the expectations hypothesis fails.

[22]This is the preferred model according to the Marginal Likelihood criterion

Stdev of Risk Premium in 5 Yr Forward Rate (in BPs)		
DSGE Weight	No News	y^n News
0% (MLE)	120.58	21.97
25%	30.70	4.45
50%	23.67	7.90
75%	39.43	6.32

The table shows that in the absence of economic information from the structural model risk premia are significantly more volatile. The linear approximation to the DSGE model's solution implies that the expectations hypothesis holds and thus risk-premia have no time variation. It is important to note that the prior doesn't directly impose the expectations hypothesis from the DSGE but instead imposes that the factor dynamics are consistent with an underlying structural model[23]. Thus, by 'shrinking' the estimation toward a region of the parameter space that is close to an economic model estimated risk premia are found to vary much less.

Incorporating news shocks into the estimation also has interesting implications. The table shows that when the DSGE weight is non-zero and news is included the estimated risk premium exhibits little variation. This result is consistent with Davis (2007) who finds the expectations hypothesis to be a good approximation in the presence of news shocks[24].

[23]The factor dynamics implied by the DSGE may be different when risk premia are incorporated. This difference is likely to be small given the evidence that suggests that linearized DSGE solutions are accurate in postwar data

[24]Davis (2007) finds that measurement error, in his estimation the difference between observed market yields and model yields constructed using the expectations hypothesis, are approximately zero.

CHAPTER 3

Two Flaws in Business Cycle Accounting

Chari, Kehoe and McGrattan (2005) (CKM) argue that a procedure they call Business Cycle Accounting (BCA) is useful for identifying promising directions for model development.[1] The key substantive finding of CKM is that financial frictions like those analyzed by Carlstrom and Fuerst (1997) (CF) and Bernanke, Gertler and Gilchrist (1999) (BGG) are not promising avenues for studying business cycles. Based on our analysis of business cycle data for the US in the 1930s and for the US and 14 other OECD countries in the postwar period, we find that the CKM conclusion is not warranted.

The BCA strategy begins with the standard real business cycle (RBC) model, augmented by introducing four shocks, or 'wedges'. A vector autoregressive representation (VAR) for the wedges is estimated using macroeconomic data on output, consumption, investment and government consumption.[2] The macroeconomic data are assumed to be observed with a small measurement error whose variance is fixed a priori. The fitted wedges have the property that when they are fed simultaneously to the augmented RBC model, the model reproduces the four macroeconomic data series up to the small measurement error. The importance of a particular wedge is determined by feeding it to the model, holding all the other wedges constant, and comparing the resulting model predictions with the data. One of the wedges, the intertemporal wedge, is the shock that enters

[1] This strategy is closely related to that advocated in Parkin (1988), Ingram, Kocherlakota and Savin (1994), Hall (1997), and Mulligan (2002).
[2] The last variable includes government consumption and net exports.

between the intertemporal marginal rate of substitution in consumption and the rate of return on capital. CKM argue that this wedge contributes very little to business cycle fluctuations, for the following two reasons: (i) the wedge accounts for only a small part of the movement in macroeconomic variables during recessions and (ii) the wedge drives consumption and investment in opposite directions, while these two variables display substantial positive comovement over the business cycle. CKM assert that their conclusions are robust to various model perturbations, including the introduction of adjustment costs in investment.

There are two reasons that BCA does not warrant being pessimistic about the usefulness of models of financial frictions such as those proposed in CF or BGG. First, CKM's conclusions are not robust to small changes in the way they implement BCA. For example, when we redo CKM's calculations for the 1982 recession, we reproduce their finding that the intertemporal wedge accounts for essentially no part of the decline in output below trend at the trough of the recession. When we introduce a modest amount of investment adjustment costs, the intertemporal wedge accounts for a substantial 34 percent of the drop in output at the trough of the recession.[3] We then consider an alternative specification of the intertemporal wedge which is at least as plausible as the one CKM work with. CKM define the intertemporal wedge as an ad valorem tax on the price of investment goods. We argue that the CF and BGG models motivate considering an alternative formulation in which the wedge is modeled as a tax on the gross rate of return on capital.

[3] Our adjustment costs are 'modest' in two senses. First, they imply a steady state elasticity of the investment-capital ratio to the price of capital equal to unity. This lies in the middle of the range of empirical estimates reported in the literature. Second, the adjustment cost function has the property that the quantity of resources lost due to investment adjustment costs is small, even in the wake of the enormous decline in investment in the early 1930s (see section **??** below for a detailed discussion).

When we work with this alternative formulation, the intertemporal wedge accounts for 22 percent of the drop in output at the trough of the 1982 recession. But, when we also drop CKM's model of measurement error, that quantity jumps to 52 percent. Notably, the CKM model of measurement error is overwhelmingly rejected in the post war US data. So, BCA actually places a range of 0 to 52 percent on the fraction of the drop in output accounted for by the intertemporal wedge in the 1982 recession. This range is sufficiently wide to comfortably include most views about the importance of the intertemporal wedge.

We show that, at a qualitative level, economic theory predicts the lack of robustness in BCA that we find. The intertemporal wedge associated with different perturbations of the RBC model represent different ways of bundling the fundamental economic shocks to the economy. As a result, the BCA experiment of feeding measured wedges to an RBC model represents fundamentally different economic experiments under alternative specifications of the RBC model. Since the experiments are different, the outcomes are expected to be different too. Our results show that these expected differences are quantitatively large enough to overturn CKM's conclusions.

Second, CKM's analysis ignores that the financial shocks which drive the intertemporal wedge may have spillover effects onto other wedges.[4] It is not possible to determine the magnitude of these effects with BCA, because BCA leaves the fundamental shocks to the economy unidentified. In fact, the VAR for the wedges estimated under BCA is consistent with a wide range of possible spillover patterns. In terms of CKM's conclusion

[4]Recent developments in economic modeling suggest a variety of mechanisms by which these spillover effects can occur. For example, it is known that in models with Calvo-style wage-setting frictions (see, e.g., Erceg, Henderson and Levin (2000)), a shock outside the labor market can trigger what looks like a preference shock for labor, or a 'labor wedge'. Similarly, variable capital utilization can have the effect that a non-technology shock triggers a move in measured TFP, or the 'efficiency wedge'.

(i) above, we show that the financial shocks which drive the intertemporal wedge could account for as much as 70-100 percent of reductions in output in US recessions, including the Great Depression. We obtain the same finding for several other countries in the OECD. Regarding CKM's conclusion (ii), we show that once spillover effects are taken into account, financial shocks which drive the intertemporal wedge can drive consumption and investment in the same direction.

CKM understand that the fundamental economic shocks are not identified under BCA. However, the implications they draw from this observation are very different from the ones we draw. They say, 'Our method is not intended to identify the primitive sources of shocks. Rather, it is intended to help understand the mechanisms through which such shocks lead to economic fluctuations.'[5] We find that, without the ability to identify the economic shocks, a potentially important part of the mechanism by which these shocks affect the economy - the spillover effects - is also not identified. In effect, BCA offers a menu of observationally equivalent assessments about the importance of shocks to the intertemporal wedge. By focusing exclusively on the extreme case of zero spillovers, CKM select the element in the menu which minimizes the role of intertemporal shocks. We show that there are other elements in that menu which assign a very large role to intertemporal shocks.

Following is an outline of the paper. In the following section, we describe the model used in the analysis. In section 3, we elaborate on the observational equivalence results discussed above. In section 4, we discuss our model solution and estimation strategy. In section 5 we discuss the lack of identification of spillover effects in BCA. In section

[5]The quote is taken from the CKM introduction. It summarizes CKM's comments in section 3 of their paper.

6 we discuss the wedge decomposition under BCA and our modification to take into account spillovers. Section 7 displays the results of implementing BCA on various data sets. Concluding remarks appear in section 8. Additional technical details appear in three Appendices.

3.1. The Model and the Wedges

This section describes the model used in the analysis. In addition, we discuss the wedges and, in particular, our two specifications of the intertemporal wedge.

According CKM's version of the RBC model, households maximize:

$$E\sum_{t=0}^{\infty}\left(\beta\left(1+g_n\right)\right)^t\left[\log c_t+\psi\log\left(1-l_t\right)\right],\ 0<\beta<1,$$

where c_t and l_t denote per capita consumption and employment, respectively. Also, g_n is the population growth rate and $\psi>0$ is a parameter. The household budget constraint is

$$c_t+\left(1+\tau_{x,t}\right)x_t\leq r_tk_t+\left(1-\tau_{l,t}\right)w_tl_t+T_t,$$

where T_t denotes lump sum taxes, x_t denotes investment and $\tau_{l,t}$ denotes the labor wedge. Here, k_t denotes the beginning-of-period t stock of capital divided by the period t population. The variable, $\tau_{x,t}$, is CKM's specification of the intertemporal wedge. The technology for capital accumulation is given by:

$$(3.1)\qquad \left(1+g_n\right)k_{t+1}=\left(1-\delta\right)k_t+x_t-\Phi\left(\frac{x_t}{k_t}\right)k_t,$$

where $\Phi(\zeta)$ is symmetric about $\zeta = b$, where b is the steady state investment-capital ratio. In addition, to ensure that Φ has no impact on steady state, we suppose that $\Phi(b) = \Phi'(b) = 0$.

The household maximizes utility by choice of $\{c_t, k_{t+1}, l_t, x_t\}$, subject to its budget constraint, the capital evolution equation, the laws of motion of the wedges and the usual inequality constraints and no-Ponzi scheme condition.

The resource constraint is:

$$c_t + g_t + x_t = y(k_t, l_t, Z_t) = k_t^{\alpha} (Z_t l_t)^{1-\alpha}, \tag{3.2}$$

where

$$Z_t = \tilde{Z}_t (1 + g_z)^t,$$

and $\tilde{Z}_t$, the 'efficiency' wedge, is an exogenous stationary stochastic process. In the resource constraint, g_t denotes government purchases of goods and services plus net exports, which is assumed to have the following trend property:

$$g_t = \tilde{g}_t (1 + g_z)^t,$$

where $\tilde{g}_t$ is a stationary, exogenous stochastic process and $g_z \geq 0$.

Combining firm and household first order necessary conditions for optimization in the case $\Phi = 0$,

$$\frac{-u_{l,t}}{u_{c,t}} = (1-\tau_{l,t})\, y_{l,t} \tag{3.3}$$

$$u_{c,t} = \beta E_t u_{c,t+1} \frac{y_{k,t+1} + (1+\tau_{x,t+1})\, P_{k',t+1}\left[1-\delta-\Phi\left(\frac{x_{t+1}}{k_{t+1}}\right)+\Phi'\left(\frac{x_{t+1}}{k_{t+1}}\right)\frac{x_{t+1}}{k_{t+1}}\right]}{(1+\tau_{x,t})\, P_{k',t}} \tag{3.4}$$

where $u_{c,t}$ and $-u_{l,t}$ are the derivatives of period utility with respect to consumption and leisure, respectively. In addition, $y_{l,t}$ and $y_{k,t}$ are the marginal products of labor and capital, respectively. Also, the price of capital, $P_{k',t}$, is

$$P_{k',t} = \frac{1}{1-\Phi'\left(\frac{x_{t+1}}{k_{t+1}}\right)}. \tag{3.5}$$

The equilibrium values of $\{c_t, k_{t+1}, l_t, x_t\}$ are computed by solving (3.1), (3.2), (3.3), (3.4), subject to the transversality condition and the following law of motion for the exogenous shocks:

$$s_t = [I-P]\, P_0 + P s_{t-1} + u_t,\ \ s_t = \begin{pmatrix} \log \tilde{Z}_t \\ \tau_{l,t} \\ \tau_{x,t} \\ \log \tilde{g}_t \end{pmatrix},\ Eu_t u_t' = QQ' = V, \tag{3.6}$$

where P_0 is the 4×1 vector of unconditional means for s_t and

$$P = \begin{bmatrix} \bar{P} & 0 \\ 0 & p_{44} \end{bmatrix},\ Q = \begin{bmatrix} \bar{Q} & 0 \\ 0 & q_{44} \end{bmatrix}. \tag{3.7}$$

Here, P is stationary and $\bar{P}$ is not otherwise restricted. The symmetric matrix, V, in (3.6) must satisfy the zero restrictions implicit in $QQ' = V$, and the zeros in the lower diagonal part of Q in (3.7). We follow CKM in implementing the zero restrictions in our analysis of the US Great Depression. We do this in our analysis of OECD countries as well. In our analysis of postwar US data, we allow all elements of P and all elements in the lower triangular part of Q to be non-zero. The parameters of (3.6) are P_0, P, and V, possibly with the indicated zero restrictions on V and the zero and stationarity restrictions on P.

We consider an alternative specification of the intertemporal wedge. Our specification is motivated by our analysis of the version of the CF model with adjustment costs and by our analysis of BGG. In the appendix of the NBER working paper version of this document (Christiano and Davis (2006)), we derive equilibrium conditions for a version of the CF model with $\Phi \neq 0$. We establish a proposition displaying a set of wedges which, if added to the RBC economy, ensure that the equilibrium allocations of the RBC economy coincide with those of our version of the CF economy with investment adjustment costs. We show that the intertemporal wedge has the following form:

$$u_{c,t} = \beta E_t u_{c,t+1} \left(1 - \tau_{t+1}^k\right) R_{t+1}^k, \tag{3.8}$$

where,

$$R_t^k = \frac{y_{k,t} + P_{k',t} \left[1 - \delta - \Phi\left(\frac{x_t}{k_t}\right) + \Phi'\left(\frac{x_t}{k_t}\right)\frac{x_t}{k_t}\right]}{P_{k',t-1}} \tag{3.9}$$

Note that in the alternative formulation, the wedge is a tax on the gross return to capital, in contrast to CKM's value-added tax on investment purchases, $\tau_{x,t}$. In Appendix A of

Christiano and Davis (2006) we show that the CF model with adjustment costs implies τ_{t+1}^{k} is a function of uncertainty realized at date t, but not at date $t+1$.[6] We follow CKM in presuming that all wedges implied by the CF financial frictions apart from the intertemporal wedge, $1-\tau_{t+1}^{k}$, are quantitatively small and can be ignored.

In Appendix B of Christiano and Davis (2006) we derive the intertemporal wedge associated with the BGG model. That model also implies that the intertemporal wedge enters as $1-\tau_{t+1}^{k}$ in (3.8). The only difference is that under BGG, τ_{t+1}^{k} is a function of the period $t+1$ realization of uncertainty.[7]

In our alternative specification of the intertemporal wedge, we allow τ_{t}^{k} to respond to current and past information. This assumption encompasses both the CF and BGG financial friction models, since the econometric estimation is free to produce a τ_{t}^{k} whose response to current information is very small.

3.2. General Observations on the Robustness of BCA to Modeling Details

In later sections, we show that the conclusions of BCA for the importance of the intertemporal wedge are not robust to alternative specifications of the intertemporal wedge, and to alternative specifications of investment adjustment costs. This finding may at first

[6] That appendix provides a careful derivation of our result, because our finding for the way the intertemporal wedge enters (3.8) differs from CKM's finding. CKM consider the case, $\Phi=0$, in deriving the wedge representation of the CF model. The results for the $\Phi=0$ and $\Phi\neq 0$ cases are qualitatively different. When $\Phi=0$ capital producers simply produce increments to the capital stock, which capital owners add to the existing undepreciated capital by themselves. When $\Phi\neq 0$, old capital is a fundamental input in the production of new capital. In this case, we assume that the capital producers must purchase the economy's entire stock of capital in order to produce new capital, so that their financing requirements and the associated frictions are different. There are perhaps other ways of arranging the production of new installed capital when $\Phi\neq 0$. We find our way convenient because it results in an intertemporal wedge that virtually coincides with the one we derive for BGG

[7] CKM derive the intertemporal wedge for a version of the BGG model in which banks have access to complete state-contingent markets. Our wedge formula applies to the model analyzed in BGG, which does not permit complete markets.

seem puzzling in light of a type of observational equivalence result emphasized in CKM. An example of this type of result which occurs when BCA is done with a linearly approximated RBC model is the following. Consider an RBC economy with, say, no investment adjustment costs (i.e., $\Phi = 0$) and a particular time series representation for the wedges. After introducing adjustment costs (i.e., $\Phi \neq 0$), one can find a new representation of the intertemporal wedge which ensures that the equilibria of the economies with and without adjustment costs coincide.[8] This is an observational equivalence result because it implies that the likelihood of a set of allocations is invariant to the presence of adjustment costs. This case of adjustment costs is just example of the type of observational equivalence result we have in mind. For example, consider an RBC economy in which the intertemporal wedge is of the $\tau_{x,t}$ type emphasized by CKM. Given a specification of the joint time series representation of $\tau_{x,t}$ and the other wedges, the $\tau_{x,t}$ RBC model implies a set of equilibrium allocations. Now consider an alternative RBC economy in which the intertemporal wedge is of the τ_t^k type. There exists a specification of the joint stochastic process for τ_t^k and the other wedges having the property that the equilibrium allocations in the τ_t^k RBC model coincide with those in the $\tau_{x,t}$ RBC model. Again, this stochastic process is identified from the requirement that the after tax rates of return in the two economies coincide. In both of the above examples, it is clear that the observational equivalence result depends on the assumption that the time series representations used for the shocks

[8] Here, we make use of our asumption that analysis is done using log-linear approximation. In this case, the only effect of the change in Φ is to change the rate of return on capital. For example, in the linear approximation the law of motion for the capital stock, (3.1), is always linear and invariant to a.

are sufficiently flexible to accommodate any specification for the stochastic process of the wedges.[9]

We wish to stress here that the equilibrium observational equivalence result does not imply a 'BCA robustness result'. In particular, the outcome of BCA (i.e., the outcome of feeding fitted wedges, one at a time, to a model) is not expected to be robust to the specification of investment adjustment costs, or to whether the intertemporal wedge is modeled as $\tau_{x,t}$ or τ_t^k. There are two reasons for this lack of robustness. One is practical and reflects that the analyst must confine him/herself to a specific parametric time series representation of the wedges, thus potentially ruling out one of the conditions of the observational equivalence result. The other, deeper, reason is the one mentioned in the introduction. Even if the analyst uses a completely flexible time series representation of the wedges, the intertemporal wedge represents a different bundle of fundamental shocks under alternative perturbations of the model. Feeding the measured intertemporal wedge to an RBC model under alternative model perturbations represents a different experiment and so is expected to produce a different outcome.

To illustrate these observations, suppose the data are generated by an RBC model in which intertemporal wedge is the τ_t^k type, with a certain specification of the adjustment cost function, Φ. The joint time series representation of the wedges is given by (3.6), in which P and Q are diagonal. Thus, each wedge is uncorrelated with all other wedges, at all leads and lags. In this case, BCA has a clear interpretation: when the estimated intertemporal wedge is fed to the baseline RBC model, the simulations display the model's response to a particular history of past innovations to that wedge alone. Suppose the

[9]This is a special case of a well-known result that econometric identification often hinges on having sufficient restrictions on the unobserved shocks.

econometrician is provided with an infinite amount of data, but misspecifies the adjustment cost function, Φ. As in BCA, the econometrician only estimates the joint time series representation of the wedges, and holds the misspecified Φ and other nonstochastic parts of the economy fixed. We assume that the econometrician's time series representation for the wedges is sufficiently flexible to encompass the quasi-true time series of the wedges that is implied by the observational equivalence result. We obtain insight into BCA by deriving that time series representation. The requirement that the after tax rates of return in the econometrician's model coincide with the true after tax rate of return implies, using (3.9):

$$1-\tau_{t+1}^{k}=\left(1-\bar{\tau}_{t+1}^{k}\right)\frac{y_{k,t}+\bar{P}_{k',t}\left[1-\delta-\bar{\Phi}\left(\frac{x_t}{k_t}\right)+\bar{\Phi}'\left(\frac{x_t}{k_t}\right)\frac{x_t}{k_t}\right]}{y_{k,t}+P_{k',t}\left[1-\delta-\Phi\left(\frac{x_t}{k_t}\right)+\Phi'\left(\frac{x_t}{k_t}\right)\frac{x_t}{k_t}\right]}. \tag{3.10}$$

Here, a $-$ over a variable indicates the value of the variable in the true model and absence of a $-$ indicates the value estimated by the econometrician who misspecifies Φ. The endogenous variables on the right side of the equality in (3.10) are specific functions of the history of the innovations driving the wedges in the actual economy. Then, according to (3.10), the adjusted time series representation of τ_t^k is the convolution of these functions with the function on the right of the equality in (3.10). We derive this map from the fundamental innovations in the economy to τ_t^k using linearization.

Consider the true specification of Φ and the true joint time series representation of the wedges, s_t, given in (3.6). Let z_t denote the list of endogenous variables in the model, i.e., $z_t=(c_t,x_t,k_{t+1},l_t,\tau_t^k)$, where the quantity variables are measured in log deviations from steady state and τ_t^k is in deviation from steady state. The equilibrium conditions of

z_t may be written in the form:

$$E_t\left[\alpha_0 z_{t+1} + \alpha_1 z_t + \alpha_2 z_{t-1} + \beta_0 s_{t+1} + \beta_1 s_t\right] = 0, \text{ with } s_t = P s_{t-1} + Q\varepsilon_t.$$

Here, $s_t = \left(\log \tilde{Z}_t, \tau_{l,t}, \bar{\tau}_t^k, \log \tilde{g}_t\right)'$. The expectational difference equation is composed of the intertemporal first order condition (3.8), the intratemporal first order condition (3.3), the law of motion for capital (3.1), the resource constraint, (3.2), and the mapping from $\bar{\tau}_t^k$ to τ_t^k, (3.10), all after suitable log-linearization. The solution to this system is written $z_t = Az_{t-1} + Bs_t$, or, when expressed in moving average form[10]:

$$z_t = [I - AL]^{-1} B [I - PL]^{-1} Q\varepsilon_t.$$

Let τ denote the 5-dimensional column vector with all zeros, except a 1 in the 5th location. Then, the time series representation for τ_t^k is

$$\tau_t^k = \tau [I - AL]^{-1} B [I - PL]^{-1} Q\varepsilon_t.$$

This is the convolution of (3.10) with the time series representation of the (linearized) variables in (3.10). Let ν denote the 3 by 4 matrix constructed by deleting the third row of the 4-dimensional identity matrix and let S_t denote the 3 dimensional vector obtained by deleting $\bar{\tau}_t^k$ from s_t. We conclude that the econometrician who misspecifies Φ will estimate the following joint time series representation for the wedges in his misspecified model:

$$\begin{pmatrix} \tau_t^k \\ S_t \end{pmatrix} = \begin{bmatrix} \tau [I - AL]^{-1} B \\ \nu \end{bmatrix} [I - PL]^{-1} Q\varepsilon_t.$$

[10]For further discussion, see Christiano (2002).

By inspection, it is clear that in general, the new joint series representation of $\left(\tau_t^k, S_t\right)$ has a moving average component. To see this, it is useful to examine the *iid* case, $P = 0$ and $Q = I$. It is easily verified that $\tau\left[I - A\right]^{-1} B$ has the following form:

$$\tau\left[I - A\right]^{-1} B = \left[\begin{array}{cccc} B_{51} - a_{63}B_{31}L & B_{52} - a_{63}B_{32}L & B_{53} - a_{63}B_{33}L & B_{54} - a_{63}B_{34}L \end{array}\right],$$

where B_{ij} denotes the ij^{th} element of B and a_{63} denotes the $6,3$ element of A. We conclude that the new joint representation of the wedges is:

$$\left(\begin{array}{c} \tau_t^k \\ S_t \end{array}\right) = \left[\begin{array}{c} \left(\begin{array}{cccc} B_{51} - a_{63}B_{31}L & B_{52} - a_{63}B_{32}L & B_{53} - a_{63}B_{33}L & B_{54} - a_{63}B_{34}L \end{array}\right) \\ \nu \end{array}\right] \varepsilon_t.$$

Note that the intertemporal wedge has a pure, first order moving average representation, even though τ_t^k in the correctly specified economy is *iid* and a function only of the third element of ε_t. Evidently, the wedges in the misspecified economy do not obey the same first order VAR(1) representation that s_t does. Thus, the analyst who is restricted VAR(1) (or, VAR(q), $q < \infty$) representations for the wedges misrepresents the reduced form of the data. Under these circumstances, it is not surprising that the conclusions of BCA will be different, across different specifications of Φ.

Now, suppose that the analyst adopts a sufficiently flexible time series representation of the wedges, so that the specification error described in the previous paragraph does not occur. The intertemporal wedge, $\bar{\tau}_t^k$, computed by the econometrician working with the correct specification of Φ is a function of just the current realization of the third element of ε_t. In the alternative specification, τ_t^k is a function of the entire history of all elements of ε_t. Clearly, feeding the estimated intertemporal wedge to the model is a different experiment across the two different specifications of Φ. This is why we do not expect the results of BCA to be robust to perturbations in the RBC model.

3.3. Model Solution and Estimation

Here, we describe how we assigned values to the model parameters. A subset of the parameters were not estimated. These were set as in CKM:

$$
\begin{aligned}
\beta &= 1/1.03,\ \alpha = 0.35,\ \delta = 0.0464,\ \psi = 2.24, \\
g_n &= 0.015,\ g_z = 0.016.
\end{aligned}
\tag{3.11}
$$

Here, β, δ, g_n, and g_z are expressed at annual rates. These are suitably adjusted when we analyze quarterly data. The first subsection below discusses the estimation of the parameters of the exogenous shocks, P_0, P, and V, using data on output, consumption, investment and government consumption plus net exports. Estimation is carried out conditional on a parameterization of the adjustment cost function. The parameterization of the adjustment cost function is discussed in the second subsection. The third subsection rebuts some criticisms of the investment adjustment cost function expressed in CKM. Their criticisms suggest that investment adjustment costs are, in effect, a 'nonstarter'. Since they are not empirically interesting, they therefore do not constitute a compelling basis for criticizing BCA. We explain why we disagree with this assessment.

3.3.1. Estimating the Parameters of the Time-Series Representation of the Wedges

For the US Great Depression, we used annual data covering the period, 1901-1940.[11] Quarterly data covering the period 1959Q1-2004Q3 were used for the US and quarterly data over various periods were used on 14 other OECD countries.[12] Following CKM, the elements of the matrices, P and V are estimated subject to the zero restrictions described in section **??**, and to the restriction that the maximal eigenvalue of P not exceed 0.995.

The first step of estimation is to set up the model's solution in state space - observer form:

$$Y_t = H(\xi_t;\gamma) + \upsilon_t \tag{3.12}$$

$$\xi_t = F\left(\xi_{t-1};\gamma\right) + \eta u_t \tag{3.13}$$

$$\gamma = (P, P_0, V), \ \eta = \begin{pmatrix} \tilde{0} \\ I \end{pmatrix}, \ E\upsilon_t\upsilon_t' = R, \ Eu_tu_t' = V,$$

where $\tilde{0}$ is a 1×4 vector of zeros and ξ_t is the state of the system:

$$\xi_t = \begin{pmatrix} \log \tilde{k}_t \\ s_t \end{pmatrix}, \tag{3.14}$$

[11] These data were taken from CKM, as supplied on Ellen McGrattan's web site.

[12] US data are the data associated with the CKM project, and were taken from Ellen McGrattan's web page. With two exceptions, data for other OECD countries were taken from Chari, Kehoe and McGrattan (2002), also on Ellen McGrattan's web site. Data on hours worked were taken from the OECD productivity database. These data are annual and were converted to quarterly by log-linear interpolation. Population data were taken from the OECD national databases and log-linearly intertpolated to quarterly.

where $\tilde{k}_t = k_t / (1+g_z)^t$. Also, Y_t is the observation vector:

$$Y_t = \begin{pmatrix} \log \tilde{y}_t \\ \log \tilde{x}_t \\ \log l_t \\ \log \tilde{g}_t \end{pmatrix}, \tag{3.15}$$

where $\tilde{x}_t = x_t / (1+g_z)^t$. Finally, υ_t is a 4×1 vector of measurement errors, with

$$R = 0.0001 \times I_4, \tag{3.16}$$

where I_4 is the four-dimensional identity matrix and CKM set the scale factor exogenously (see CKM (technical appendix, page 16)). We refer to this specification of R as the 'CKM measurement error assumption'. We repeat the analysis under CKM measurement error, as well as with $R = 0$.

As noted in the introduction, the CKM specification of measurement error has an impact on the analysis. CKM do not explain why they include measurement error, nor do they discuss the a priori evidence which leads them to the specific values they choose for the measurement error variance.[13] We do have reason to believe the data are measured with error. However, we know of no reason to take seriously the notion that CKM's specification even approximately captures actual data measurement error.[14]

---

[13]As already noted, other parameter values are also fixed in the analysis, such as production function parameters. Dogmatic priors like this can perhaps be justified by appealing to analyses based on other data, such as observations on income shares. We are not aware of any such argument, however, that can be used as a basis for adopting the dogmatic priors in (3.16).

[14]Based on what we know about the way data are collected, there is strong a priori reason to question the CKM model of measurement error. For a careful discussion, see Sargent (1989).

We implement BCA using first and second-order approximations to the model's equilibrium conditions. Consider the first order approximation. In this case, the representation of the policy rule is:

$$\log \tilde{k}_{t+1} = (1-\lambda)\,\lambda_0 + \lambda \log \tilde{k}_t + \psi s_t, \tag{3.17}$$

where λ_0 and λ are scalars and ψ is a 1×4 row vector. Then, (3.12)-(3.13) can be written:

$$\begin{aligned} \xi_t &= F_0 + F_1 \xi_{t-1} + \eta \varepsilon_t, \\ F_0 &= \begin{bmatrix} (1-\lambda)\,\lambda_0 \\ (I-P)\,P_0 \end{bmatrix},\ F_1 = \begin{bmatrix} \lambda & \psi \\ 0 & P \end{bmatrix}, \end{aligned}$$

where F_0 is a 5×1 column vector, and F_1 is a 5×5 matrix. Also,

$$Y_t = H_0 + H_1 \xi_t + \upsilon_t, \tag{3.18}$$

where H_0 is a 4×1 column vector and H_1 is a 4×4 matrix. The Gaussian likelihood is constructed using F_0, F_1, H_0, H_1, V, R, and $Y = (Y_1, ..., Y_T)$ (see Hamilton (1994)). These in turn can be constructed using γ, R. Thus, the likelihood can be expressed as $L\,(Y|\gamma; R)$.

For the nonlinear case, we use the algorithm in Schmitt-Grohe and Uribe (2004) to obtain second order approximations to the functions, F and H in (3.12) and (3.13). It is easy to see that even if u_t is Normally distributed, Y_t will not be Normal in this nonlinear system. We nevertheless proceed to form the Gaussian density function using the unscented filter described in Wan and van der Merwe (2001). It is known that under certain conditions, Gaussian maximum likelihood estimation has the usual desirable properties, even when the data are not Gaussian.

3.3.2. Investment Adjustment Costs

To analyze the version of the model with adjustment costs, we must parameterize the investment adjustment cost function, Φ. Our calibration is based on our interpretation of the variable, $P_{k',t}$. On this dimension, the CF and BGG models differ slightly (for details, see Appendices A and B). Both agree that $P_{k',t}$ is the marginal cost, in units of consumption goods, of producing new capital when only (3.1) is considered.[15] However, in the CF model, financial frictions introduce a wedge between the market price of capital and $P_{k',t}$. Still, in practice the discrepancy between $P_{k',t}$ and the market price of new capital in the CF model with adjustment costs may be quantitatively small. To see this, it is instructive to consider the response of the variables in the CF model (where $P_{k',t} = 1$ always) to a technology shock. According to CF (see Figure 2 in CF), the contemporaneous response of the market price of capital is only one-tenth the contemporaneous response of investment. That simulation suggests that the distinction between $P_{k',t}$ and the market price of capital may not be large in the CF model.

In the BGG model, financial frictions arise inside the relationship between the managers of capital and banks, and so the frictions do not open wedge between the marginal cost of capital and $P_{k',t}$. As a consequence, $P_{k',t}$ corresponds to the market price of capital in the BGG model.

[15]It is easy to verify that $P_{k',t}$ in (3.1) corresponds to the price of investment goods (i.e., unity) divided by the marginal product of investment goods in producing end of period capital.

Under the interpretation of $P_{k',t}$ as the market price of capital, we can calibrate Φ based on empirical estimates of the elasticity of investment with respect to the price of capital (i.e., Tobin's q). From (3.5), this is

$$\frac{d\log\left(x_t/k_t\right)}{d\log P_{k',t}}=\frac{1}{\Phi''\left(b\right)b}. \tag{3.19}$$

According to estimates reported in Abel (1980) and Eberly (1997), Tobin's q lies in a range of 0.6 to 1.4. Interestingly, if we just consider the period of largest fall in the Dow Jones Industrial average during the Great Depression, 1929Q4 to 1932Q4, the ratio of the percent change in investment to the percent change in the Dow is 0.68.[16] This is an estimate of Tobin's q under the assumption that the movement in the Dow reflects primarily the price of capital, and not its quantity.[17] This estimate lies in the middle of the Abel-Eberly range of estimates. A unit Tobin's q elasticity implies $\Phi''\left(b\right)=1/b$.

Another factor impacting on our choice of $\Phi''\left(b\right)$ is the model's implication for the rate of return on capital, R^k. Figure 1A shows the results corresponding to Tobin's q elasticities 1/2, 1, 3 and ∞ (the latter corresponds to $\Phi''\left(b\right)=0$). For each elasticity, the model was estimated using the linearization strategy and using quarterly US data covering the period 1959QIV-2003QI. For these calculations, the only feature of Φ that is required is the value of $\Phi''\left(b\right)$. The model-based estimate of R_t^k, (3.9), was computed using the two-sided Kalman smoother.[18] The US data on R_t^k were constructed using Robert

[16]This is the ratio of the log difference in investment to the log difference in the Dow, over the period indicated. Both variables were in nominal terms.

[17]By associating the model's capital stock with what is priced in the Dow, we are implicitly taking the position that capital in the model corresponds to both tangible and intangible capital.

[18]See Hamilton (1994) for a discussion. The two-sided smoother is required because we do not use empirical data on the capital stock, which is an input in (3.9). Presumably, the smoother estimates the capital stock by combining the investment data with the capital accumulation equation.

Shiller's data on real dividends and real stock prices for the $S\&P$ composite index. In the case of both model-based and actual R_t^k, we report centered, equally weighted, 5 quarter moving averages. Note that without adjustment costs, the model drastically understates the volatility in R_t^k. With a Tobin's q elasticity of 3 (i.e., $\Phi''(b) = 1/(3b)$) the model still substantially understates that volatility. With an elasticity around unity, the model begins to reproduce the volatility of R^k, though it is still somewhat low. Only with an elasticity around 1/2 does the model nearly replicate the volatility of R^k. These results reinforce our impression that the data suggest a Tobin's q elasticity of unity or less. To be conservative, we work with an elasticity of unity.

3.3.3. Responding to CKM's Criticism about Adjustment Costs

CKM criticize the use of adjustment costs with a unit Tobin's q elasticity for two reasons. According to their first critique, adjustment costs with a unit Tobin's q elasticity imply that an unreasonably large amount of resources are absorbed by adjustment costs during collapse of investment in the Great Depression. This conclusion is based on the arbitrary assumption that the adjustment cost function, Φ, is globally quadratic. But, we show that other functional forms for Φ can be found with the property, $\Phi''(b) = 1/b$, whose global properties do not imply that an inordinate amount of resources were used up in investment adjustment costs in the Great Depression. Second, CKM assert that an adjustment cost formulation which implies a static relationship between the investment-capital ratio and Tobin's q is empirically implausible. But, we show that BCA lacks robustness even with the specification of adjustment costs proposed in Christiano, Eichenbaum and Evans (2005), which does not imply a static relationship the investment-capital ratio and Tobin's

q. This adjustment cost function, in which adjustment costs are a function of the change in the flow of investment, also does not imply that an inordinate amount of resources were used up in adjustment costs during the collapse of investment in the 1930s.[19]

The globally quadratic adjustment cost formulation adopted by CKM is:

$$\Phi\left(\frac{x_t}{k_t}\right) = \frac{a}{2}\left(\frac{x_t}{k_t} - b\right)^2,$$

so that $\Phi''(b) = a$. Imposing that Tobin's q elasticity is unity, the resources lost to adjustment costs, as a fraction of output, is given by:

$$\Phi\left(\frac{x_t}{k_t}\right) = \frac{1}{2}\left(\lambda_t - 1\right)^2 \frac{x}{y\mu_t}, \tag{3.20}$$

according to (3.1). Here, x/y is the steady state investment to output ratio. In (3.20), we have used $x = bk$ in the steady state. Here, λ_t is the time t investment-capital ratio, expressed as a ratio to its steady state value, b. Also, μ_t is the output-capital ratio, expressed as a ratio to its steady state value, y/k. Figure 6 indicates that output was 10 percent below trend in 1930, and then fell another 10 percent in each of 1931 and 1932. In 1933, the trough of the Depression, it fell yet another 5 percent, so that by 1933 output was a full 35 percent below trend. The drop in investment was even more dramatic. In 1930, 1931, 1932 and 1933 it was about 30, 50, 70 and 70 percent below trend, respectively. Using our capital accumulation equation, we infer that the stock of capital was 10 percent below trend in 1933.

[19]This adjustment cost function has the additional advantage that it receives empirical support from the analysis of housing investment (see Rosen and Topel (1988)) and aggregate Tobin's q data (see Matsuyama (1984)), in addition to the empirical evidence in Christiano, Eichenbaum and Evans (2006). Also, this adjustment cost formulation has economically interesting microfoundations, as shown in Lucca (2006) and Matsuyama (1984).

Since investment was 70 percent below its trend in 1933 and the capital stock was 10 percent its trend then, we infer that the investment to capital ratio is 60 percent below steady state, i.e., $\lambda_{1933} = 0.40$. Output was 35 percent below steady state in 1933, and we infer that the output-capital ratio was 25 percent below trend, so that $\mu_{1933} = 0.75$. Substituting these into (3.20),

$$\Phi\left(\frac{x_t}{k_t}\right) = \frac{1}{2}\left(0.40 - 1\right)^2 \left(0.23\right)/0.75 = 0.055,$$

or 5.5 percent. Given that output was 35 percent below trend in 1933, the implication is that 16 percent of the drop in output reflected resources lost to adjustment costs associated with the low level of investment. To see how sensitive this conclusion is to the choice of functional form for Φ, consider Figure 1B, which graphs (3.20) for $100\lambda_t$ ranging from 40 percent to 160 percent, holding $x/\left(y\mu_t\right)$ fixed at 0.31. Note how the quadratic curve hits the vertical axis at 5.5 percent. The other curve in Figure 1B coincides with the quadratic function for λ_t roughly in its range for postwar business cycles. Outside this range, the alternative function is flatter than the quadratic, and it hits the vertical axis at 2.5 percent. The alternative adjustment cost function has a much more modest implication for the amount of resources lost to adjustment costs as investment collapsed in the Great Depression. Yet, the implications of the model with the alternative adjustment cost function for postwar business cycles coincides with the implications of the model with the quadratic adjustment cost function.[20]

[20]The alternative adjustment cost function is a 10^{th} degree polynomial, and so it has a continuous derivative of every order. It was constructed as follows. We constructed a 'target' function by splicing the quadratic function in the range, $\lambda \in (0.85, 1.15)$, with straight lines on either end. The straight lines have slope equal to that of the quadratic function at the point where they meet. The 10^{th} degree polynomial was fit by standard Chebyshev interpolation.

To address CKM's second concern about adjustment costs, we also considered the following formulation:

$$(1 + g_n)\, k_{t+1} = (1 - \delta)\, k_t + \left[1 - \frac{a}{2}\left(\frac{x_t}{x_{t-1}} - 1\right)^2\right] x_t.$$

With this formulation of adjustment costs, investment responds differently to permanent and temporary changes in the price of capital. This addresses one of CKM's concerns about investment adjustment costs. To address the other concern, we needed to assign a value to a. For this, we estimated the parameters of the joint time series representation of the wedges for various values of a, using postwar US data. We found that with $a = 3.75$ the model's implications for the volatility of the rate of return on capital virtually coincides with the implications of our baseline model with a unit Tobin's q elasticity. We then used the Balke and Gordon quarterly data on investment and output in the 1930s to compute the fraction of output lost due to adjustment as investment plunged at the start of the Great Depression. We found that the largest fraction of output lost due to adjustment costs in the period 1929Q1-1933Q1 was 1.46 percent. According to the Balke and Gordon data, investment rose sharply starting in 1933Q2. Adjustment costs were larger then, but adjustment costs in expansions are less of a concern to CKM.[21] We conclude that with the alternative adjustment costs, neither of CKM's two objections apply.

[21] According to Balke and Gordon's data, per capita real investment, including durable goods, (1929 dollars), was 44, 65, 119, and 83 in the first to fourth quarters of 1933. Our estimate of the percent of aggregate output lost to adjustment costs is 0.77, 3.09, 17.04, and 1.69 for each of the four quarters in 1933. The number for 1933Q3 is very large. However, we note that it is generated by a rise in investment, not a fall. In addition, we are suspicious that investment rose 83 percent in 1933Q3 and then fell about 30 percent in 1933Q4. This sharp volatility is consistent with the possibility that measurement error overstated the level of invesment in 1933Q3.

Significantly, our finding that BCA is sensitive to the presence of adjustment costs is also true when the adjustment costs are in terms of the change in investment. Ignoring the spillover effects between wedges, as CKM do, we calculated the percent of the fall in output due to the intertemporal wedge at the trough of five postwar US recessions. For the 1970, 1974, 1980, 1990 and 2000 recessions, the percentages are 17, 30, 14, 26, and 43, respectively. All these are substantial amounts and certainly do not warrant the CKM conclusion that financial frictions which manifest themselves primarily in the intertemporal wedge are not worth pursuing.

3.4. Identification, the Importance of Financial Frictions and BCA

In the introduction we discussed the sense in which the importance of financial frictions is not identified under BCA. We explain this here. We describe a statistic which we use to characterize the importance of financial frictions. We show that a range of values for this statistic is consistent with the same value of the likelihood function.

Until now, the basic shocks driving the system have been u_t in (3.6). The interactions among these shocks are left almost completely unrestricted under BCA. In part, this is because the u_t's are found to be highly correlated in practice. This correlation is assumed to reflect that the elements of u_t are overlapping combinations of different fundamental economic shocks. Because fundamental economic shocks are assumed to be primitive and to have separate origins, they are often assumed to be uncorrelated. We make this uncorrelatedness assumption here. Denote the 5×1 vector of fundamental economic shocks by e_t. We normalize their variances to unity, so that $Ee_te_t' = I$. We assume that the fundamental shocks are related to the u_t's by the following invertible relationship:

$$u_t = Ce_t,\ Ee_te_t' = I,\ CC' = V, \tag{3.21}$$

where C has the structure of Q in (3.7).[22] It is well known that even with a particular estimate of V in hand, there are many C's that satisfy $CC' = V$. Alternative specifications of C that preserve the property, $CC' = V$, are observationally equivalent with respect to a set of observations, $Y = (Y_1, ..., Y_T)$. Because this property plays a key role in our analysis, it is useful to state it as a proposition:

Proposition 1. Consider a set of model parameter values, $\gamma = (P, P_0, V)$, with likelihood value, $L(Y|\gamma; R)$. Perturbations of C such that $CC' = V$ have no impact on the likelihood, L.

Obviously, Proposition 1 applies for both the linear and the nonlinear strategies we use to approximate the likelihood. Although BCA makes many detailed economic assumptions (e.g., details about utility and technology), it does not make the assumptions needed to identify the fundamental economic disturbances, e_t, to the economy.

We suppose, for the purpose of our discussion, that the third element in e_t corresponds to the financial frictions shock which originates in the intertemporal wedge, which is the third element of s_t.[23] To discuss the difficulty of pinning down the importance of financial

[22] We are assuming that the fundamental economic shocks can be recovered from the space of current and past shocks. Lippi and Reichlin (1993) challenge this assumption and discuss some of the implications of its failure. See also Sims and Zha (1996) and Fernandez-Villaverde, Rubio-Ramirez and Sargent (2006).

[23] In an agency cost model, these shocks could be perturbations to the variance of idiosyncratic disturbances affecting entrepreneurs, or to the survival rate of entrepreneurs. See Christiano, Motto and Rostagno (2004, 2006) for examples.

frictions, it is useful to develop a constructive characterization of the family of C's that satisfy (3.21).[24] Write

$$C = \bar{C}W, \tag{3.22}$$

where W is any orthonormal matrix and $\bar{C}$ is the unique lower diagonal matrix with non-negative diagonal elements having the property that $\bar{C}\bar{C}' = V$. Although each C in (3.22) is observationally equivalent by Proposition 1, each C implies a different e_t. To see this, note that for any sequence of fitted disturbances, u_t, one can recover a time series of e_t using

$$e_t = C^{-1}u_t = W'\bar{C}^{-1}u_t. \tag{3.23}$$

To see how many e_t's there are, for given V and sequence u_t, let

$$W = \frac{1}{2a}\begin{bmatrix} a & b & c & d \\ -b & a & e & f \\ -c & -e & a & g \\ -d & -f & -g & a \end{bmatrix},$$

where $g = (cf - de)/b$. It is easy to verify that W is orthonormal for each $\theta = (a, b, c, d, e, f)$. For a fixed set of observed u_t, $t = 1, ..., T$, there is a different sequence, e_t, $t = 1, ..., T$, associated with almost all $\theta \in R^6$. According to Proposition 1, the likelihood of the data based on the linear approximation is constant with respect to variations in θ.

[24]Here, we follow the strategy pursued in Uhlig (2002).

We are now in a position to describe our measure of the importance of financial frictions. This measure combines the two mechanisms by which financial frictions can matter. The first is that financial frictions represent a source of shocks (see Figure 2). For us, the stand-in for these shocks is e_{3t}. These operate on the economy by driving the intertemporal wedge, s_{3t}, (see (i) in Figure 2) and through spillover effects onto other wedges ((iii) in Figure 2). The second mechanism reflects that financial frictions modify the way non-financial friction shocks, e_{1t}, e_{2t}, e_{4t}, affect the economy. They do so by inducing movements in the intertemporal wedge (see (ii) in Figure 2). Our measure of the importance of financial frictions is the ratio of what the variance of HP-filtered output would be if only the financial frictions were operative, to the total variance of HP-filtered output. We construct this formally as follows. The wedges, s_t, have the following moving average representation (here, we ignore constant terms):

$$s_t = [I - PL]^{-1} Q\varepsilon_t = F(L)\varepsilon_t,$$

say. Define

$$\tilde{s}_t = \tilde{F}(L)\,\varepsilon_t.$$

Here, $\tilde{F}(L)$ denotes the version of $F(L)$ in which all elements have been set to zero, except those in the third column and the third row (i.e., $\tilde{F}(L)$ is $F(L)$ with the (i, j) elements set to zero, for $i, j = 1, 2, 4$.) The dynamics of $\tilde{s}_t$ reflect the mechanisms by which the financial frictions affect the wedges. The fact that the $3, 3$ element of $F(L)$ is kept in $\tilde{F}(L)$ means that the financial friction shock is permitted to exert its effect on the intertemporal wedge, s_{3t}. The fact that we keep the other elements of the third column of $F(L)$ means that we include in $\tilde{F}(L)$ the spillover effects from the financial friction shock to the other wedges. Regarding the other elements of ε_t, $\tilde{F}(L)$ only includes their spillover effects onto the intertemporal wedge. This is our way of capturing the notion that financial frictions modify the transmission of non-financial shocks. Although $\tilde{s}_t$ represents the component of s_t corresponding to financial frictions, it is important to bear in mind that it is not an orthogonal decomposition of s_t.[25] For example, it is possible for the variance of $\tilde{s}_t$ to exceed that of s_t.

Write (3.17) in lag operator form:

$$\log \tilde{k}_t = \frac{\gamma L}{1 - \lambda L} s_t,$$

and express the linearized observer equation, (3.18), as follows:

$$Y_t = h_0 s_t + h_1 \log \tilde{k}_t + v_t$$

where h_0 is a 4×4 matrix and h_1 is a 4×1 column vector. Then,

$$Y_t = H(L) F(L) \varepsilon_t + v_t,$$

where

$$H(L) = h_0 + h_1 \frac{\gamma L}{1 - \lambda L}.$$

[25] That is, $\tilde{s}_t$ and $s_t - \tilde{s}_t$ are correlated. Since $var(s_t) = var(\tilde{s}_t) + var(s_t - \tilde{s}_t) + 2cov(\tilde{s}, s_t - \tilde{s}_t)$, it is possible for $var(\tilde{s}_t) > var(s_t)$ if the covariance term is sufficiently negative.

The representation of Y_t that reflects only the financial frictions is denoted $\tilde{Y}_t$, and is as follows:

$$\tilde{Y}_t = H(L)\tilde{F}(L)\,\varepsilon_t + \upsilon_t. \tag{3.24}$$

The spectral densities of $\tilde{Y}_t$ and Y_t are, respectively,

$$S_{\tilde{Y}}(\omega) = H(e^{-i\omega})\tilde{F}\left(e^{-i\omega}\right)\tilde{F}\left(e^{i\omega}\right)' H\left(e^{i\omega}\right)' + R$$

$$S_Y(\omega) = H(e^{-i\omega})F\left(e^{-i\omega}\right)F\left(e^{i\omega}\right)' H\left(e^{i\omega}\right)' + R.$$

The variance of Y_t, denoted C_0, can be computed by solving the following expression for large N :

$$C_0 = \frac{1}{N}S_Y(\omega_0) + \frac{2}{N}\sum_{k=1}^{\frac{N}{2}-1} re\left(S_Y(\omega_k)\right) + \frac{1}{N}S_Y(\omega_{N/2}),\ \omega_j = \frac{2\pi j}{N}.$$

The variance of $\tilde{Y}_t$, $\tilde{C}_0$, can be computed in an analogous way.

Our measure of the importance of financial frictions, f, is the 1,1 element of $\tilde{C}_0$, which we denote $\tilde{C}_0^{11}$. Our measure of financial frictions scales this by the 1,1 element of C_0 :

$$f = \frac{\tilde{C}_0^{11}}{C_0^{11}}. \tag{3.25}$$

Since it is a ratio of variances, f must be positive. However, because (3.25) is not based on an orthogonal decomposition, f may be larger than unity. The importance of financial frictions is not identified, because almost all perturbations in θ imply different values of f, but the same value of the likelihood, by Proposition 1.

3.5. Wedge Decompositions

We describe decompositions of the data during a recession which begins in period $t = t_1$ and ends in period $t = t_2$. CKM's strategy, which we call the 'baseline decomposition', is as follows. CKM ask how the recession would have unfolded if only the wedge, s_{3t}, evolved as it did and the other wedges remained constant at their values at the start of the recession. We find the sequence, ε_t, $t = t_1, ..., t_2$ which has the property that when this is input into (3.6), the third element of the simulated s_t, $t = t_1, ..., t_2$, coincides with its estimated values and the other elements of s_t are fixed at their value at $t = t_1$.

We investigate an alternative strategy for assessing the role of financial frictions, which recognizes the roles played by these frictions discussed in the introduction and in section **??**. Such a strategy would choose a value for the rotation parameter, θ and use the implied sequence of e_t's to simulate (3.24). However, because $\tilde{F}$ is an infinite-ordered moving average representation, we decided this strategy is impractical and we devised a closely related one instead. The strategy we implemented ('rotation decomposition') recognizes that financial shocks drive both the intertemporal wedge and have spillover effects on other wedges. But, it does not capture the spillover effects from other shocks onto the intertemporal wedge. In this sense our rotation decomposition understates the role of financial frictions. However, we mitigate the latter effect by working with the rotation, θ, which maximizes the role of financial frictions, f.

The rotation decomposition is constructed as follows. We compute u_t, $t = t_1, ..., t_2$, and the value, θ^*, of $\theta \in R^6$ which maximizes f in (3.25). Then, we fix W and compute the implied sequence, e_t, for $t = t_1, ..., t_2$ using (3.23) and the value of C implied by (3.22). Next, set to zero all but the third element in e_t. After that, we compute the implied sequence of disturbances, $u_t^{\theta^*}$, $t = t_1, ..., t_2$ using (3.21). Here, the superscript θ^* highlights the dependence on the rotation parameter, θ^*. For input into our state space - observer system, (3.13)-(3.12), we require ε_t. We compute a sequence, $\varepsilon_t^{\theta^*}$, $t = t_1, ..., t_2$ using $\varepsilon_t^{\theta^*} = C^{-1} u_t^{\theta^*}$.

3.6. Empirical Results

This section documents two problems with BCA: conclusions are sensitive to modeling details and to the position one takes on spillover effects. In the first subsection we discuss the results for US postwar recessions. We then consider postwar recessions in the remaining OECD countries. Finally, we consider the US in the Great Depression.

3.6.1. US Postwar Recessions

3.6.1.1. Sensitivity of Baseline Decomposition to Modeling Details. In our analysis of the post-war US data, we examine five recessions. The 1982 recession, which is emphasized in CKM, is highlighted in the text. Details about the other post war US recessions are provided in Appendix C of Christiano and Davis (2006). Consider Table 1, which presents summary results for the 1982 US recession. The statistic reported in Table 1 is the fraction of the decline in output at the recession trough which is accounted for by the intertemporal wedge. The trough of the recession is defined as the quarter when detrended output achieves its minimum value. Panel 1a displays results based on the CKM specification of the intertemporal wedge (i.e., τ_{xt}) and Panel 1b displays results for the alternative specification (τ_t^k). In addition, results based on the baseline and rotation decompositions and with and without investment adjustment costs are reported. Finally, the table shows the impact of including CKM measurement error at the estimation stage of computing the wedges.

Turning to the CKM version of the wedge in Panel 1a we see that, regardless of whether measurement error is included in the analysis, adjustment costs make a substantial difference. Without investment adjustment costs, the intertemporal wedge contributes essentially nothing to the decline in output (or investment) in the 1982 recession. With adjustment costs, the intertemporal wedge accounts for roughly 30 percent of the decline in output at the trough of that recession. Evidently, adjustment costs have a very large impact on inference. At the same time, the impact of measurement error is nil, when we work with the CKM version of the intertemporal wedge.

Turning to the alternative specification of the intertemporal wedge, in Panel 1b we see that measurement error now matters a great deal. For example, with no measurement error and with adjustment costs, the intertemporal wedge accounts for over half the decline in output at the trough of the 1982 recession. With measurement error, that number falls to a much smaller (though still substantial!) 22 percent. The first column in the table shows that the CKM measurement error specification is strongly rejected by a likelihood ratio test whether or not adjustment costs are included in the analysis. So, the likelihood directs us to pay attention to the results without measurement error.

The results in Panel 1b show how much the specification of the intertemporal wedge matters. When CKM measurement error is used and there are no adjustment costs, the alternative formulation of the intertemporal wedge accounts for a substantial 24 percent of the drop in output at the trough of the 1982 recession. This stands in sharp contrast with the nearly zero percent drop implied by the CKM measure of that wedge. Interestingly, with the alternative measure of the wedge and with CKM measurement error, adjustment costs matter very little. When we set measurement error to zero (inducing a very large jump in the likelihood!) then adjustment costs matter a great deal, even with the alternative specification of the intertemporal wedge.

A more complete representation of our findings is reported in Figure 3, which displays results for the baseline decomposition of US data in the 1982 recession. To save space, Figure 3 reports results only for the alternative specification of the intertemporal wedge. The alternative version of the wedge is of special interest because of its conformity with the model in BBG.

In Figure 3, the circles indicate the zero line. The line with diamonds indicates the evolution of the data in response to all the wedges. By construction, the line with diamonds corresponds to the actual (detrended) data. The line marked with stars indicates the baseline decomposition when we estimated the model with the CKM specification of measurement error. The left column of graphs indicates results based on setting adjustment costs in investment to zero (i.e., $\Phi = 0$). The right column of graphs indicates results based on setting adjustment costs in investment to a level which implies a Tobin's q elasticity of unity. Note that for results based on estimation using the CKM measurement error specification, the intertemporal wedge accounts for relatively little of the movement in output, investment, hours worked and consumption. This conclusion is not sensitive to the introduction of adjustment costs in investment.

The line in Figure 3 indicated by pluses displays results based on estimation with measurement error set to zero. In the left column, we see that if the only wedge that had been active in the 1982 recession had been the intertemporal wedge, the US economy would have experienced a substantial boom (this can also be seen in Table 1). Investment would have been massively above trend, and consumption would have been massively below trend. These results show how sensitive BCA can be to seemingly minor details. Measurement error is very small under the CKM measurement error specification, yet it has a large impact on the outcome of BCA.

Measurement error also has a big impact on the assessment of the importance of adjustment costs. Comparing results in the left and right columns of Figure 3, we see that when measurement error is set to zero in estimation, then adjustment costs make a big difference to the assessment of the importance of the intertemporal wedge. The boom in output produced by the intertemporal wedge in the absence of adjustment costs becomes a recession when adjustment costs are turned on. As noted above, with adjustment costs the intertemporal wedge accounts for a very substantial 52 percent of the drop in output at the trough of the 1982 recession.

Results for four other US postwar recessions are presented in the appendix of Christiano and Davis (2006), and they generally support our findings for the 1982 recession: BCA results sensitive to the position taken on measurement error, the specification of the intertemporal wedge and on adjustment costs in investment.

3.6.1.2. The Potential Importance of Spillovers. The evidence for the 1982 recession in Figure 3 and for the other recession episodes is that the intertemporal wedge, when it has any impact at all, drives consumption and investment in opposite directions. At first, this may seem damaging to the proposition that shocks which drive the intertemporal wedge are important in business cycles, because consumption and investment are both procyclical in the data. This section shows that the opposite-signed response of consumption and investment is simply an artifact of ignoring spillover effects. Once spillover effects are taken into account, the evidence from BCA is consistent with consumption and investment responding with the same sign to an intertemporal wedge shock.

We quantify the potential importance of spillover effects by considering our rotation decomposition, discussed in section **??**. Table 1 indicates that the intertemporal wedge accounts for almost the whole of the 1982 recession under the rotation decomposition, under almost all model perturbations. The one exception occurs in the case of no measurement error, no adjustment costs and τ_t^k intertemporal wedge.

We can see the results more completely for the alternative representation of the wedge, in Figure 4 (from here on, only results for the alternative representation of the wedge are presented). The left column of that figure reproduces the results of CKM's baseline decomposition from Figure 3. The right column displays the results based on the rotation decomposition. All results in Figure 4 are based on setting measurement error to zero. This is consistent with our remarks above, according to which CKM's measurement error specification has no a priori appeal, and it is overwhelmingly rejected in the post war data.

What we see in the right column of Figure 4 is that the estimated financial shock accounts for nearly the whole of the 1982 recession. Also, the financial shock drives consumption and investment in the same directions. This reflects the operation of spillover effects. We stress that the likelihood of the model on which the results in the left and right columns are based is the same. BCA provides no way to select between the two.

3.6.2. OECD Postwar Recessions

The results for postwar recessions in OECD countries for which we have data are summarized in Table 2, panel A (no adjustment costs) and Table 2, panel B (adjustment costs). For each country the entry represents the average of a statistic over all the recessions for which we have data. The statistic is the fraction of the decline in output in the trough of a recession, due to the intertemporal wedge. This is measured, as indicated in the table, according to the baseline or rotation decomposition.[26] In each panel, the bottom row is the weighted mean of the corresponding column entries. The weight for a given country is proportional to the number of recessions in that country's data.[27]

[26]The numbers for the United States are different from what is reported in Table 1, because all results in Table 2 are based on P and Q matrices with the zero restrictions indicated in (3.7). In addition, the numbers in Table 2 reflect an average over all recessions in the sample for each country, while Table 1 only pertains to the 1982 recession.

[27]For Belgium, we only have data for the 1990 recession; for Canada, the 1980 and 1990 recessions; for Denmark, the 1990 recession; for Finland, the 1974 and 1990 recessions; for France, the 1980 and 1990 recessions; for Germany, the 1990 recession; for Italy, the 1980 and 1990 recessions; for Japan, the 1990 recession; for Mexico, the 1990 recession; for Holland, the 1980 and 1990 recessions; for Norway, the 1990 recession; for Spain, the 1974 and 1990 recessions; for Switzerland, the 1990 recession; for the UK, the 1974, 1980, and 1990 recessions.

Consider first the case where the BCA methodology is closest to CKM, i.e., the case with measurement error, no investment adjustment costs and the baseline wedge decomposition.[28] Note that there are numerous countries with fractions that are well above zero. Some are even above unity, which means that when the intertemporal wedge is fed to the RBC model, the model on average predicts bigger recessions than actually occurred. Overall, the average contribution of the intertemporal wedge to the fall in output in a trough is a substantial 22 percent.

As we found for the United States in the 1982 recession, when we then drop measurement error we find that the intertemporal wedge on average predicts an output boom in the OECD recessions for which we have data (Panel A, right portion). Although the measurement error used in the analysis is quite small, the outcome of BCA is evidently very sensitive to it.

Now consider what happens when we introduce adjustment costs, in Panel B. When we include measurement error in the analysis, there are several countries in which the intertemporal wedge plays a substantial role in recessions. However, there are several where the intertemporal wedge actually predicts a significant boom. As a result, the average contribution of the intertemporal wedge to business cycles across all countries is now about zero. When we now drop measurement error, the importance of the wedge jumps substantially for several countries. For example, it jumps from 15 percent to 46 percent in the United States and 33 percent to 75 percent in Canada. Some, however, such as Switzerland, go from 31 percent to -14 percent when measurement error is dropped. As a result, the overall average is a more modest jump of 16 percent.

[28]However, recall that we now consider the alternative type of wedge, the τ_t^k wedge motivated by the CF and BGG models.

Turning to the rotation wedge, we see that under that decomposition, the intertemporal wedge assumes a very large role in most countries. It is logically possible that the entire effect of this substantial importance assigned to financial shocks is due to spillover effects. In this case, one might be tempted to conclude that these are not actually shocks to the intertemporal wedge itself, and are better thought of as shocks to other wedges. To investigate this, we computed the ratio of the variance in HP filtered output due only to the spillover effects of financial shocks, to the total variance in HP filtered output due to financial frictions. This ratio is reported in the column, 'ratio', in Table 2. Note that in the case of no measurement error and investment adjustment costs, the ratio is only 30 percent for the US. Evidently, in US business cycles, the great importance assigned to financial shocks is not coming primarily from spillover effects. In other countries, the ratio is greater than unity, suggesting that spillovers are substantial (see Belgium, Germany and the UK). However, on average the ratio is only 60 percent, suggesting that the financial shocks identified in our rotation decomposition operate on the economy primarily by their direct impact on the intertemporal wedge.

We conclude that our findings for the postwar US also hold up on average across the other countries in the OECD.

3.6.3. US Great Depression

We now consider results for the US Great Depression. In this episode, the data exhibit substantial fluctuations and so it is perhaps not surprising that there is evidence of inaccuracy in the linear approximation of our model's solution. To quantify the degree approximation error we first estimate the capital stock for each date in the sample, by a two-sided Kalman projection using the state-space representation of our model.[29] This, together with the realized wedges for each date, provided us with an estimate of the model's state for each date in the sample. Then, for each t we used the approximate policy rule to compute $(c_t, k_{t+1}, l_t, x_t, y_t)$ as a function of the date t state. We then computed the percent change in each of these 5 variables required for the four equilibrium conditions, (3.1), (3.2), (3.3), (3.8), plus the production function to be satisfied as a strict equality at t. For each t these calculations were done under the assumption that the period $t+1$ decisions are made using the approximate solution. Figure 5 shows that outside the 1930s, the approximation error associated with the linearized policy rule is for the most part fairly small. In the period of the 1930s, however, the approximation error becomes large, briefly reaching 65 percent for investment. We report the same measure of approximation error for the second order approximation to the model solution. In this case, the approximation errors are considerably smaller. Because of this evidence that the first order approximation has substantial approximation error, and because the second order approximation appears to be noticeably more accurate, we only display results for the Great Depression based on the second order approximation.

Consider the results in Figure 6. The left column displays the baseline decomposition and shows that the intertemporal wedge accounts for a substantial 21 percent of the fall in output in the Great Depression. In addition, that wedge drives consumption and investment in opposite directions. When we allow for spillovers using the rotation decomposition, we find that financial shocks may account for as much as 92 percent

[29] Essentially, this involves using measured investment to compute the capital stock using the capital accumulation equation.

of the fall in output at the trough of the Great Depression.[30] Moreover, shocks to the intertemporal wedge drive consumption and investment in the same direction. We also did the calculations using the CKM measurement error and the results appear in Figure A9 in Appendix C of Christiano and Davis (2006). The results reported there are qualitatively similar to what emerges from Figure 6.

We conclude that results for the Great Depression are consistent with the findings for the postwar period. Taken as a whole, the evidence from BCA is consistent with the proposition that shocks to the intertemporal wedge play a significant role in business fluctuations.

3.7. Conclusion

Chari, Kehoe and McGrattan (2006) advocate the use of business cycle accounting to identify directions for improvement in equilibrium models. As a demonstration of the power of the approach, they argue that BCA can be used to rule out a prominent class of financial friction models. In particular, they conclude that models of financial frictions which create wedges in the intertemporal Euler equation are not promising avenues for understanding business cycle dynamics.

We have described two flaws in BCA which undermine its usefulness. First, consistent with economic theory, the results of BCA are not robust to small changes in the modeling environment. Second, BCA necessarily misses key mechanisms by which financial shocks which drive the intertemporal wedge affect the economy. The empirical correlations among wedges are consistent with the possibility that the financial shocks which

[30]Given the nonlinearity of the model, we could not compute the rotation decomposition as we did for postwar data. Instead, we computed the rotation that minimized the sum of squared deviations between the actual data and the predicted data using the estimated wedges.

drive the intertemporal wedge have important spillover effects on other wedges. These spillover effects are not identified under BCA. However, spillover effects are potentially so important that the evidence is consistent with the proposition that financial shocks are the major driving force in postwar recessions in the US and many OECD countries, as well as in the US Great Depression.

Fortunately, there are alternative ways to investigate whether given model features are useful in business cycle analysis. An approach which does not involve so many of the detailed model assumptions used by BCA, but which does incorporate the sort of assumptions needed to identify spillover effects, uses vector autoregressions.[31] An alternative approach works with fully specified, structural models. With the recent advances in computational technology and in economic theory, exploration of alternative models is relatively costless. A full set of references to the literature that explores the sort of financial frictions which are the object of interest in CKM would be too lengthy to include here. See Carlstrom and Fuerst (1997), Bernanke, Gertler and Gilchrist (1999), Christiano, Motto and Rostagno (2004, 2006) and Queijo (2005), and the references they cite.

Another approach uses a natural way to confront one of the identification problems with BCA. Absent direct observations, it is difficult to identify the intertemporal wedge and the rate of return of capital separately. However, as stressed in Cochrane and Hansen (1992), rates of return are the one type of economic variable on which we have excellent observations. For example, rates of return do not have the problems of interpretation associated with wages and they do not have the measurement error problems associated

[31]For a recent review, see Christiano, Eichenbaum and Vigfusson (2006).

with observations on quantities like consumption and investment. The recent work of Primiceri, Schaumburg and Tambalotti (2005) carries out an analysis that is similar to business cycle accounting, except that they make use of direct measures of rates of return. They find that the estimates of τ_t^k (which they call 'preference shocks') assign that variable an important role in business cycle fluctuations.[32] A related approach is taken recently in Christiano, Motto and Rostagno (2006), who also include rates of return in the analysis. In addition, they integrate an explicit model of financial frictions and so are able to relate τ_t^k directly to primitive, uncorrelated financial shocks. When they feed the individual shocks to the model, holding other shocks fixed, they find that the financial shocks are an important driving force in business cycles.

[32]This approach is related to that of Hansen and Singleton (1982, 1983).

References

[1] NOTE: Items 2-46 correspond to Chapters 1 and 2, 47-77 to Chapter 3

[2] Alexopoulos, M. (2004): "Read All about it!! What Happens Following a Technology Shock?," mimeo, University of Toronto, 2004.

[3] Ang, A. and M. Piazzesi (2003): "A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables," *Journal of Monetary Economics*, 50, 745-787.

[4] Arouba, B., F. Diebold and G. Rudebusch (2006): "The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach," Journal of Econometrics, 131, 309-338.

[5] Banerjee, A. (1992): " A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 107, 797-817

[6] Bekaert, G., S Cho and A. Moreno (2006): "New-Keynesian Macroeconomics and the Term Structure," mimeo, Columbia University

[7] Beaudry, Paul and F. Portier (2004): "An Exploration into Pigou's Theory of Cycles," *Journal of Monetary Economics*, 51: 1183-1216.

[8] Beaudry, P. and F. Portier (2006): "News, Stock Prices and Economic Fluctuations," Working Paper 10548, National Bureau of Economic Research, forthcoming *American Economic Review*.

[9] Bikbov, R. and M. Chernov (2005), "No -Arbitrage Macroeconomic Determinants of the Yield Curve," Working Paper, Columbia University

[10] Bussie, M. and C. Mulder (2000), "Political Instability and Economic Vulnerability," *International Journal of Finance and Economics* 5: 309-330

[11] Calvo, G. (1983), "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics 12*, 383-398

[12] Campbell, J.Y. and R.J. Shiller (1991), "Yield Spreads and Interest Rate Movements: A Bird's Eye View," *Review of Economic Studies*, 57, 495-514.

[13] Caplin, A., and J. Leahy (1993): "Sectoral Shocks, Learning and Aggregate Fluctuations," *Review of Economic Studies*, 60, 1065-85

[14] Christiano, L., M. Eichenbaum and C. Evans (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113: 1-45, 2005.

[15] Clarida, R., J. Gali and M. Gertler (2000), "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics*, vol. CXV, issue 1, 147-180

[16] Christiano, L., C. Ilut, R. Motto and M. Rostagno (2005): "Monetary Policy and Stock Market Boom-Bust Cycle," mimeo, Northwestern University.

[17] Christiano, L., R. Motto and M. Rostagno (2007): "Shocks, Structures, of Monetary Policies? The EA and US after 2001," mimeo, Northwestern University

[18] Cochrane, J. (1994): "Shocks," *Carnegie-Rochester Conference Series on Public Policy*, 41, 295-364.

[19] Cox, J.C., J.E. Ingersoll Jr. and S.A. Ross (1985), "A Theory of the Term Structure of Interest Rates in General Equilibrium," *Econometrica* 53, 385-408.

[20] Dejon, D. B. Ingram and C. Whiteman (2000), "A Bayesian Approach to Dynamic Macroeconomics," *Journal of Econometrics*, 98(2), 203-223

[21] Del Negro, M., F. Schorheide, F. Smets and R. Wouters (2007), "On the Fit and Forecasting Performance of New-Keynesian Models," *Journal of Business and Economic Statistics* 25(2), 123-162.

[22] Erceg, C., D. Henderson and A. Levin (2000), "Optimal Monetary Policy with Staggered Wage and Price Contracts," Journal of Monetary Economics 46, 281-313

[23] Engel, C. and H. Rogers (2006), "The U.S. Current Account Deficit and Expected Share of World Output," *Journal of Monetary Economics* 53, 1063-1093.

[24] Estrella, A. and G.A. Hardouvelis (1991): "The Term Structure as a Predictor of Real Economic Activity," Journal of Finance, 46, 555-576

[25] Estrella, A. and F.S. Mishkin (1998): "Predicting U.S. Recessions: Financial Variables as Leading Indicators," *Journal of Finance*, 46, 555-576.

[26] Evans, C.L. and D. Marshall (1998): "Monetary Policy and the Term Structure of Nominal Interest Rates: Evidence and Theory," Carnegie-Rochester Conference Series on Public Policy, 49, 53-111

[27] Fama, E. and R. Bliss (1987), "The Information in Long-Maturity Forward Rates," *American Economic Review* 77, 680-692.

[28] Francis, N. and V. Ramey (2006), "Measures of Hours Per-Capita and their Implications for the Technology-Hours Debate," mimeo, University of California, San Diego.

[29] Fisher, J. (2006): "The Dynamic Effects of Neutral and Investment-Specific Technology Shocks," *Journal of Political Economy*, 114: 413-451, 2006.

[30] Geweke, J.F. (1992), Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments," in J.O. Berger, J.M. Bernardo, A.P. Dawid and A.F.M. Smith (editors), *Proceedings of the Fourth Valencia International Meeting on Bayesian Statistics*, Oxford University Press, 169-194

[31] Harrison, M. and D. Kreps (1979), "Martingales and Arbitrage in Multiperiod Security Markets," *Journal of Economic Theory* 20, 381-408.

[32] Hordahl, P., O. Tristani and D. Vestin (2002), "A Joint Econometric Model of Macroeconomic and Term Structure Dynamics," mimeo, European Central Bank

[33] Jaimovich, N. and S. Rebelo (2006), "Do News Shocks Drive the Business Cycle?," Northwestern University, mimeo

[34] Kadiyala, K. and S Karlsson (1997), "Numerical Methods for Estimation and Inference in Bayesian VAR-Models," *Journal of Applied Econometrics*, vol. 12(2), 99-132

[35] King, R., C Plosser and S. Rebelo (1988), " Production, Growth and Business Cycles: I. the Bastic Neoclassical Model," *Journal of Monetary Economics* 21, 195-232.

[36] Kydland, F.E. and E. Prescott (1982), "Time to Build and Aggregate Fluctuations," *Econometrica* 50, 1345-1370.

[37] Litterman, R.B. (1986), "Forecsating with Bayesian Vector Autoregressions–Five Years of Experience," *Journal of Business, Economics and Statistics* 4, 25-38.

[38] Otrok, C. (2001), "On Measuring the Welfare Cost of Business Cycles," *Journal of Monetary Economics*, vol 47, 61-92

[39] Pigou, A. (1926): *Industrial Fluctuations*. MacMillan, London.

[40] Rotemberg, J. (2003): "Stochastic Technical Progress, Smooth Trends, and Nearly Distinct Business Cycles," *American Economic Review* 93: 1543-1559, 2003.

[41] Schmitt-Grohe, S. and M Uribe, "Optimal Operational Monetary Policy in the Christiano-Eichenbaum-Evans Model of the U.S. Business Cycle," mimeo, Duke University

[42] Smets, F. and R. Wouters (2004), "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," European Central Bank, mimeo

[43] Stock, J.H. and M.W. Watson (1999): "Business Cycle Fluctuations in US Macroeconomic Time Series," in Woodford and Taylor (eds.) *Handbook of Macroeconomics*, Amsterdam, Holland, 1999.

[44] Uhlig, H. (2004), "Do Technology Shocks lead to a fall in Total Hours Worked?" Journal of the European Economic Association, 361-371 (2004)

[45] Wu, T. (2002): "Monetary Policy and the Slope Factors in Empirical Term Structure Estimations," Federal Reserve Bank of San Francisco Working Paper 02-07

[46] Zeira, J. (1994): "Informational Cycles," *Review of Economic Studies*, 61(1) 31-44

[47] Abel, Andrew, 1980, "Empirical investment equations: An integrative framework," Carnegie-Rochester Conference Series on Public Policy, vol. 12, pp. 39–91.

[48] Bernanke, Ben, Mark Gertler, and Simon Gilchrist, 1999, "The Financial Accelerator in a Quantitative Business Cycle Framework," *Handbook of Macroeconomics*, edited by John B. Taylor and Michael Woodford, pp. 1341-1393, Amsterdam, New York and Oxford, Elsevier Science, North-Holland.

[49] Boldrin, Michele, Lawrence J. Christiano and Jonas Fisher, 2001, 'Habit Persistence, Asset Returns and the Business Cycle', *American Economic Review*, vol. 91, no. 1, pp. 149-166. March.

[50] Carlstrom, Charles, and Timothy Fuerst, 1997, "Agency Costs, Net Worth, and Business Fluctuations: A Computable General Equilibrium Analysis," *American Economic Review*, 87, 893-910.

[51] Chari, V.V., Patrick Kehoe and Ellen McGrattan, 2006, "Business Cycle Accounting," Federal Reserve Bank of Minneapolis Staff Report 328, revised February.

[52] Christiano, Lawrence J., 2002, 'Solving Dynamic Equilibrium Models by a Method of Undetermined Coefficients,' *Computational Economics,* October, Volume 20, Issue 1-2.

[53] Christiano, Lawrence J., Martin Eichenbaum and Charles Evans, 2005, 'Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy', *Journal of Political Economy.*

[54] Christiano, Lawrence J., Martin Eichenbaum and Robert Vigfusson, 2006, 'Assessing Structural VARs,' forthcoming, *Macroeconomics Annual*, edited by George-Marios Angeletos, Kenneth Rogoff, and Michael Woodford.

[55] Christiano, Lawrence J., Roberto Motto and Massimo Rostagno, 2004, 'The US Great Depression and the Friedman-Schwartz Hypothesis,' *Journal of Money, Credit and Banking.*

[56] Christiano, Lawrence J., Roberto Motto and Massimo Rostagno, 2006, 'Financial Factors in Business Cycles,' manuscript.

[57] Christiano, Lawrence J., and Joshua Davis, 2006, 'Two Flaws in Business Cycle Accounting,' NBER Working paper.

[58] Cochrane, John H., and Lars Peter Hansen, 1992, 'Asset pricing explorations for macroeconomics,' *NBER Macroeconomics Annual*, pp. 115–165.

[59] Eberly, Janice C., 1997, 'International Evidence on Investment and Fundamentals,' European Economic Review, vol. 41, no. 6, June, pp. 1055-78.

[60] Erceg, Christopher J., Henderson, Dale, W. and Andrew T. Levin, 2000, 'Optimal Monetary Policy with Staggered Wage and Price Contracts,' *Journal of Monetary Economics*, 46(2), October, pages 281-313.

[61] Fernandez-Villaverde, Jesus, Juan F. Rubio-Ramirez and Thomas J. Sargent, 2006, 'A, B, C's (and D)'s For Understanding VARs,' manuscript.

[62] Gali, Jordi, Mark Gertler and J. David Lopez-Salido, 2003, 'Markups, Gaps, and the Welfare Costs of Business Fluctuations,' forthcoming, *Review of Economics and Statistics.*

[63] Hall, Robert, 1997, 'Macroeconomic Fluctuations and the Allocation of Time,' *Journal of Labor Economics*, 15 (part 2) , S223-S250.

[64] Hamilton, James B. (1994). Time Series Analysis. Princeton: Princeton University Press.

[65] Ingram, Beth, Narayana Kocherlakota and N. Savin, 1994, "Explaining Business Cycles: A Multiple-Shock Approach," *Journal of Monetary Economics*, 34, 415-428.

[66] Lippi, Marco, and Lucrezia Reichlin, 1993, 'The Dynamic Effects of Aggregate Demand and Supply Disturbances: Comment,' *The American Economic Review,* Vol. 83, No. 3 (June), pp. 644-652.

[67] Lucca, David, 2006, 'Resuscitating Time to Build,' manuscript, Board of Governors of the Federal Reserve.

[68] Matsuyama, Kiminori, 1984, 'A Learning Effect Model of Investment: An Alternative Interpretation of Tobin's Q', unpublished manuscript, Northwestern University.

[69] Mulligan, Casey, 2002, 'A Dual Method of Empirically Evaluating Dynamic Competitive Models with Market Distortions, Applied to the Great Depression and World War II,' National Bureau of Economic Research Working Paper 8775.

[70] Parkin, Michael, 1988, 'A Method for Determining Whether Parameters in Aggregative Models are Structural,' *Carnegie-Rochester Conference Series on Public Policy*, 29, 215-252.

[71] Queijo, Virginia, 2005, "How Important are Financial Frictions in the US and the Euro Area?", Seminar paper no. 738, Institute for International Economic Studies, Stockholm University, August.

[72] Rosen, Sherwin and Robert Topel, 1988, 'Housing Investment in the United States,' *Journal of Political Economy*, Vol. 96, No. 4 (Aug.), pp. 718-740

[73] Sargent, Thomas, 1989, 'Two Models of Measurements and the Investment Accelerator,' *Journal of Political Economy*, vol. 97, issue 2, April, pp. 251-287.

[74] Schmitt-Grohe, Stephanie, and Martin Uribe, 2004, 'Solving Dynamic General Equilibrium Models Using a Second-Order Approximation to the Policy Function,' *Journal of Economic Dynamics and Control*, 28.

[75] Sims, Christopher A. and Tao Zha, 1996, 'Does Monetary Policy Generate Recessions?', forthcoming as a 'vintage paper' in *Macroeconomic Dynamics*.

[76] Uhlig, Harald, 2002, 'What Moves Real GNP?', manuscript, Humboldt Unversity Berlin.

[77] Wan, E.A. and R. van der Merwe, 2001, "Kalman Filtering and Neural Networks", chap. Chapter 7 : *The Unscented Kalman Filter*, (50 pages), Wiley Publishing, Eds. S. Haykin.

APPENDIX A

Chapter 1

A.0.1. Bayesian Vector-Autoregression

Bayesian VAR's have been shown to provide superior forecasting perfomance relative to their non-Bayesian counterparts. In addition, they are often used to avoid problems such as collinearity and over-paramaterization. The BVAR is estimated using a rolling sample of 50 quarterly observations and 3 lags. The use of a relatively small window allows for agility in the way the model captures agents' expectations. This sample size, roughly 4 observations per coefficient, does create some problems since there is a reasonable amount of sampling uncertainty in the parameter estimates. The prior used in this analysis corresponds to a random walk process[1], further focusing the parameters governing agents' expectations on current values. The prior shrinks the parameters of the BVAR toward those of an independent unit root process consistent with the behavior of many economic time series. Letting $\rho_{i,j}^{k}$ represent the coefficient of the k^{th} lag of variable j in equation i, the BVAR makes the following assumption about the variance of the autoregressive coefficients:

$$
\begin{aligned}
var(\rho_{i,j}^{k}) &= \begin{array}{c} \frac{\pi_1}{k^2} \text{ if } i = j \\ \frac{\pi_2 \widehat{\sigma}_i^2}{k^2 \widehat{\sigma}_j^2} \text{ if } i \neq j \end{array} \\
var(const_i) &= \pi_0 \widehat{\sigma}_i^2
\end{aligned}
$$

[1]This prior is known as the Minnesotta prior and was first introduced by Litterman (1986)

Here the variable $\widehat{\sigma}_i$ corresponds to the estimated standard error of the residuals in a regression of series i on 3 lags of itself[2]. I follow Kadiyala and Karlsson (1997) by setting $\pi_0 = 1e4$, $\pi_1 = 0.8$ and π_2 to 0.0016.

A.0.2. Scaling and Equilibrium

The following transformation places the DSGE into stationary form

$$
\begin{aligned}
\widetilde{\lambda}_t &= z_t \lambda_t \\
\widetilde{k}_{t+1} &= k_{t+1}/z_t \\
\widetilde{c}_t &= c_t/z_t \\
\widetilde{i}_t &= i_t/z_t \\
\widetilde{m}_t^h &= m_t^h z_t^{-1/\sigma_m} \\
\widetilde{f}_t^1 &= z_t f_t^1 \\
\widetilde{x}_t^1 &= x_t^1/z_t \\
\widetilde{x}_t^2 &= x_t^2/z_t \\
\widetilde{w}_t &= w_t/z_t \\
\widetilde{\overline{w}}_t &= \overline{w}_t/z_t
\end{aligned}
$$

[2]Litterman proposes the use of this prior to account for time-series of different scale

Note that this scaling implies the following equivalence in the observation equation of the state space system

$$\begin{aligned}\log\left(\frac{y_t}{y_{t-1}}\right) &= \log\left(\frac{y_t}{z_t}\frac{z_{t-1}}{y_{t-1}}\frac{z_t}{z_{t-1}}\right) \\ &= \log(\widetilde{y}_t) - \log(\widetilde{y}_{t-1}) + \log(\chi_t)\end{aligned}$$

Demeaning the left hand side (the data) and replacing the variables on the right hand side with their 'hat' counterparts (corresponding model parameters) provides the mapping used in the estimation of the model. All other variables used in the observation equation of the model correspond to their model counterparts in the conventional manner.

The following equations correspond to a stationary competitive symmetric equilibrium

$$\widetilde{k}_{t+1} = (1-\delta)\frac{\widetilde{k}_t}{\chi_t} + \tau_t^k \widetilde{i}_t \left[1 - S(\frac{\widetilde{i}_t}{\widetilde{i}_{t-1}}\frac{1}{\chi_t})\right]$$

$$\frac{1}{\widetilde{c}_t - b\widetilde{c}_{t-1}/\chi_t} - E_t\left[\frac{b}{\widetilde{c}_{t+1}\chi_{t+1} - b\widetilde{c}_t}\left(\beta\frac{\phi_{\beta,t+1}}{\phi_{\beta,t}}\right)\right] = \widetilde{\lambda}_t$$

$$\phi_0 h_t = \frac{\widetilde{\lambda}_t \widetilde{w}_t}{\bar{\mu}_t}$$

$$\widetilde{\lambda}_t q_t = E_t\left[\left(\beta\frac{\phi_{\beta,t+1}}{\phi_{\beta,t}}\right)\frac{1}{\chi_{t+1}}\widetilde{\lambda}_{t+1}\left[r_{t+1}^k u_{t+1} - a(u_{t+1}) + q_{t+1}(1-\delta)\right]\right]$$

$$\begin{aligned}\widetilde{\lambda}_t &= \widetilde{\lambda}_t q_t \tau_t^k \left[1 - S\left(\frac{\widetilde{i}_t}{\widetilde{i}_{t-1}}\chi_t\right) - \frac{\widetilde{i}_t}{\widetilde{i}_{t-1}}\chi_t S'\left(\frac{\widetilde{i}_t}{\widetilde{i}_{t-1}}\chi_t\right)\right] + ... \\ & ... + E_t\left[\left(\beta\frac{\phi_{\beta,t+1}}{\phi_{\beta,t}}\right)\frac{1}{\chi_{t+1}}\widetilde{\lambda}_{t+1} q_{t+1}\tau_{t+1}^k\left(\frac{\widetilde{i}_{t+1}}{\widetilde{i}_t}\chi_{t+1}\right)^2 S'\left(\frac{\widetilde{i}_{t+1}}{\widetilde{i}_t}\chi_{t+1}\right)\right]\end{aligned}$$

$$r_t^k = a'(u_t)$$

$$\widetilde{\lambda}_t = \phi_1\left(\widetilde{m}_t^h\right)^{-\sigma_m} + E_t\left[\left(\beta\frac{\phi_{\beta,t+1}}{\phi_{\beta,t}}\right)\frac{\widetilde{\lambda}_{t+1}}{\chi_{t+1}\pi_{t+1}}\right]$$

$$\widetilde{f}_t^1 = \widetilde{\lambda}_t\left(\frac{\widetilde{w}_t}{\widetilde{\bar{w}}_t}\right)^{\bar{\eta}} h_t^d + E_t\left[\left(\beta\frac{\phi_{\beta,t+1}}{\phi_{\beta,t}}\right)\frac{\bar{\alpha}}{\chi_{t+1}}\left(\frac{\pi_{t+1}}{\pi_t^{\bar{\chi}}}\right)^{\bar{\eta}-1}\left(\frac{\widetilde{\bar{w}}_{t+1}}{\widetilde{\bar{w}}_t}\chi_{t+1}\right)^{\bar{\eta}}\widetilde{f}_{t+1}^1\right]$$

$$f_t^2 = \frac{\widetilde{\lambda}_t}{\bar{\mu}_t}\widetilde{w}_t\left(\frac{\widetilde{w}_t}{\widetilde{\bar{w}}_t}\right)^{\bar{\eta}} h_t^d + E_t\left[\bar{\alpha}\left(\beta\frac{\phi_{\beta,t+1}}{\phi_{\beta,t}}\right)\left(\frac{\pi_{t+1}}{\pi_t^{\bar{\chi}}}\right)^{\bar{\eta}}\left(\frac{\widetilde{\bar{w}}_{t+1}}{\widetilde{\bar{w}}_t}\chi_{t+1}\right)^{\bar{\eta}} f_{t+1}^2\right]$$

$$0 = (1-\bar{\eta})\widetilde{\bar{w}}_t\widetilde{f}_t^1 + \bar{\eta} f_t^2$$

$$\widetilde{\lambda}_t = R_t E_t\left[\left(\beta\frac{\phi_{\beta,t+1}}{\phi_{\beta,t}}\right)\frac{\widetilde{\lambda}_{t+1}}{\chi_{t+1}\pi_{t+1}}\right]$$

$$\widetilde{y}_t = \widetilde{c}_t + \widetilde{i}_t + \widetilde{g}_t + a(u_t)\widetilde{k}_t/\chi_t$$

$$\widetilde{x}_t^1 = \widetilde{y}_t mc_t \bar{p}_t^{-\eta-1} + \alpha E_t \left[\left(\beta \frac{\phi_{\beta,t+1}}{\phi_{\beta,t}} \right) \frac{\widetilde{\lambda}_{t+1}}{\widetilde{\lambda}_t} (\bar{p}_t/\bar{p}_{t+1})^{-\eta-1} (\frac{\pi_t^{\chi}}{\pi_{t+1}})^{-\eta} \widetilde{x}_{t+1}^1 \right]$$

$$\widetilde{x}_t^2 = \widetilde{y}_t \bar{p}_t^{-\eta} + \alpha E_t \left[\left(\beta \frac{\phi_{\beta,t+1}}{\phi_{\beta,t}} \right) \frac{\widetilde{\lambda}_{t+1}}{\widetilde{\lambda}_t} \left(\frac{\pi_t^{\chi}}{\pi_{t+1}} \right)^{1-\eta} (\frac{\bar{p}_t}{\bar{p}_{t+1}})^{-\eta} \widetilde{x}_{t+1}^2 \right]$$

$$\eta \widetilde{x}_t^1 = (\eta - 1) \widetilde{x}_t^2$$

$$\widehat{R}_t = a_\pi E_t \left[\widehat{\pi}_{t+1} - \widehat{\pi}_t^* \right] + a_y E_t \widehat{y}_t + \alpha_R \widehat{R}_{t-1} + mp_t$$

$$1 = \alpha \pi_t^{\eta-1} \pi_{t-1}^{\chi(1-\eta)} + (1-\alpha) \bar{p}_t^{1-\eta}$$

$$\left(u_t \widetilde{k}_t / \chi_t \right)^{\theta} \left(h_t^d \right)^{1-\theta} - \psi = (\widetilde{c}_t + \widetilde{i}_t + \widetilde{g}_t + a(u_t) \widetilde{k}_t / \chi_t) s_t$$

$$s_t = (1-\alpha) \bar{p}_t^{-\eta} + \alpha (\frac{\pi_t}{\pi_{t-1}^{\chi}})^{\eta} s_{t-1}$$

$$(1-\theta)\, mc_t \left(\frac{u_t k_t}{\chi_t h_t^d} \right)^{\theta} = \widetilde{w}_t \left[1 + \nu \frac{R_t - 1}{R_t} \right]$$

$$\theta mc_t \left(\frac{u_t \widetilde{k}_t}{\chi_t h_t^d} \right)^{\theta-1} = r_t^k$$

$$\bar{h}_t = \left(\frac{\widetilde{\bar{w}}_t}{\widetilde{w}_t} \right)^{-\bar{\eta}} h_t^d$$

$$h_t = \bar{s}_t h_t^d$$

$$\bar{s}_t = (1-\bar{\alpha})(\frac{\widetilde{\widetilde{w}}_t}{\widetilde{w}_t})^{-\bar{\eta}} + \bar{\alpha}\left(\frac{\widetilde{w}_{t-1}}{\widetilde{w}_t \chi_t}\right)^{-\bar{\eta}}(\frac{\pi_t}{\pi_{t-1}^{\bar{\chi}}})^{\bar{\eta}}\bar{s}_{t-1}$$

$$\widetilde{w}_t^{1-\bar{\eta}} = (1-\bar{\alpha})\widetilde{\widetilde{w}}_t^{1-\bar{\eta}} + \bar{\alpha}\left(\frac{\widetilde{w}_{t-1}}{\chi_t}\right)^{1-\bar{\eta}}\left(\frac{\pi_{t-1}^{\bar{\chi}}}{\pi_t}\right)^{1-\bar{\eta}}$$

A.0.3. Steady-State

The following equations describe the steady state of the model.

$$x^1 = \frac{y * mc * \bar{p}^{-\eta-1}}{1-\alpha\beta\pi^{(1-\chi)\eta}}$$

$$x^2 = \frac{y * \bar{p}^{-\eta}}{1-\alpha\beta\pi^{(\chi-1)(1-\eta)}}$$

$$\bar{p} = \left(\frac{\eta}{\eta-1}\right)\left(\frac{1-\alpha\beta\pi^{(1-\chi)(\eta-1)}}{1-\alpha\beta\pi^{(1-\chi)\eta}}\right)mc$$

$$\begin{aligned} f^1 &= \left(\frac{1}{1-\bar{\alpha}\beta\pi^{(1-\chi)(\bar{\eta}-1)}\chi^{\bar{\eta}-1}}\right)\lambda\left(\frac{w}{\bar{w}}\right)^{\bar{\eta}} h^d \\ &= \Psi^1 \lambda \left(\frac{w}{\bar{w}}\right)^{\bar{\eta}} h^d \end{aligned}$$

$$
\begin{aligned}
f^2 &= \left(\frac{1}{1-\bar{\alpha}\beta\pi^{(1-\chi)\bar{\eta}}\chi^{\bar{\eta}}}\right)\frac{\lambda}{\bar{\mu}}w\left(\frac{w}{\bar{w}}\right)^{\bar{\eta}}h^d \\
&= \Psi^2\frac{\lambda}{\bar{\mu}}w\left(\frac{w}{\bar{w}}\right)^{\bar{\eta}}h^d
\end{aligned}
$$

$$\frac{w}{\bar{w}} = \left[\frac{(1-\bar{\alpha})}{1-\bar{\alpha}\pi^{(\bar{\chi}-1)(1-\bar{\eta})}\chi^{\bar{\eta}-1}}\right]^{\frac{1}{1-\bar{\eta}}}$$

$$\overline{\mu} = \left(\frac{\Psi^2}{\Psi^1}\right)\left(\frac{\bar{\eta}}{\bar{\eta}-1}\right)\frac{w}{\bar{w}}$$

$$\bar{s} = \left(\frac{1-\bar{\alpha}}{1-\bar{\alpha}\pi^{(1-\overline{\chi})\overline{\eta}}\chi^{\overline{\eta}}}\right)(\frac{w}{\bar{w}})^{\bar{\eta}}$$

$$\bar{p} = \left[\frac{1-\alpha\pi^{(1-\chi)(\eta-1)}}{1-\alpha}\right]^{\frac{1}{1-\eta}}$$

$$w = mc\left(\frac{k}{\chi h^d}\right)^{\theta}\left[\frac{1-\theta}{1+\nu\frac{R-1}{R}}\right]$$

$$\frac{k}{h^d} = \chi\left[\frac{r^k}{\theta * mc}\right]^{\frac{1}{\theta-1}}$$

$$s = \frac{(1-\alpha)\bar{p}^{-\eta}}{1-\alpha\pi^{(1-\chi)\eta}}$$

Solve for c as a function of h^d using the resource constraint

$$\begin{aligned}
\left(\frac{k}{\chi h^d}\right)^\theta h^d - \psi &= (c + \frac{i}{k}\frac{k}{h^d}h^d + g + a(u)k/\chi)s \\
c(h^d) &= \left[\frac{1}{s}\left(\frac{k}{\chi h^d}\right)^\theta - \frac{i}{k}\frac{k}{h^d}\right] h^d - \frac{\psi}{s} - g \\
&= C_1 h^d + C_2
\end{aligned}$$

Now use the wage equation to solve

$$\lambda = \left[\frac{1}{1-b/\chi} - \frac{b\beta}{\chi - b}\right]\frac{1}{c}$$

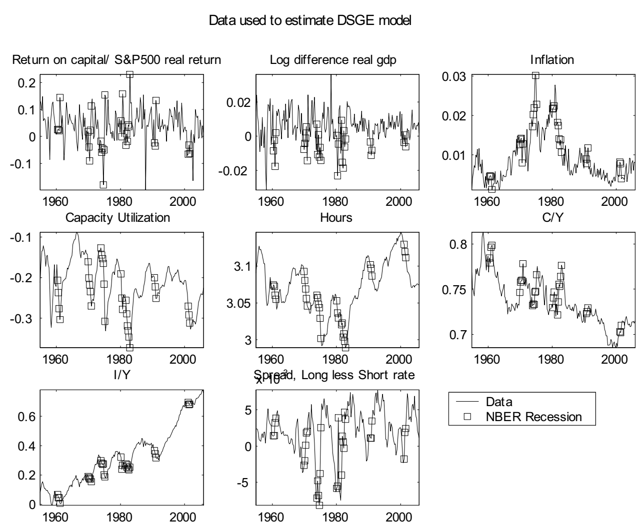
$$\phi_0 h = \frac{\lambda w}{\bar{\mu}}$$

$$\begin{aligned}
\phi_0 \overline{s} h^d c(h^d) &= \underbrace{\left(\frac{w}{\bar{\mu}}\right)\left[\frac{1}{1-b/\chi} - \frac{b\beta}{\chi - b}\right]}_{C_3} \\
\phi_0 \overline{s}\left(C_1\left(h^d\right)^2 + C_2 h^d\right) &= C_3 \\
\left[\phi_0 \overline{s} C_1\right]\left(h^d\right)^2 + \left[\phi_0 \overline{s} C_2\right] h^d - C_3 &= 0
\end{aligned}$$

A.0.4. Data

The data on consumption, investment, capacity utilization, output, inflation, population and the short term interest rate are taken from the FRED database at the Federal Reserve Bank of St. Louis. The inflation measure used in estimating the model is calculated from the GDP deflator. The short term interest rate corresponds to the Federal Funds Rate and is used to proxy for the 90-day tbill in constructing the term structure implications. The data on per-capita hours worked come from and Francis and Ramey and can be found at: http://www.econ.ucsd.edu/~vramey/research/Francis-Ramey_Hours_Data_Public.xls

The term-structure data comes from the Fama-Bliss dataset obtained via Wharton Research Data Service. It contains artificial yield data on zero coupon bonds expiring 1,2,3,4 and 5 years out constructed from a panel of existing treasury bills. A description of the data can be found in Fama and Bliss **??**. Interest rates of maturities between 3 months and 5 years are constructed via hermite interpolation of the federal funds rate and the Fama Bliss data. This results in a series of 20 different interest rates which are used in the estimation of the model.



The figure contains subplots of each time-series that is used to estimate the various versions of the model

A.0.5. Marginal Likelihood Approximation

The marginal likelihood is calculated using Geweke's modified harmonic mean. The fundemental economic impulses are assumed to be distributed gaussian and the steady-state kalman filter is used for computing the likelihood. Taking the product of the likelihood function and the priors yields the posterior distribution.

The modified harmonic mean estimator uses draws from the posterior distribution, obtained via markov-chain monte carlo starting from the mode. Robustness checks on the mode are made by running a maximization routine starting from 30 different random draws from the prior distribution. 110,000 draws are made from the posterior of which the first 10,000 are discarded. The results correspond to the remaining 100,000 draws. The modified harmonic mean estimator can be written as:

$$P(data|Model_i) \approx \left[\frac{1}{n_{sim}} \sum_{s=1}^{n_{sim}} \frac{f(\theta^{(s)})}{L(\theta^{(s)}|Y)p(\theta^{(s)})} \right]^{-1}$$

where $L(\theta^{(s)}|Y)p(\theta^{(s)})$ represents the value of the posterior at the point $\theta^{(s)}$, n_{sim} is the number of draws from the posterior and $f(\theta^{(s)})$ is given by:

$$\begin{aligned} f(\theta) &= \tau^{-1}(2\pi)^{-d/2} \left|V_\theta\right|^{-1/2} \exp\left\{ -\frac{\left(\theta - \overline{\theta}\right)' V_\theta^{-1} \left(\theta - \overline{\theta}\right)}{2} \right\} ... \\ & ... \times \left\{ \left(\theta - \overline{\theta}\right)' V_\theta^{-1} \left(\theta - \overline{\theta}\right) \leq F_{\chi_d^2}^{-1}(\tau) \right\} \end{aligned}$$

Note that τ is a hyperparamater that determines a cutoff for points sufficiently far from the mean of the draws. It can be interpreted as a penalty for distributions with a large number of outliers.

It is useful to now describe a method for computing the marginal likelihood when the numbers are of extreme size and dimensionality is large. Geweke's modified harmonic mean operates by approximating the actual density with a normal density in the neighborhood of the mode. Using the gaussian components of f we can standardize the mcmc draws $\theta^{(s)}$ which alleviates the problem of calculating the determinant of V_θ with high precision. A large number of parameters leads to large matrices V_θ that present a problem for computing the inverse, and associated determinant, with high precision. The transformation works by noting that the calculation can be carried out in the context of a standard multivariate random normal distribution. First define $z^{(s)} = C^{-1}\left(\theta^{(s)} - \overline{\theta}\right)$ as the standardized value of θ_i relative to the distribution of the mcmc draws (mean zero, unit variance). Here C^{-1} is the inverse of the cholesky factorization of the matrix V_θ. Substituting this into the formula for $f(\theta^{(s)})$ we get a simple expression

$$f(\theta^{(s)}) = \tau^{-1}(2\pi)^{-d/2}\exp\left\{-\frac{z^{(s)\prime}z^{(s)}}{2}\right\}$$

where $|V_\theta|^{-1/2}$ and V_θ^{-1} drop out since they become 1 and the identity matrix under the transformation. Another scale transformation is useful for calculating the marginal likelihood. First note the following equivalence where c is a constant

$$\begin{aligned}
\frac{f(\theta^{(s)})}{L(\theta^{(s)}|Y)p(\theta^{(s)})} &= \exp\left\{\ln\left(\frac{f(\theta^{(s)})}{L(\theta^{(s)}|Y)p(\theta^{(s)})} * c\right)\right\}/c \\
&= \exp\left\{\ln f(\theta^{(s)}) - \ln\left(L(\theta^{(s)}|Y)p(\theta^{(s)})\right) + \ln c\right\}/c
\end{aligned}$$

The resulting estimator, equal to Geweke's modified harmonic mean, can be written as:

$$p(Y) = c * \left[\frac{1}{n_{sim}} \sum_{s=1}^{n_{sim}} \exp \left\{ \ln f(\theta^{(s)}) - \ln \left(L(\theta^{(s)}|Y) p(\theta^{(s)}) \right) + \ln c \right\} \right]^{-1}$$

These numbers are quite large for the model estimated herein so I report the ratio of the log of the marginal data densities in the paper.

A.0.6. Additional Figures and Tables for Chapter 1

Estimation Results for Alternative Observation Vectors			
	$\log\left(\frac{P(Data\|Model_5)}{P(Data\|Model_1)}\right)$	News Contribution Uncond VD.	News Contribution 1-period VD
No yields or Stock Returns	54.40	34.83	11.12
Yields (Baseline)	530.04	49.71	34.16
Yields and Stock Returns	1142.82	59.98	50.19

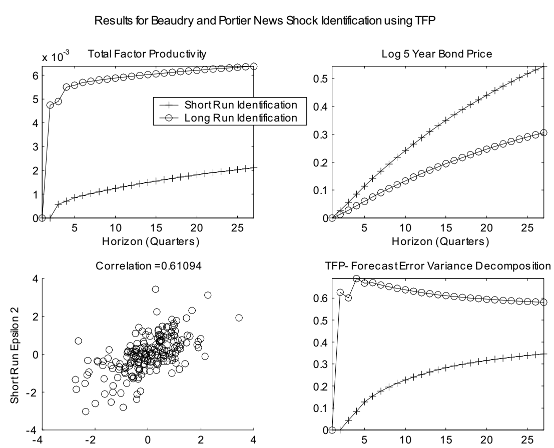
Preferred Model, Interest rates included in estimation		
Parameter Description	Notation	Post Mean
Firm Borrowing Constraint	ν	1
Price Stickiness	α	0.79
Wage Stickiness	$\overline{\alpha}$	0.36
Investment Adj. Cost	κ	2.03
Price Indexing	χ	0.16
Wage Indexing	$\overline{\chi}$	0.03
Capacity Utlilization Param.	γ_2	0.01
Inflation Feedback	α_π	1.20
Output Gap Feedback	α_y	0.06
Interest Rate Feedback	α_R	0.89
Autoreg. Tech Shock	ρ_z	0.70
Autoreg. Gov. Shock	ρ_g	0.96
Autoreg. Pref. Shock	ρ_β	0.99
Autoreg. Labor Pref. Shock	ρ_{ϕ_0}	0.81
Autoreg. Inv. Specific Shock	ρ_k	0.98
Autoreg. Mon. Pol. Shock	ρ_{mp}	-.0167
Sdev Tech Shock	σ_z	0.0039
Sdev. Gov. Shock	σ_g	0.0301
Sdev. Pref. Shock	σ_β	0.0311
Sdev. Labor Pref. Shock	σ_{ϕ_0}	0.0686
Sdev. Inv. Specific Shock	σ_k	0.0173
Sdev. Mon. Pol. Shock	σ_{mp}	0.0019

1 Period Variance Decomposition of Output Growth						
Observations include yields but no stock returns						
Model		1	2	3	4	5
log ratio Marg Lik		0	474.79	388.74	505.01	530.04
Neutral tech.	χ_t	0.73	0.10	0.16	0.24	0.36
Government	g_t	22.53	37.25	68.48	14.57	15.40
Demand	$\varphi_{b,t}$	5.74	2.47	1.50	6.89	6.29
Labor supply	$\varphi_{0,t}$	66.50	24.41	18.90	18.18	19.97
Investment	$\tau_{k,t}$	0.03	8.81	0.06	0.85	0.887
Monetary policy	mp_t	4.47	3.02	7.09	27.11	22.93
Tech. news	ξ_χ	-	7.80	-	-	0.36
Gov. news	ξ_{mp}	-	-	3.72	-	0
Inv. news	ξ_{τ_k}	-	-	-	32.16	33.80

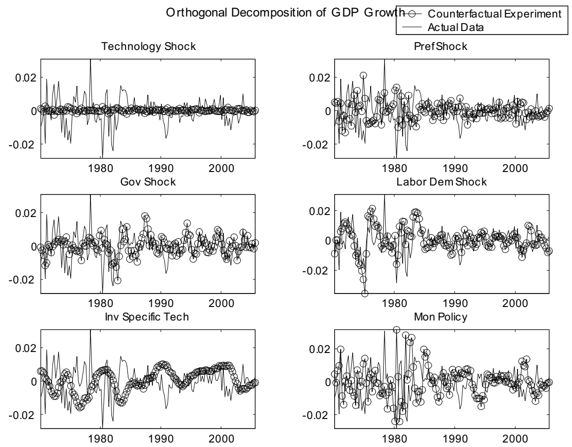
Unconditional Variance Decomposition of Output Growth						
Observations include stock returns and yields						
Model		1	2	3	4	5
log ratio Marg Lik		0	713.89	1070.73	886.18	1142.82
Neutral tech.	χ_t	0.25	56.66	0.59	1.27	1.45
Government	g_t	1.64	1.33	17.84	1.50	0.04
Demand	$\varphi_{b,t}$	2.14	6.99	18.02	4.68	0.09
Labor supply	$\varphi_{0,t}$	46.85	1.98	0.82	1.36	0.17
Investment	$\tau_{k,t}$	0.47	1.50	9.68	9.06	41.07
Monetary policy	mp_t	48.65	1.78	1.64	2.29	3.39
Tech. news	ξ_χ	-	29.77	-	-	6.41
Gov. news	ξ_{mp}	-	-	51.41	-	0
Inv. news	ξ_{τ_k}	-	-	-	79.84	52.57

1 Period Variance Decomposition of Output Growth						
Observations include stock returns and yields						
Model		1	2	3	4	5
log ratio Marg Lik		0	713.89	1070.73	886.18	1142.82
Neutral tech.	χ_t	0.76	58.95	0.30	3.56	11.27
Government	g_t	0.18	2.41	51.94	6.95	10.12
Demand	$\varphi_{b,t}$	0	9.54	29.55	7.96	5.41
Labor supply	$\varphi_{0,t}$	42.10	2.36	1.26	0.04	1.79
Investment	$\tau_{k,t}$	14.07	0.26	7.67	0.24	7.46
Monetary policy	mp_t	42.90	2.52	3.33	8.43	13.72
Tech. news	ξ_χ	-	23.96	-	-	3.49
Gov. news	ξ_{mp}	-	-	5.93	-	0.01
Inv. news	ξ_{τ_k}	-	-	-	72.81	46.70

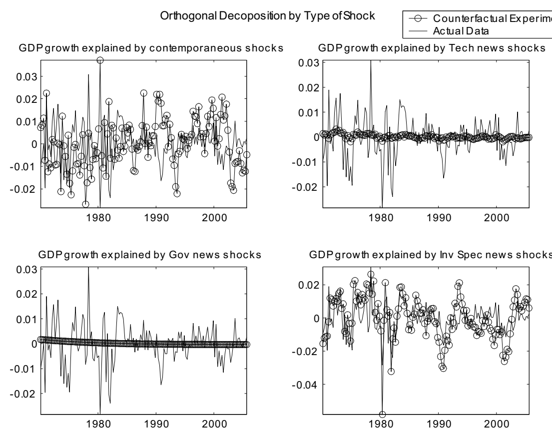
1 Period Variance Decomposition of Yield Curve Factors				
Model #5, Observation vector excludes stock returns				
Yield Curve Factor		Level	Slope	Curvature
Neutral tech.	χ_t	0.50	0.20	0.23
Government	g_t	0.02	0.14	0.01
Demand	$\varphi_{b,t}$	26.90	27.40	16.54
Labor supply	$\varphi_{0,t}$	14.59	0.12	1.24
Investment	$\tau_{k,t}$	0.34	0	0.22
Monetary policy	mp_t	42.51	69.8	81.52
Tech. news	ξ_χ	0.30	0.02	0.02
Gov. news	ξ_{mp}	0	0	0
Inv. news	ξ_{τ_k}	14.84	2.30	0.21



Beaudry and Portier news decomposition using total factor productivity



GDP growth due to various contemporaneous shocks



GDP growth due to news shocks

APPENDIX B

Chapter 2

B.0.7. Derivation of a Simple DSGE Model

Here I present a simple New-Keynesian model based on the work of Bekaert, Cho and Moreno (2007) featuring AS, IS and monetary policy equations. Two additional features, the assumption of an exogenous natural rate process and a time-varying inflation target, complete the set of equations that form a rational expectations equilibrium. Bekaert, Cho and Moreno fit this model to several macroeconomic variables as well as interest rates of various maturities under the expectations hypothesis. These authors argue that the introduction of interest rates to the estimation helps to identify parameters of the Phillips curve and that the model does relatively well at describing yield curve dynamics. The following sections briefly derive the equations describing the model equilibrium. For additional details and discussion refer to Bekaert, Cho and Moreno (2007).

B.0.7.1. The IS Equation. The IS equation is based on the optimizing behavior of a representative agent whose preferences exhibit habit formation. The agent seeks to maximize the following objective

$$E_t\left[\sum_{s=t}^{\infty}\psi^{s-t}\left(\frac{F_sC_s^{1-\sigma}-1}{1-\sigma}\right)\right]$$

where C_t is consumption, F_t an aggregate demand shifting factor, ψ a time discount factor and σ the inverse of the intertemporal elasticity of consumption. In this model

$F_t = H_t G_t$ where H_t is an external level of habit, based on levels of past consumption, and G_t is an exogenous aggregate demand shock that can also be interpreted as a preference shock. Following BCM I assume that $H_t = C_{t-1}^{\eta}$ where η measures the degree of habit dependence on the past level of consumption.

Imposing the resource constraint (that consumption equals output) and assuming log normality, the Euler equation for the interest rate yields the IS equation:

$$y_t = \alpha_{IS} + \mu E_t y_{t+1} + (1-\mu) y_{t-1} - \phi(r_t - E_t \pi_{t+1}) + \xi_t^{is}$$

where y_t is detrended log output, r_t is the short term interest rate, $\phi = \frac{1}{\sigma+\eta}$ and $\mu = \sigma\phi$. Note that $\mu = 1$ when $\eta = 0$ and in this case lagged output does not enter the equation. Thus, habit formation play a role in effecting the endogenous persistence of output. The IS shock, $\xi_t^{is} = \phi \ln G_t$ is assumed to follow an $AR(1)$ process given by

$$\xi_t^{is} = \alpha_{is} + \rho^{is} \xi_{t-1}^{is} + \varepsilon_t^{is} + \eta_t^{is}$$

B.0.7.2. The AS Equation (Phillips Curve). The Phillips curve in BCM's model builds on the sticky-pricing framework of Calvo (1983). Assuming monopolistic competition in the intermediate goods markets, a fraction of price setters do not reoptimize their prices but instead index to lagged inflation and that the real marginal cost is proportional to the output gap, the New-Keynesian aggregate supply curve relating inflation to the output gap is[1]

$$\pi_t = \delta E_t \pi_{t+1} + (1-\delta)\pi_{t-1} + \kappa(y_t - y_t^n) + \xi_t^{as}$$

[1]This is derived in detail in Woodford (2003)

where π_t is inflation, y_t^n is the natural rate of output that would arise in the case of flexible prices, $y_t - y_t^n$ is the output gap, ξ_t^{as} is an exogenous supply shock, κ captures the short run tradeoff between inflation and the output gap and $(1-\delta)$ captures endogenous persistence in inflation. Futhermore it is assumed that the natural rate of output follows an exogenous stochastic process (augmented by 'news' shocks)

$$y_t^n = \alpha_{y^n} + \lambda y_{t-1}^n + \varepsilon_t^{y^n} + \eta_t^{y^n}$$

In this framework innovations to the natural rate process can be thought of as shocks to the markup or market power of intermediate goods firms. The exogenous supply shock is assumed to follow an $AR(1)$ stochastic process

$$\xi_t^{as} = \alpha_{is} + \rho^{as}\xi_{t-1}^{as} + \varepsilon_t^{as} + \eta_t^{as}$$

B.0.7.3. The Monetary Policy Equation. I assume that the monetary authority specifies a nominal interest rate target, r_t^*, as follows[2]

$$r_t^* = r + \beta\left(E_t\pi_{t+1} - \pi_t^*\right) + \gamma(y_t - y_t^n)$$

where π_t^* is a time-varying inflation target derived below. The observed nominal rate is set by smoothing between the current and target interest rates:

$$r_t = \rho r_{t-1} + (1-\rho)r_t^* + \varepsilon_t^{mp}$$

[2] This forward-looking specification is based on that proposed by Clarida, Gali and Gertler (1999)

Here ε_t^{mp} is a monetary policy shock that captures unanticipated deviations from the Taylor rule. Expectations about long run inflation are not well understood. BCM posit a simple structure for long run inflation as the conditional expectation of a weighted average of future inflation rates

$$\pi_t^{LR} = (1-d)\sum_{j=0}^{\infty} d^j E_t \pi_{t+j} \tag{B.1}$$

where d is a number between 0 and 1. Note that $\pi_t = \pi_t^{LR}$ when $d = 0$ and as d approaches 1 long run inflation approaches constant steady-state inflation. Following BCM, I assume that the monetary authority anchors its inflation target around π_t^{LR}, but smooths target changes (similar to the specification for monetary policy):

$$\pi_t^* = \omega \pi_{t-1}^* + (1-\omega)\pi_t^{LR} + \varepsilon_t^{\pi^*} \tag{B.2}$$

Combining B.1 and B.2:

$$\pi_t^* = \varphi_1 E_t \pi_{t+1}^* + \varphi_2 \pi_{t-1}^* + \varphi_3 \pi_t + \varepsilon_t^{\pi^*}$$

where $\phi_1 = \frac{d}{1+d\omega}$, $\varphi_2 = \frac{w}{1+dw}$ and $\varphi_3 = 1 - \varphi_1 - \varphi_2$

B.0.8. Mapping the State Vector to the Observed Factors

Recall that linearly approximating the small DSGE presented in the previous section results in a solution of the form

$$\begin{aligned} S_t &= F(\theta) S_{t-1} + Q(\theta) \varepsilon_t \\ \widehat{F}_t &= H(\theta) S_t \end{aligned}$$

where S_t is the 'state' vector associate with the DSGE model and $\widehat{F}_t$ is a vector of factors measured in deviations from their mean[3]. It is useful to derive everything in terms of deviations from the mean, hence the 'hatted' variables. This allows for a reduction in the number of variables to be estimated since the steady state parameters of the DSGE model can be fixed to coincide with the historical means.

Now I will derive the components of $H(\theta)$ that map the state into the model's implications for the factors. The mappings for year-on-year inflation and year-on-year output growth are simply obtained by augmenting the state equation, obtained from the solution, to include lags of inflation and output. The excess return forecasting factor of Cochrane and Piazzesi (2006) and level and slope factors are slightly more sophisticated and thus are derived below.

To begin define the mapping from the state to the short rate as H_r, given as part of the linear solution, as

$$\widehat{r}_t = H_r S_t$$

[3]For details on how this solution is obtained see Klein (2000)

then bond prices can be found by the relation (note the expectations hypothesis underlying this equivalence)

$$\text{(B.3)} \qquad \begin{aligned} \widehat{p}_t^{(n)} &= -\sum_{j=0}^{n-1} E_t\left[\widehat{r}_{t+j}\right] = -H_r\left(\sum_{j=0}^{n-1} F(\theta)^j\right) S_t \\ &= H_{p^n} * S_t \end{aligned}$$

and thus forward rates are simply defined as the difference between bond prices of different maturities: $f_t^{(n)} = p_t^{(n-1)} - p_t^{(n)}$. Note that the estimation uses annual forward rates though the model is estimated at a quarterly frequency.

B.0.8.1. Cochrane and Piazzesi (CP) Risk Factor Mapping. Now I will describe the excess return forecasting factor of Cochrane and Piazzesi as well as the method for obtaining a counterpart for this variable in the DSGE model. Note that the linear approximation to the solution in the DSGE model exhibits the certainty equivelence property and thus risk premia are not time-varying. This is of course a result of the solution method and not of the model. To obtain the DSGE model's counterpart to the excess return forecasting factor I utilize the fact that the factor is a particular linear combination of forward rates. Thus, I find the linear combination following the strategy of Cochrane and Piazzesi (2006) which, in addition to the forward rate mappings, can be used to construct a proxy for the excess return forecasting factor.

CP obtain the excess return forecasting factor first consider the following regression (spreads specification)

$$rx_{t+1}^{(n)} = \alpha^{(n)} + \beta^{(n)}\widetilde{f}_t + \varepsilon_{t+1}$$

where

$$\widetilde{f}_t^{(n)} = f_t^{(n)} - y_t^{(1)} = p_t^{(n-1)} - p_t^{(n)} + p_t^{(1)}$$
$$\widetilde{f}_t = \begin{bmatrix} \widetilde{f}_t^{(1)} & \widetilde{f}_t^{(2)} & ... \end{bmatrix}$$

then the excess return forecasting factor corresponds to the largest eigenvalue of the resulting variance-covariance matrix of expected returns

$$Q\Lambda Q = cov\left(E_t\left[rx_{t+1}\right]\right) = cov\left(\alpha + \beta\widetilde{f}_t\right)$$

Letting q_r denote the eigenvector corresponding to the largest eigenvalue, define the excess return forecasting factor as

$$\widetilde{x}_t = q_r' * \left(\alpha + \beta\widetilde{f}_t\right)$$

which can be written conveniently as

$$\widetilde{x}_t = \gamma_{x,0} + \gamma_{x,1}' * \widetilde{f}_t$$

where

$$\gamma_{x,0} = q_r'\alpha$$
$$\gamma_{x,1}' = q_r'\beta$$

Note that the excess return forecasting factor is simply a linear function of difference spread forward rates (a vector of forward rate less the 1 year yield). Once $\gamma_{x,1}$ is obtained from data on bond prices it is used in conjuction with B.3 to obtain the DSGE implications for the risk factor. In detail, note that B.3 can be used to construct a matrix $H_{\widetilde{f}}$ such that

$$\widehat{\widetilde{f}}_t = H_{\widetilde{f}} * S_t$$

and thus

$$\widehat{\widetilde{x}}_t = \gamma'_{x,1} * H_{\widetilde{f}} * S_t$$

where hatted variables represents deviations from the mean.

B.0.8.2. Level and Slope Mappings. Level is defined as the average of the 2, 5 and 10 year yields. This can be found using B.3 and

$$\widehat{level}_t = -\frac{p_t^{(2)}}{2} - \frac{p_t^{(5)}}{5} - \frac{p_t^{(10)}}{10}$$

The slope is defined as the difference between the 2 and 10 year yields

$$\widehat{slope}_t = -\frac{p_t^{(10)}}{10} + \frac{p_t^{(2)}}{2}$$

B.0.9. Prior Distributions

The prior distribution for the variance-covariance matrix of the VAR process, Σ (dimension $p \times p$), is assumed to be distributed Inverse Wishart. Evaluating the prior for the candidate covariance matrix Σ centered at $\lambda T \Sigma^* (\theta)$ and having degrees of freedom $\lambda T - k$ can be done using the followng formula

$$P(\Sigma|\theta, \lambda) = \frac{|\lambda T \Sigma^* (\theta)|^{\lambda T - k} |\Sigma|^{-((\lambda T - k)+p+1)/2} \exp\left(-trace\left(\Sigma * (\lambda T \Sigma^* (\theta))^{-1}\right)/2\right)}{2^{(\lambda T - k)p/2} \Gamma_p \left((\lambda T - k)/2\right)}$$

where Γ_p is the multivariate gamma function:

$$\Gamma_p(a) = \pi^{p(p-1)/4} \prod\nolimits_{j=1}^{p} \Gamma\left(a + (1-j)/2\right)$$

The prior distribution for the VAR coefficients Φ is Gaussian with mean corresponding to the parameters from the VAR approximating the DSGE(θ) model and variance $\frac{1}{\lambda T}\left[\Sigma^{-1} \otimes \Gamma_{XX}(\theta)\right]^{-1}$. Recall that the VAR describing the Factor dynamics (written in terms of deviation from the mean) is

$$\widehat{F}_t = \widehat{X}'_{t-1}\Phi + v_t$$

Define the stacked matrix, having dimension N, as

$$\widetilde{\Phi} = vec\left(\Phi\right)$$

where vec is an operator that maps a matrix to a column vector by stacking the columns of the matrix. Then the prior is given by

$$\begin{aligned} P(\Phi|\Phi^*(\theta), \Sigma) &= \frac{\left|\frac{1}{\lambda T}\left[\Sigma^{-1} \otimes \Gamma_{XX}(\theta)\right]^{-1}\right|^{-1/2}}{(2\pi)^{N/2}} \times \cdots \\ &\cdots \times \exp\left(-\frac{1}{2}\left(\Phi - \Phi^*(\theta)\right)' \left(\frac{1}{\lambda T}\left[\Sigma^{-1} \otimes \Gamma_{XX}(\theta)\right]^{-1}\right)^{-1} \left(\Phi - \Phi^*(\theta)\right)\right) \end{aligned}$$

Further details can be found in Hamilton (1994) Chapter 11 where the distribution of the VAR parameters is derived. Now note the following useful formula for evaluating the matrix determinant above (assume Σ is $p \times p$ and $\Gamma_{XX}(\theta)$ is $q \times q$)

$$\begin{aligned} \left| \frac{1}{\lambda T} \left[\Sigma^{-1} \otimes \Gamma_{XX}(\theta) \right]^{-1} \right| &= \left(\frac{1}{\lambda T} \right)^{p*q} \left| \left[\Sigma^{-1} \otimes \Gamma_{XX}(\theta) \right]^{-1} \right| \\ &= \left(\frac{1}{\lambda T} \right)^{p*q} \frac{1}{\left| \Sigma^{-1} \otimes \Gamma_{XX}(\theta) \right|} \\ &= \left(\frac{1}{\lambda T} \right)^{p*q} \frac{1}{\left| \Sigma^{-1} \right|^{q} * \left| \Gamma_{XX}(\theta) \right|^{p}} \\ &= \left(\frac{1}{\lambda T} \right)^{p*q} \frac{\left| \Sigma \right|^{q}}{\left| \Gamma_{XX}(\theta) \right|^{p}} \end{aligned}$$

as well as the matrix $\left(\frac{1}{\lambda T} \left[\Sigma^{-1} \otimes \Gamma_{XX}(\theta) \right]^{-1} \right)^{-1}$

$$\left(\frac{1}{\lambda T} \left[\Sigma^{-1} \otimes \Gamma_{XX}(\theta) \right]^{-1} \right)^{-1} = \lambda T * \Sigma^{-1} \otimes \Gamma_{XX}(\theta)$$

B.0.10. Marginal Likelihood Approximation

The marginal likelihood is calculated using Geweke's modified harmonic mean. The modified harmonic mean estimator uses draws from the posterior distribution, obtained via markov-chain monte carlo starting from the mode. The modified harmonic mean estimator can be written as:

$$P(data|Model_i) \approx \left[\frac{1}{n_{sim}} \sum_{s=1}^{n_{sim}} \frac{f(\theta^{(s)})}{L(\theta^{(s)}|Y)p(\theta^{(s)})} \right]^{-1}$$

where $L(\theta^{(s)}|Y)p(\theta^{(s)})$ represents the value of the posterior at the point $\theta^{(s)}$, n_{sim} is the number of draws from the posterior and $f(\theta^{(s)})$ is given by:

$$\begin{aligned} f(\theta) &= \tau^{-1}(2\pi)^{-d/2}\left|V_\theta\right|^{-1/2}\exp\left\{-\frac{\left(\theta-\overline{\theta}\right)' V_\theta^{-1}\left(\theta-\overline{\theta}\right)}{2}\right\}... \\ & \quad ... \times\left\{\left(\theta-\overline{\theta}\right)' V_\theta^{-1}\left(\theta-\overline{\theta}\right) \leq F_{\chi_d^2}^{-1}(\tau)\right\} \end{aligned}$$

Note that τ is a hyperparamater that determines a cutoff for points sufficiently far from the mean of the draws. It can be interpreted as a penalty for distributions with a large number of outliers. A value of $\tau = 0.25$ is used in the calculation of the marginal likelihood

It is useful to now describe a method for computing the marginal likelihood when the numbers are of extreme size and dimensionality is large. Geweke's modified harmonic mean operates by approximating the actual density with a normal density in the neighborhood of the mode. Using the gaussian components of f we can standardize the mcmc draws $\theta^{(s)}$ which alleviates the problem of calculating the determinant of V_θ with high precision. A large number of parameters leads to large matrices V_θ that present a problem for computing the inverse, and associated determinant, with high precision. The transformation works by noting that the calculation can be carried out in the context of a standard multivariate random normal distribution. First define $z^{(s)} = C^{-1}\left(\theta^{(s)} - \overline{\theta}\right)$ as the standardized value of θ_i relative to the distribution of the mcmc draws (mean zero, unit variance). Here C^{-1} is the inverse of the cholesky factorization of the matrix V_θ.

Substituting this into the formula for $f(\theta^{(s)})$ we get a simple expression

$$f(\theta^{(s)}) = \tau^{-1}(2\pi)^{-d/2}\exp\left\{-\frac{z^{(s)\prime}z^{(s)}}{2}\right\}$$

where $|V_\theta|^{-1/2}$ and V_θ^{-1} drop out since they become 1 and the identity matrix under the transformation. Another scale transformation is useful for calculating the marginal likelihood. First note the following equivalence where c is a constant

$$\begin{aligned}\frac{f(\theta^{(s)})}{L(\theta^{(s)}|Y)p(\theta^{(s)})} &= \exp\left\{\ln\left(\frac{f(\theta^{(s)})}{L(\theta^{(s)}|Y)p(\theta^{(s)})} * c\right)\right\}/c \\ &= \exp\left\{\ln f(\theta^{(s)}) - \ln\left(L(\theta^{(s)}|Y)p(\theta^{(s)})\right) + \ln c\right\}/c\end{aligned}$$

The resulting estimator, equal to Geweke's modified harmonic mean, can be written as:

$$p(Y) = c * \left[\frac{1}{n_{sim}}\sum_{s=1}^{n_{sim}}\exp\left\{\ln f(\theta^{(s)}) - \ln\left(L(\theta^{(s)}|Y)p(\theta^{(s)})\right) + \ln c\right\}\right]^{-1}$$

These numbers are quite large for the model estimated herein so I report the ratio of the log of the marginal data densities in the paper.

B.0.11. Additional Figures and Tables for Chapter 2

The results presented here correspond to the 'preferred' model: DSGE weight = 75% and news shocks arriving via the stochastic process for y^n

Parameter	Description	Median MCMC	Prior Distribution
δ	Phillips coeff: $E_t\pi_{t+1}$	0.63	Beta(0.6,0.1)
κ	Phillips coeff: $(y_t - y_t^n)$	0.00	Unif(0,0.3)
σ	Risk Aversion	5.44	Normal(2,1)[4]
η	Habit Formation	4.23	Normal(3,1)
ρ	I.R. Persistence	0.88	Beta(0.8,0.1)[5]
β	Taylor Rule Infl	1.39	Normal(1.5,0.2)[6]
γ	Taylor Rule Out. Gap	0.14	Beta(0.08,0.03)
λ	Persistence Nat. Rate	0.55	Beta(0.9,0.1)
w	Infl Target Param	0.29	Beta(0.85,0.1)
d	Long Run Infl Exp Param	0.99	Beta(0.85,0.1)
ρ_{as}	AS shock Persistence	0.01	Unif(0,1)
ρ_{is}	IS shock Persistence	0.38	Unif(0,1)

Variance Decomposition of Yield Curve Factors

Shock \ Yield Factor:	*CP*	*Level*	*Slope*
$\varepsilon_t^{y^n}$ (Nat. Rate)	14.16	6.35	7.63
ε_t^{as} (AS)	8.25	76.86	10.30
ε_t^{is} (IS)	4.14	1.04	2.00
ε_t^{mp}(Mon. Pol.)	0.60	0.30	1.29
$\varepsilon_t^{\pi^*}$(Infl. Target)	46.81	9.30	61.25
$\eta_t^{y^n}$(News)	26.03	6.12	17.51

[4]This Normal distribution is truncated to only have mass above 1
[5]Truncated to have mass only on positive numbers
[6]This prior is truncated to have mass above 1

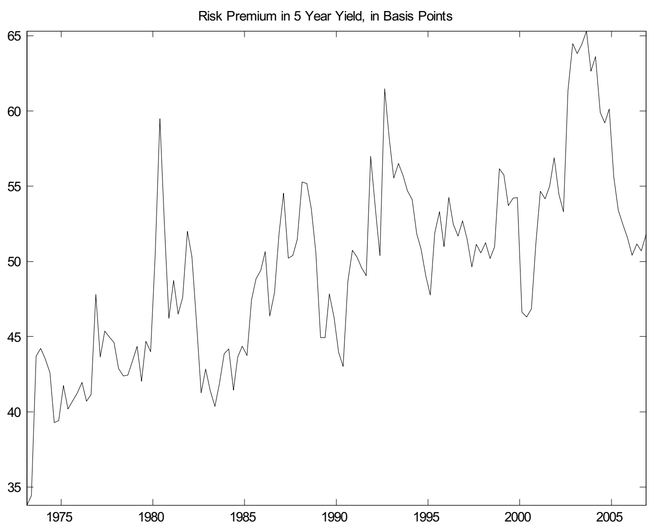


Figure B.1. Model implied risk premium

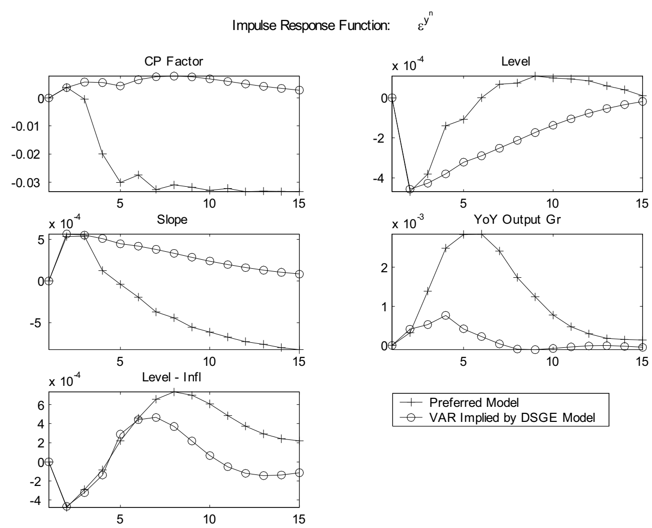


Figure B.2. Impulse response to a 1 standard deviation shock in natural rate of output process

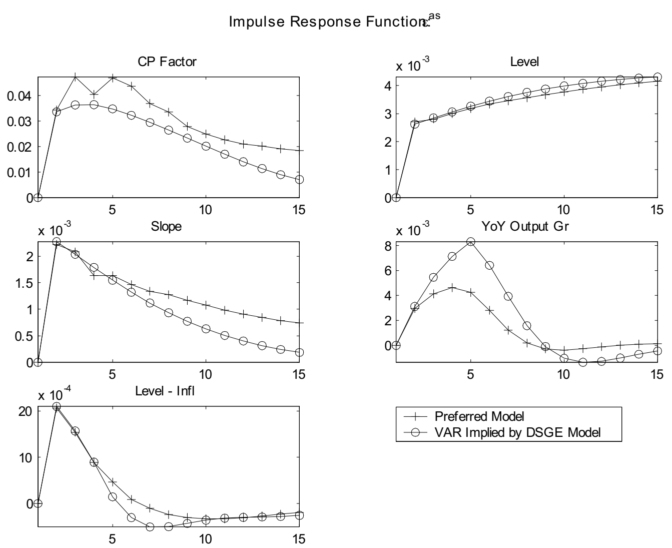


Figure B.3. Impulse response to a one standard deviation supply shock

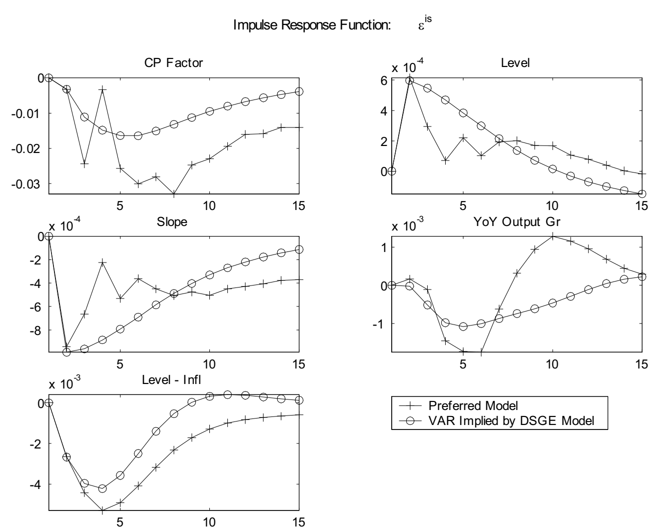


Figure B.4. Impulse response to a one standard deviation aggregate demand shock

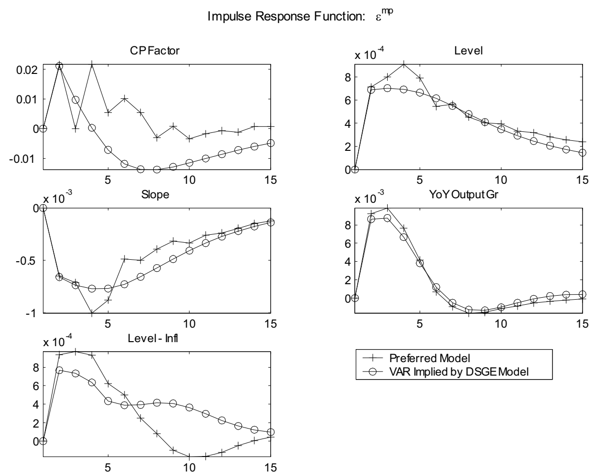


Figure B.5. Impulse response to a one standard deviation monetary policy shock

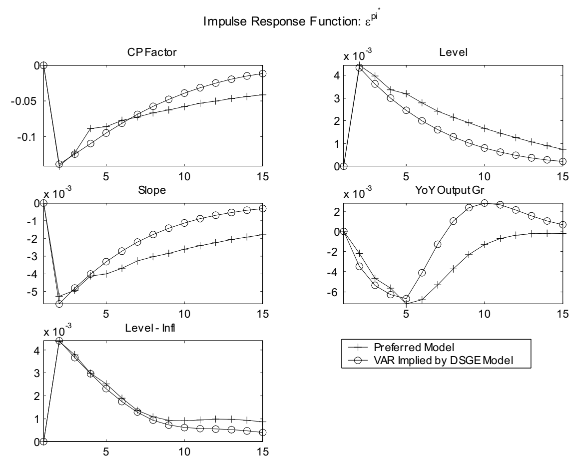


Figure B.6. Impulse response to a one standard deviation inflation target shock

APPENDIX C

Chapter 3

C.1. The Carlstrom-Fuerst Financial Friction Wedge

This section considers a version of the CF model, modified to include the adjustment costs in capital studied in CKM. We identify the version of the RBC model with wedges whose equilibrium coincides with that of the CF model with adjustment costs. We state the result as a proposition. For the RBC wedge economy to have literally the same equilibrium as the CF economy with adjustment costs requires several wedges and other adjustments. We then describe the parameter settings required in the original CF model to ensure that the adjustments primarily take the form of a wedge in the intertemporal Euler equation, and nowhere else. In this respect we follow the approach taken in CKM. To simplify the notation, we set the population growth rate to zero throughout the discussion in the appendix.

C.1.1. RBC Model With Adjustment Costs

To establish a baseline, we describe the version of the RBC model with adjustment costs. Preferences are:

$$E_0 \sum_{t=0}^{\infty} \beta^t u\left(c_t, l_t\right).$$

The resource constraint and the capital accumulation technology are, respectively,

$$c_t + x_t \leq k_t^{\alpha} \left(Z_t l_t\right)^{1-\alpha} \tag{C.1}$$

and

$$k_{t+1} = \left(1-\delta\right) k_t + x_t - \Phi\left(\frac{x_t}{k_t}\right) k_t. \tag{C.2}$$

The first order necessary conditions for optimization are:

$$\frac{-u_{l,t}}{u_{c,t}} = (1-\alpha)\left(\frac{k_t}{l_t}\right)^{\alpha} Z_t^{1-\alpha} \tag{C.3}$$

$$1 = \beta E_t \frac{u_{c,t+1}}{u_{c,t}}\left(1+R_{t+1}^k\right), \tag{C.4}$$

where the gross rate of return on capital is:

$$1 + R_{t+1}^k = \frac{\alpha\left(\frac{k_{t+1}}{Z_{t+1} l_{t+1}}\right)^{\alpha-1} + P_{k,t+1}}{P_{k't}}.$$

where $P_{k,t}$ and $P_{k',t}$ are given in (**??**) and (**??**).

In the following two subsections, we argue that the CF financial frictions act like a tax on the gross return on capital, $1+R_{t+1}^k$, in (C.4). In particular, $1+R_{t+1}^k$ is replaced by

$$\left(1+R_{t+1}^k\right)\left(1-\tau_t^k\right).$$

This statement is actually only true as an approximation. Below we state, as a proposition, what the exact RBC model with wedges is, which corresponds to the CF model. We then explain the sense in which the wedge equilibrium just described is an approximation.

C.1.2. The CF Model With Adjustment Costs

Here, we develop the version of the CF model in which there are adjustment costs in the production of new capital. The economy is composed of firms, an η mass of entrepreneurs and a mass, $1 - \eta$, of households. The sequence of events through the period proceeds as follows. First, the period t shocks are observed. Then, households and entrepreneurs supply labor and capital to competitive factor markets. Because firm production functions are homogeneous, all output is distributed in the form of factor income. Households and entrepreneurs then sell their used capital on a capital market. The total net worth of households and entrepreneurs at this point consists of their earnings of factor incomes, plus the proceeds of the sale of capital. Households divide this net worth into a part allocated to current consumption, and a part that is deposited in the bank. Entrepreneurs apply their entire net worth to a technology for producing new capital. They produce an amount of capital that requires more resources than they can afford with only their own net worth. They borrow the rest from banks. At this point the entrepreneur experiences an idiosyncratic shock which is observed to him, while the bank can only see it by paying a monitoring cost. This creates a conflict between the entrepreneur and the bank which is mitigated by the bank extending the entrepreneur a standard debt contract. After capital production occurs, entrepreneurs sell the new capital, and pay off their bank loan. Households receive a return on their deposits at the bank, and use the proceeds to purchase new capital. Entrepreneurs use their income after paying off the banks to buy consumption goods and new capital. All the newly produced capital is purchased by households and entrepreneurs, and all the economy's consumption goods are consumed. The next period, everything starts all over.

We now provide a formal description of the economy. The household problem is

$$\max_{\{c_{c,t},k_{ct+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t u\left(c_t, l_t\right),$$

subject to:

$$c_t + q_t k_{c,t+1} \leq w_t^c l_t + \left[r_t + P_{k,t}\right] k_{c,t} \tag{C.5}$$

where c_t and $k_{c,t}$ denote household consumption and the household stock of capital, respectively. In addition, and l_t denotes household employment, w_t^c denotes the household's competitive wage rate, $P_{k,t}$ denotes the price of used capital and q_t denotes the price of capital available for production in the next period (the reason for not denoting this price by $P_{k',t}$ will be clear momentarily). After receiving their period t income, households allocate their net worth (the right side of (C.5)) to c_t and the rest, $w_t^c l_t + \left[r_t + P_{k,t}\right] k_{c,t} - c_t$, is deposited in a bank. These deposits earn a rate of return of zero. This is because markets are competitive and the opportunity cost to the household of the output they lend to the bank is zero. Later in the period, when the deposit matures, the households use the principal to purchase $k_{c,t+1}$ units of capital. The first order conditions of the household are (C.5) with the equality strict and:

$$1 = E_t \beta \frac{u_{c,t+1}}{u_{c,t}} \left[\frac{r_{t+1} + P_{k,t+1}}{q_t}\right] \tag{C.6}$$

$$\frac{-u_{l,t}}{u_{c,t}} = w_t^e, \tag{C.7}$$

where $u_{c,t}$ and $-u_{l,t}$ denote the time t marginal utilities of consumption and leisure, respectively.

The η entrepreneurs' present discounted value of utility is:

$$E_0 \sum_{t=0}^{\infty} (\beta\gamma)^t \, c_{et}.$$

After the period t shocks are realized, the net worth of entrepreneurs, a_t, is:

$$a_t = w_t^e + [r_t + P_{k,t}] \, k_{et},$$

where w_t^e is the wage rate earned by the entrepreneur, who inelastically supplies his one unit of labor. The entrepreneur uses the a_t consumption goods, together with a loan from the bank to purchase the inputs into the production of capital goods. Entrepreneurs have access to the technology for producing capital, (**??**) and (**??**). The technology proceeds in two stages. In the first stage, the entrepreneur produces an intermediate input, i_t. In the second stage, that input results in ωi_t units of capital, which has a price, in consumption goods, q_t. The random variable, ω, is independently distributed across entrepreneurs, has mean unity, and cumulative density function,

$$\Psi(z) \equiv prob\,[\omega \leq z]\,.$$

The entrepreneur who wishes to produce i_t units of the capital input faces the following cost function:

$$C(i_t; P_{k,t}) = \min_{\varphi_{k,t}, \varphi_{x,t}} P_{k,t}\varphi_{k,t} + \varphi_{x,t} + \lambda_t \left[i_t - (1-\delta)\,\varphi_{k,t} - \varphi_{x,t} + \Phi\left(\frac{\varphi_{x,t}}{\varphi_{k,t}}\right)\varphi_{k,t} \right],$$

where the constraint is that the object in square brackets is no less than zero. In addition, $\varphi_{k,t}$ and $\varphi_{x,t}$ denote the quantity of old capital and investment goods, respectively, purchased by the entrepreneur. The first order conditions for $\varphi_{k,t}$ and $\varphi_{x,t}$ are:

$$P_{k,t} = \lambda_t \left[1 - \delta - \Phi\left(\frac{\varphi_{x,t}}{\varphi_{k,t}}\right) + \Phi'\left(\frac{\varphi_{x,t}}{\varphi_{k,t}}\right)\frac{\varphi_{x,t}}{\varphi_{k,t}}\right] \tag{C.8}$$

$$1 = \lambda_t \left[1 - \Phi'\left(\frac{\varphi_{x,t}}{\varphi_{k,t}}\right)\right], \tag{C.9}$$

respectively. The reason for denoting the time t price of old capital by $P_{k,t}$ is now apparent. Substituting out for λ in (C.8) from (C.9), we see that the formula for $P_{k,t}$ here coincides with with the one implied by (**??**)-(**??**). The reason for not denoting the price of new capital by $P_{k',t}$ is also apparent. Comparing (C.9) with (**??**)-(**??**) we see that the formula for λ_t coincides with the formula for $P_{k',t}$ implied by (**??**)-(**??**). However, the equilibrium value of q_t will not coincide with λ_t here. This is because λ_t does not capture all the costs of producing new capital. It measures the marginal costs implied by the production technology. However, as discussed in detail in CF, it is missing the marginal cost that arises from the conflict between entrepreneurs and banks. This has the consequence that the production of capital necessarily involves some monitoring, and therefore also involves some destruction of capital.

Solving (C.9) for x_t/k_t in terms of λ_t, and using the result to substitute out for $\varphi_{x,t}/\varphi_{k,t}$ in (C.8):

$$P_{k,t} = \lambda_t \left[(1-\delta) - \frac{a}{2}\left(\frac{\frac{1}{\lambda_t} - 1}{a}\right)^2 + a\left(\frac{\frac{1}{\lambda_t} - 1}{a}\right)\left(\frac{\frac{1}{\lambda_t} - 1}{a} + b\right)\right]$$

Solving this for λ_t, provides the marginal cost function for producing i_t :

$$\lambda_t = \lambda\left(P_{k,t}\right). \tag{C.10}$$

Because all entrepreneurs face the same $P_{k,t}$, they will all choose the same ratio, $\varphi_{x,t}/\varphi_{k,t}$, regardless of the scale of production, i_t. Moreover, that ratio must be equal to the ratio of aggregate investment to the aggregate stock of capital.

The constant returns to scale feature of the production function implies that the total cost of producing i_t is:

$$C\left(i_t; P_{k,t}\right) = \begin{cases} \lambda\left(P_{k,t}\right) i_t & a > 0 \\ i_t & a = 0 \end{cases}$$

Consider an entrepreneur who has a_t units of goods and wishes to produce $i_t \geq a_t$, so that the entrepreneur must borrow $\lambda\left(P_{k,t}\right) i_t - a_t$ from the bank. Following CF, we suppose that the entrepreneur receives a standard debt contract. This specifies a loan amount and an interest rate, R_t^a, in consumption units. If the revenues of the entrepreneur turn out to be too low for him to repay the loan, then the entrepreneur is 'bankrupt' and he simply provides everything he has to the bank. In this case, the bank pays a monitoring cost which is proportional to the scale of the entprepeneur's project, μi_t^a. We now work out the equilibrium value of the parameters of the standard debt contract.

The value of ω such that entrepreneurs with smaller values of ω are bankrupt, is $\bar{\omega}_t^a$, where

$$\left[\lambda\left(P_{k,t}\right) i_t - a_t\right] R_t^a = \bar{\omega}_t^a i_t q_t.$$

Using this we find that the average, across all entrepreneurs with asset level a_t, of revenues is:

$$\begin{aligned} & i_t q_t \int_0^\infty \omega dF\left(\omega\right) - \int_{\bar{\omega}_t^a}^\infty R_t^a \left(\lambda\left(P_{k,t}\right) i_t^a - a_t\right) dF\left(\omega\right) - i_t q_t \int_0^{\bar{\omega}_t^a} \omega dF\left(\omega\right) \\ = \; & i_t q_t f\left(\bar{\omega}_t^a\right), \end{aligned}$$

where

$$f\left(\bar{\omega}_t^a\right)=\int_{\bar{\omega}_t^a}^{\infty}\omega d\Phi\left(\omega\right)-\bar{\omega}_t^a\left(1-\Phi\left(\bar{\omega}_t^a\right)\right).$$

The average receipts to banks, net of monitoring costs, across loans to all entprepreneurs with assets a_t is:

$$\begin{aligned} & q_t i_t^a\left[\int_0^{\bar{\omega}_t^a}\omega d\Phi\left(\omega\right)-\mu\Phi\left(\bar{\omega}_t^a\right)\right]+\left[\lambda\left(P_{k,t}\right)i_t-a_t\right]R_t^a\left[1-\Phi\left(\bar{\omega}_t^a\right)\right] \\ = \; & q_t i_t^a g\left(\bar{\omega}_t^a\right), \end{aligned}$$

where

$$g\left(\bar{\omega}_t^a\right)=\int_0^{\bar{\omega}_t^a}\omega d\Phi\left(\omega\right)-\mu\Phi\left(\bar{\omega}_t^a\right)+\bar{\omega}_t^a\left[1-\Phi\left(\bar{\omega}_t^a\right)\right].$$

The contract with entrepreneurs with asset levels, a_t, that is assumed to occur in equilibrium is the one that maximizes the expected state of the entrepreneur at the end of the contract, subject to the requirement that the bank be able to pay the household a gross rate of interest of unity. The interval during which the entrepreneur produces capital is one in which there is no alternative use for the output good. So, the condition that must be satisfied for the bank to participate in the loan contract is:

$$q_t i_t g\left(\bar{\omega}_t^a\right)\geq\lambda\left(P_{k,t}\right)i_t-a_t.$$

The contract solves the following Lagrangian problem:

$$\max_{\bar{\omega}_t^a,i_t} i_t q_t f\left(\bar{\omega}_t^a\right)+\mu\left[q_t i_t g\left(\bar{\omega}_t^a\right)-\lambda\left(P_{k,t}\right)i_t+a_t\right].$$

The first order conditions for $\bar{\omega}_t^a$ and i_t are, after solving out for μ and rearranging:

$$q_t f\left(\bar{\omega}_t^a\right) = \frac{f'\left(\bar{\omega}_t^a\right)}{g'\left(\bar{\omega}_t^a\right)}\left[q_t g\left(\bar{\omega}_t^a\right) - \lambda\left(P_{k,t}\right)\right] \tag{C.11}$$

$$i_t = \frac{a_t}{\lambda\left(P_{k,t}\right) - q_t g\left(\bar{\omega}_t^a\right)} \tag{C.12}$$

From (C.11), we see that $\bar{\omega}_t^a = \bar{\omega}_t$ for all a_t, so that $R_t^a = R_t$ for all a_t. It then follows from (C.12) that the loan amount is proportional to a_t. As in the no adjustment cost case in CF, these two properties imply that in studying aggregates, we can ignore the distribution of assets across entrepreneurs.

The expected net revenues of the entrepreneurs, expressed in terms of a_t, are, after making use of (C.12):

$$i_t q_t f\left(\bar{\omega}_t\right) = \frac{q_t f\left(\bar{\omega}_t\right)}{\lambda\left(P_{k,t}\right) - q_t g\left(\bar{\omega}_t\right)} a_t. \tag{C.13}$$

At the end of the period, after the debt contract with the bank is paid off, the entrepreneurs who do not go bankrupt in the process of producing capital have income that can be used to buy consumption goods and new capital goods:

$$c_{et} + q_t k_{et+1} \leq \begin{cases} i_t q_t \omega - R_t\left(\lambda\left(P_{k,t}\right) i_t - a_t\right) & \omega \geq \bar{\omega}_t \\ 0 & \omega < \bar{\omega}_t \end{cases}. \tag{C.14}$$

An entrepreneur who is bankrupted in period t must set $c_{et} = 0$ and $k_{et+1} = 0$. In period $t+1$, these entrepreneurs start with net worth $a_{t+1} = w_{t+1}^e$. Entrepreneurs who are not bankrupted in period t can purchase positive amounts of c_{et} and $k_{e,t+1}$ (except in the non-generic case, $\omega = \bar{\omega}_t$). For these entrepreneurs, the marginal cost of purchasing $k_{e,t+1}$ is q_t units of consumption. The time t expected marginal payoff from $k_{e,t+1}$ at the beginning of period $t+1$ is $E_t\left[r_{t+1}+P_{k,t+1}\right]$. In each aggregate state in period $t+1$, the entrepreneur expands his net worth by the value of $\left[r_{t+1}+P_{k,t+1}\right]$ in that state. This extra net worth can be leveraged into additional bank loans, which in turn permit an expansion in the entrepreneur's payoff by investing in the capital production technology. The expected value of this additional payoff (relative to date $t+1$ idiosyncratic uncertainty) corresponds to the coefficient on a_t in (C.13). So, the expected rate of return available to entrepreneurs who are not bankrupt in period t is:

$$E_t\left[\frac{r_{t+1}+P_{k,t+1}}{q_t}\times\zeta_{t+1}\right], \tag{C.15}$$

which they equate to $1/(\beta\gamma)$. Here,

$$\zeta_{t+1}=\max\left[\frac{q_{t+1}f\left(\bar{\omega}_{t+1}\right)}{\lambda\left(P_{k,t+1}\right)-q_{t+1}g\left(\bar{\omega}_{t+1}\right)},1\right].$$

The expression to the left of '$\times$' in (C.15) is the rate of return enjoyed by ordinary households. The reason that ζ_{t+1} cannot be less than unity is that and entprepeneur can always obtain unity, simply by consuming his net worth in the following period and not producing any capital. Averaging over all budget constraints in (C.14):

$$c_{et}+q_tk_{et+1}=\frac{q_tf\left(\bar{\omega}_t\right)}{\lambda\left(P_{k,t}\right)-q_tg\left(\bar{\omega}_t\right)}a_t.$$

Here, c_{et} and $k_{e,t+1}$ refer to averages across all entrepreneurs.

Output is produced by goods-producers using a linear homogeneous technology,

$$y\left(k_t,l_t,\eta,Z_t\right)=k_t^{\alpha}\left(\left(1-\eta\right)Z_tl_t\right)^{1-\alpha-\zeta}\eta^{\zeta}, \tag{C.16}$$

where k_t is the sum of the capital owned by households and the average capital held by entrepreneurs:

$$k_t = (1-\eta)\, k_{ct} + \eta k_{e,t}.$$

The argument, η, in y is understood to apply to the second occurrence of η. The arguments in the production function reflect our assumption that the entrepreneur supplies one unit of labor, and households supply l_t units of labor. Profit maximization implies:

$$y_{k,t} = r_t,\ y_{l,t} = w_t^c,\ y_{3,t} = w_t^e. \tag{C.17}$$

We now collect the equilibrium conditions for the economy. The production of new capital goods by the average entrepreneur is:

$$\begin{aligned} & i_t \int_0^{\infty} \omega dF(\omega) - \mu i_t \int_0^{\bar{\omega}_t} dF(\omega) \\ = \ & i_t \left[1 - \mu F(\bar{\omega}_t)\right]. \end{aligned}$$

Since there are η entpreprenuers, the total new capital produced is $k_{t+1} = \eta i_t \left[1 - \mu F(\bar{\omega}_t)\right]$, so that

$$k_{t+1} = \left[(1-\delta)\, k_t + x_t - \Phi\left(\frac{x_t}{k_t}\right) k_t\right] \left[1 - \mu F(\bar{\omega}_t)\right]. \tag{C.18}$$

The resource constraint is:

$$(1-\eta)\, c_t + \eta c_t^e + x_t = k_t^{\alpha} \left((1-\eta)\, Z_t l_t\right)^{1-\alpha-\zeta} \eta^{\zeta}. \tag{C.19}$$

Substituting (C.17) into (C.6) and (C.7):

$$1 = \beta E_t \frac{u_{c,t+1}}{u_{c,t}} \frac{y_{k,t+1} + P_{k,t+1}}{q_t} \tag{C.20}$$

$$\frac{-u_{l,t}}{u_{c,t}} = y_{l,t}. \tag{C.21}$$

The budget constraint of the entrepreneur is:

$$c_{et} + q_t k_{et+1} = \lambda\left(P_{k,t}\right) \frac{k_{t+1}}{\eta\left[1 - \mu F\left(\bar{\omega}_t\right)\right]} q_t f\left(\bar{\omega}_t\right) \tag{C.22}$$

The efficiency conditions associated with the contract are:

$$q_t f\left(\bar{\omega}_t\right) = \frac{f'\left(\bar{\omega}_t\right)}{g'\left(\bar{\omega}_t\right)} \left[q_t g\left(\bar{\omega}_t\right) - \lambda\left(P_{k,t}\right)\right] \tag{C.23}$$

$$\frac{k_{t+1}}{\eta\left[1 - \mu F\left(\bar{\omega}_t\right)\right]} = \frac{a_t}{\lambda\left(P_{k,t}\right) - q_t g\left(\bar{\omega}_t\right)} \tag{C.24}$$

$$a_t = y_{3,t} + y_{k,t} k_{et} + P_{k,t} k_{et} \tag{C.25}$$

The intertemporal efficiency condition for the entrepreneur is (assuming the condition, $\zeta_{t+1} \geq 1$, is not binding):

$$E_t \left[\frac{F_{k,t+1} + P_{k,t+1}}{q_t} \times \frac{q_{t+1} f\left(\bar{\omega}_{t+1}\right)}{\lambda\left(P_{k,t+1}\right) - q_{t+1} g\left(\bar{\omega}_{t+1}\right)}\right] = \frac{1}{\gamma\beta} \tag{C.26}$$

Taking the ratio of (C.8) to (C.9), we obtain:

$$P_{k,t} = \frac{1 - \delta - \Phi\left(\frac{x_t}{k_t}\right) + \Phi'\left(\frac{x_t}{k_t}\right) \frac{x_t}{k_t}}{1 - \Phi'\left(\frac{x_t}{k_t}\right)} \tag{C.27}$$

The 10 variables to be determined with the 10 equations, (C.18)-(C.27) are: c_t, $c_{e,t}$, x_t, k_t, $k_{e,t}$, l_t, $P_{k,t}$, q_t, $\bar{\omega}_t$, a_t.

It is convenient to define a sequence of markets equilibrium formally. Let s^t denote a history of realizations of shocks. Then,

Definition 1. *An equilibrium of the CF economy with adjustment costs is a sequence of prices,* $\{P_k(s^t), q(s^t), w^e(s^t), w^c(s^t), r(s^t)\}$ *, quantities,* $\{c(s^t), c_e(s^t), x(s^t), k(s^t), k_e(s^t), l(s^t), a(s^t)$ *and* $\{\bar{\omega}(s^t)\}$ *such that:*

(i) Households optimize (see (C.20), (C.21))

(ii) Entrepreneurs optimize (see (C.22), (C.25), (C.26), (C.27))

(iii) Firms optimize (see (C.17))

(iv) Conditions related to the standard debt contract are satisfied (see (C.23), (C.24))

(v) The resource constraint and capital accumulations equations are satisfied (see (C.18), (C.19))

C.1.3. The CF Model as an RBC Model with Wedges

We now construct a set of wedges for the RBC economy in section C.1.1, such that the equilibrium for the distorted version of that economy coincides with the equilibrium for the CF economy. We begin by constructing the following state-contingent sequences:

$$
\begin{aligned}
\psi\left(s^{t}\right) &= 1-\mu F\left(\tilde{\bar{\omega}}\left(s^{t}\right)\right), \\
\tau_{x}\left(s^{t}\right) &= \frac{\psi\left(s^{t}\right)\tilde{q}\left(s^{t}\right)}{\lambda\left(\tilde{P}_{k,t}\left(s^{t}\right)\right)}-1, \\
\theta\left(s^{t}\right) &= \frac{\tilde{P}_{k}\left(s^{t}\right)\tau_{x}\left(s^{t}\right)}{\tilde{r}\left(s^{t}\right)}, \\
G\left(s^{t}\right) &= \eta\left(\tilde{c}^{e}\left(s^{t}\right)-\tilde{c}\left(s^{t}\right)\right), \\
T\left(s^{t}\right) &= G\left(s^{t}\right)-\tau_{x}\left(s^{t}\right)\tilde{x}\left(s^{t}\right)-\theta\left(s^{t}\right)\tilde{r}\left(s^{t}\right)\tilde{k}\left(s^{t-1}\right), \\
D\left(s^{t}\right) &= \tilde{w}^{e}\left(s^{t}\right)\eta,
\end{aligned}
\tag{C.28}
$$

where $\tilde{q}$, $\tilde{c}^{e}$, $\tilde{c}$, $\tilde{w}^{e}$, $\tilde{r}$, $\tilde{k}$, $\tilde{x}$, $\tilde{\bar{\omega}}$ and $\tilde{P}_k$ correspond to the objects without ' $\tilde{\cdot}$ ' in a CF equilibrium. Also, λ is the function defined in (C.10). In this subsection, we treat D, ψ, θ, τ_x, G and T as given exogenous stochastic processes, outside the control of agents. Here, D, G, and T represent exogenous sequences of profits, government spending and lump sum taxes. Also, θ and τ_x are tax rates on capital rental income and investment good purchases. Finally, ψ is a technology shock in the production of physical capital.

Consider the following budget constraint for the household:

$$
\begin{aligned}
& c\left(s^{t}\right)+\left(1+\tau_{x}\left(s^{t}\right)\right)x\left(s^{t}\right) \\
\leq\ & \left(1-\theta\left(s^{t}\right)\right)r\left(s^{t}\right)k\left(s^{t-1}\right)+w\left(s^{t}\right)l\left(s^{t}\right)-T\left(s^{t}\right)+D\left(s^{t}\right).
\end{aligned}
\tag{C.29}
$$

Here, r is the rental rate on capital, w is the wage rate, and l measures the work effort of the household. Each of these variables is a function of s^t and is determined in an RBC wedge economy. At time 0 the household takes prices, taxes and $k\left(s^{-1}\right)$ as given and chooses c, k and l to maximize utility:

$$\sum_{t=0}^{\infty}\sum_{s^t}\beta^t\pi\left(s^t\right)u\left(c\left(s^t\right),l\left(s^t\right)\right)$$

subject to the budget constraint, no-Ponzi game and non-negativity constraints. Here, $\pi\left(s^t\right)$ is the probability of history, s^t.

Households operate the following backyard technology to produce new capital:

$$k\left(s^t\right)=\left[(1-\delta)\,k\left(s^t\right)+x\left(s^t\right)-\Phi\left(\frac{x\left(s^t\right)}{k\left(s^{t-1}\right)}\right)k\left(s^{t-1}\right)\right]\psi\left(s^t\right). \tag{C.30}$$

The first order necessary conditions for household optimization are:

$$u_l\left(s^t\right)+u_c\left(s^t\right)w\left(s^t\right)=0, \tag{C.31}$$

$$\begin{aligned}&\frac{1+\tau_x\left(s^t\right)}{\psi\left(s^t\right)}P_{k'}\left(s^t\right) \qquad \text{(C.32)}\\ =\;&\sum_{s^{t+1}|s^t}\beta\pi\left(s^{t+1}|s^t\right)\frac{u_c\left(s^{t+1}\right)}{u_c\left(s^t\right)}\left[r\left(s^{t+1}\right)\left(1-\theta\left(s^{t+1}\right)\right)+\left(1+\tau_x\left(s^{t+1}\right)\right)P_k\left(s^{t+1}\right)\right],\end{aligned}$$

where

$$P_k\left(s^t\right)\equiv\frac{1-\delta-\Phi\left(\frac{x\left(s^t\right)}{k\left(s^t\right)}\right)+\Phi'\left(\frac{x\left(s^t\right)}{k\left(s^t\right)}\right)\frac{x\left(s^t\right)}{k\left(s^t\right)}}{1-\Phi'\left(\frac{x\left(s^t\right)}{k\left(s^{t-1}\right)}\right)}. \tag{C.33}$$

Equation (C.31) is the first order condition associated with the optimal choice of $l\left(s^t\right)$. Equation (C.32) combines the first order order conditions associated with the optimal choice of $x(s^t)$ and $k(s^t)$. Also,

$$P_{k'}\left(s^t\right) \equiv \frac{1}{1-\Phi'\left(\frac{x(s^t)}{k(s^{t-1})}\right)}, \tag{C.34}$$

is the pre-tax marginal cost of producing new capital, in units of the consumption good. In addition, $\pi(s^{t+1}|s^t) \equiv \pi(s^{t+1})/\pi(s^t)$ is the conditional probability of history s^{t+1} given s^t.

The technology for firms is taken from (C.16):

$$y(k,l,\eta,Z) = k^\alpha\left((1-\eta)Zl\right)^{1-\alpha-\zeta}\eta^\zeta,$$

where, as before, the third argument in y refers only to the second occurrence of η. There are three inputs: physical capital, household labor and another factor whose aggregate supply is fixed at η. Profit maximization leads to:

$$r\left(s^t\right) = y_k\left(s^t\right),\ w\left(s^t\right) = y_l\left(s^t\right),\ w^e\left(s^t\right) = y_\eta\left(s^t\right). \tag{C.35}$$

The household is assumed to own the representative firm, and it receives the earnings of η in the form of lump-sum profits, $D(s^t)$. We do allow allow trade in claims on firms, a restriction that is non-binding on allocations because the households are identical.

We now state the equilibrium for the RBC wedge economy:

Definition 2. *An RBC wedge equilibrium is a set of quantities,* $\{c(s^t), l(s^t), k(s^t), x(s^t)\}$, *and prices* $\{P_k(s^t), P_{k'}(s^t), r(s^t), w(s^t)\}$, *and a set of taxes, profits and government spending,* $\{G(s^t), \tau_x(s^t), \theta(s^t), T(s^t)\}$, *technology shocks,* $\{Z(s^t), \psi(s^t)\}$, *such that* *(i) The quantities solve the household problem given the prices, taxes, profits, government spending and the shock to the backyard investment technology*

(ii) Firm optimization is satisfied

(iii) Relations (C.28) is satisfied, for given state-contingent sequences, $\tilde{q}$, $\tilde{c}^e$, $\tilde{c}$, $\tilde{w}^e$, $\tilde{r}$, $\tilde{k}$, $\tilde{x}$, $\tilde{\tilde{\omega}}$ *and* $\tilde{P}_k$.

The variables to be determined in an RBC wedge equilibrium are c, l, k, x, P_k, $P_{k'}$, r and w. The 8 equations that can be used to determine these are (C.29)-(C.35). It is easily verified that c, l, k, x, P_k, r and w coincide with the corresponding objects in a CF equilibrium. In addition, $P_{k'}$ coincides with $\lambda\left(P_k\right)$ in a CF equilibrium. To see this, one verifies that the equilibrium conditions in the RBC wedge economy coincides with the equilibrium conditions in the CF economy. First, (C.30) coincides with (C.18). After using (C.35), we see that (C.31) coincides with (C.21). Consider the household budget equation evaluated at equality. Substituting out for lump sum transfers:

$$\begin{aligned}
& c\left(s^t\right)+\left(1+\tau_x\left(s^t\right)\right)x\left(s^t\right) \\
= \; & \left(1-\theta\left(s^t\right)\right)r\left(s^t\right)k\left(s^{t-1}\right)+w\left(s^t\right)l\left(s^t\right) \\
& \tau_x\left(s^t\right)x\left(s^t\right)+\theta\left(s^t\right)r\left(s^t\right)k\left(s^{t-1}\right)+w^e\left(s^t\right)\eta+G\left(s^t\right),
\end{aligned}$$

or,

$$\begin{aligned}
& (1-\eta)\,c\left(s^t\right)+x\left(s^t\right)+\eta c^e\left(s^t\right) \\
= \; & r\left(s^t\right)k\left(s^{t-1}\right)+w\left(s^t\right)l\left(s^t\right)+w^e\left(s^t\right)\eta \\
= \; & y\left(k\left(s^{t-1}\right),l\left(s^t\right),\eta,Z\left(s^t\right)\right),
\end{aligned} \tag{C.36}$$

by linear homogeneity. Here, $Z(s^t) = Z(s_t)$, where s_t is the realization of period t uncertainty. Equation (C.36) coincides with (C.19). Substitute out for θ and τ_x from (C.28) into (C.32), and rearranging, we obtain:

$$1 = E_t \frac{u_c(s^{t+1})}{u_c(s^t)} \left[\frac{r(s^{t+1}) + P_k(s^{t+1})}{\frac{1+\tau_x(s^t)}{\psi(s^t)} P_{k'}(s^t)} \right].$$

Note that by definition of $1 + \tau_x(s^t)$ in (C.28),

$$\frac{1+\tau_x(s^t)}{\psi(s^t)} P_{k'}\left(s^t\right) = \frac{P_{k'}(s^t)\, q(s^t)}{\lambda\left(P_{k,t}(s^t)\right)}.$$

Combining (C.10) and (C.9), we find that $\lambda\left(P_{k,t}(s^t)\right) = P_{k'}(s^t)$, so that the household's intertemporal Euler equation reduces to (C.20), or (after making use of (C.35)):

$$E_t \frac{u_c(s^{t+1})}{u_c(s^t)} \left[\frac{y_k(s^{t+1}) + P_k(s^{t+1})}{P_{k'}(s^t)} \left(1 - \tau^k\left(s^t\right)\right) \right] = 1, \tag{C.37}$$

where

$$1 - \tau^k\left(s^t\right) = \frac{\psi(s^t)}{1 + \tau_x(s^t)}.$$

We conclude that conditions (C.18)-(C.21) in the CF economy are satisfied. The remaining equilibrium conditions are satisfied, given (C.28). We state this result as a proposition:

Proposition 1. *Consider a CF equilibrium, and a set of taxes, technology shocks and transfers computed in (C.28). The objects,* $\{c(s^t), l(s^t), k(s^t), x(s^t)\}$, $\{P_k(s^t), r(s^t), w(s^t)\}$ *and* $P_{k'}(s^t) = \lambda(P_k(s^t))$ *in the CF equilibrium correspond to an RBC wedge equilibrium.*

For η and ζ close to zero and ψ close to unity, the RBC wedge equilibrium converges to the equilibrium conditions of the RBC model with adjustment costs in section C.1.1 with a wedge, $1-\tau^k$, in the intertemporal Euler equation, (C.4).

C.2. The Bernanke-Gertler-Gilchrist Financial Friction Wedge

In this section we briefly review the BGG model and derive the RBC wedge model to which it corresponds. In the model there are households, capital producers, entrepreneurs and banks. At the beginning of the period, households supply labor to factor markets, and entrepreneurs supply capital. Output is then produced and an equal amount of income is distributed among households and entrepreneurs. Households then make a deposit with banks, who lend the funds on to entrepreneurs. Entrepreneurs have a special expertise in the ownership and management of capital. They have their own net worth with which to acquire capital. However, it is profitable for them to leverage this net worth into loans from banks, and acquiring more capital than they can afford with their own resources. The source of friction is a particular conflict between the entrepreneur and the bank. In the management of capital, idiosyncratic things happen, which either make the management process more profitable than expected, or less so. The problem is that the things that happen in this process are observed only by the entrepreneur. The bank can observe what happens inside the management of capital, but only at a cost. As a result, the entrepreneur has an incentive to underreport the results to the bank, and thereby attempt to keep a greater share of the proceeds for himself. To mitigate this conflict, it is assumed that entrepreneurs receive a standard debt contract from the bank.

The capital that entrepreneurs purchase at the end of the period is sold to them by capital producers. The latter use the old capital used within the period, as well as investment goods, to produce the new capital that is sold to the entrepreneurs. Capital producers have no external financing need. They finance the purchase of used capital and investment goods using the revenues earned from the sale of new capital.

The budget constraint of households is:

$$c_t + B_{t+1} \leq (1 + R_t)\, B_t + w_t l_t + T_t,$$

where R_t denotes the interest earned on deposits with the bank, b_t denotes the beginning-of-period t stock of those deposits, w_t denotes the wage rate, l_t denotes employment and T_t denotes lump sum transfers. Subject to this budget constraint and a no-Ponzi condition, households seek to maximize utility:

$$E_0 \sum_{t=0}^{\infty} \beta^t u\,(c_t, l_t)\,.$$

Households' first order conditions, in addition to the transversality condition, are:

$$\begin{aligned} u_{c,t} &= \beta E_t u_{c,t}\,(1 + R_{t+1}) \\ \frac{-u_{l,t}}{u_{c,t}} &= w_t. \end{aligned}$$

Firms have the following production function:

$$y_t = k_t^{\alpha}\,(Z_t l_t)^{1-\alpha} = y\,(k_t, l_t, Z_t)\,.$$

They rent capital and hire labor in perfectly competitive markets at rental rate, r_t, and wage rate, w_t, respectively. Optimization implies:

$$y_{k,t} = r_t,\ y_{l,t} = w_t.$$

Capital producers purchase investment goods, x_t, and old capital, k_t, to produce new capital, k_{t+1}, using the following linear homogeneous technology:

$$k_{t+1} = (1-\delta)\,k_t + x_t - \Phi\left(\frac{x_t}{k_t}\right)k_t.$$

The competitive market prices of k_t and k_{t+1} are $P_{k,t}$ and $P_{k',t}$, respectively. Capital producer optimization leads to the following conditions:

$$\begin{aligned} P_{k,t} &= \frac{1}{1-\Phi'\left(\frac{x_t}{k_t}\right)}\left[1-\delta-\Phi\left(\frac{x_t}{k_t}\right)+\Phi'\left(\frac{x_t}{k_t}\right)\frac{x_t}{k_t}\right] \\ P_{k',t} &= \frac{1+g_n}{1-\Phi'\left(\frac{x_t}{k_t}\right)}. \end{aligned}$$

At the end of period t, entrepreneurs have net worth, N_{t+1}, and it is assumed that $N_{t+1} < P_{k',t}k_{t+1}$. As a result, in an equilibrium in which the entire stock of capital is to be owned and operated, entprepreneurs must borrow:

$$b_{t+1} = P_{k',t}k_{t+1} - N_{t+1}. \tag{C.38}$$

As soon as an individual entrepreneur purchases k_{t+1}, he experiences a shock, and k_{t+1} becomes $k_{t+1}\omega$. Here, ω is a random variable that is iid across entrepreneurs and has mean unity. The realization of ω is unknown before the loan is made and it is known only to the entrepreneur after it is realized. The bank which extends the loan to the entrepreneur must pay a monitoring cost in order to observe the realization of ω. The cumulative distribution function of ω is F, where

$$\Pr ob\left[\omega < x\right] = F\left(x\right).$$

Entrepreneurs receive a standard debt contract from their bank, which specifies a loan amount, b_{t+1}, and a gross rate of return, Z_{t+1}, in the event that it is feasible for the entrepreneur to repay. The lowest realization of ω for which it is feasible to repay is $\bar{\omega}_{t+1}$, where

$$\bar{\omega}_{t+1}\left(1+R_{t+1}^{k}\right)P_{k',t}k_{t+1} = Z_{t+1}b_{t+1}. \tag{C.39}$$

For $\omega < \bar{\omega}_{t+1}$ the entrepreneur simply pays all its revenues to the bank:

$$\left(1+R_{t+1}^{k}\right)\omega P_{k',t}k_{t+1}.$$

In this case, the bank monitors the entrepreneur. Following BGG, we assume that monitoring costs are a fraction, μ, of the total earnings of the entrepreneur:

$$\mu\left(1+R_{t+1}^{k}\right)\omega P_{k',t}k_{t+1}.$$

At time t the bank borrows b_{t+1} from households. In each state of $t+1$ the bank pays households

$$\left(1+R_{t+1}\right)b_{t+1} \tag{C.40}$$

units of currency. The bank's source of funds in each state of period $t+1$ is the earnings from non-bankrupt entrepreneurs plus the earnings of bankrupt entrepreneurs, net of monitoring costs:

$$[1-F(\bar{\omega}_{t+1})]\, Z_{t+1}b_{t+1} + (1-\mu)\int_0^{\bar{\omega}_{t+1}} \omega dF(\omega)\left(1+R^k_{t+1}\right) P_{k',t}k_{t+1}. \tag{C.41}$$

We follow BGG, who implicitly suppose that at date t there are no state-contingent markets for currency in date $t+1$. As a consequence, (C.40) must not exceed (C.41) in any date $t+1$ state. This condition, together with competition among banks, leads to:

$$[1-F(\bar{\omega}_{t+1})]\, Z_{t+1}b_{t+1} + (1-\mu)\int_0^{\bar{\omega}_{t+1}} \omega dF(\omega)\left(1+R^k_{t+1}\right) P_{k',t}k_{t+1} = (1+R_{t+1})\, b_{t+1}.$$

Substituting from (C.39) for $Z_{t+1}b_{t+1}$ and dividing by $\left(1+R^k_{t+1}\right) P_{k',t}k_{t+1}$:

$$[1-F(\bar{\omega}_{t+1})]\, \bar{\omega}_{t+1} + (1-\mu)\int_0^{\bar{\omega}_{t+1}} \omega dF(\omega) = \left(\frac{1+R_{t+1}}{1+R^k_{t+1}}\right)\frac{b_{t+1}}{P_{k',t}k_{t+1}}.$$

We conclude that the gross return on capital can be expressed:

$$1+R_{t+1} = \left(1-\tau^k_{t+1}\right)\left(1+R^k_{t+1}\right),$$

where the 'wedge', $1-\tau^k_{t+1}$, satisfies:

$$1-\tau^k_{t+1} = \frac{P_{k',t}k_{t+1}}{P_{k',t}k_{t+1}-N_{t+1}}\left([1-F(\bar{\omega}_{t+1})]\, \bar{\omega}_{t+1} + (1-\mu)\int_0^{\bar{\omega}_{t+1}} \omega dF(\omega)\right).$$

The wedge, τ^k_t, contains two additional endogenous variables, N_{t+1} and $\bar{\omega}_{t+1}$. These are determined in general equilibrium by the introduction of two additional equations: the condition associated with the fact that the standard debt contract maximizes the utility of the entrepreneur, as well as the law of motion for entrepreneurial net worth.

The resource constraint for this economy is:

$$c_t + G_t + x_t = k_t^{\alpha} \left(Z_t l_t \right)^{1-\alpha},$$

where G_t includes any consumption of entrepreneurs, as well as monitoring costs incurred by banks. As long as these latter can be ignored, then the BGG financial friction is to, in effect, introduce a tax on the rate of return on capital in, $1+R_{t+1}^k$, in (C.4). In particular, $1+R_{t+1}^k$ is replaced by

$$\left(1+R_{t+1}^k\right)\left(1-\tau_{t+1}^k\right).$$

Note there is a slight difference with CF financial frictions in that the latter imply the tax rate is not a function of period $t+1$ uncertainty, while the BGG frictions imply that in general it is a function of this uncertainty.

C.3. The Linearized Model

Preferences are:

$$E \sum_{t=0}^{\infty} \beta^t \left[\log c_t + \psi \log \left(1 - l_t\right) \right],$$

and the resource constraint is:

$$c_t + G_t + x_t \leq y\left(k_t, l_t, Z_t\right) = k_t^{\alpha} \left(Z_t l_t \right)^{1-\alpha},$$

where

$$Z_t = \tilde{Z}_t \left(1 + g_z\right)^t.$$

Government spending is G_t, where

$$G_t = \tilde{g}_t \left(1 + g_z\right)^t .$$

The law of motion for capital is:

$$k_{t+1} = (1 - \delta) k_t + x_t - \Phi \left(\frac{x_t}{k_t} \right) k_t,$$

where

$$\Phi \left(\frac{x_t}{k_t} \right) = \frac{a}{2} \left(\frac{x_t}{k_t} - b \right)^2 .$$

Here, b is the steady state investment to capital ratio.and $\tilde{Z}_t$, the efficiency wedge, is an exogenous stationary stochastic process.

The law of motion of the exogenous shocks is:

$$s_t = P_0 + P s_{t-1} + Q \varepsilon_t, \; s_t = \begin{pmatrix} \log \tilde{Z}_t \\ \tau_{l,t} \\ \tau_{x,t} \\ \log \tilde{g}_t \end{pmatrix}, \; E \varepsilon_t \varepsilon_t' = I,$$

where

$$P = \begin{bmatrix} \bar{P} & 0 \\ 0 & p_{44} \end{bmatrix}, \; Q = \begin{bmatrix} \bar{Q} & 0 \\ 0 & q_{44} \end{bmatrix}.$$

The intratemporal Euler equation is:

$$\frac{-u_{l,t}}{u_{c,t}} = (1 - \tau_{l,t}) y_{l,t},$$

or,

$$\frac{\psi c_t}{1-l_t} = (1-\tau_{l,t})\,(1-\alpha)\left(\frac{k_t}{l_t}\right)^{\alpha} Z_t^{(1-\alpha)}$$

The intertemporal Euler equation is:

$$\frac{1}{c_t} = \beta E_t \frac{1}{c_{t+1}} \left(1-\tau_{t+1}^{k}\right) \frac{y_{k,t+1}+P_{k,t+1}}{P_{k',t}},$$

where,

$$\begin{aligned} P_{k',t} &= \frac{1}{1-\Phi'\left(\frac{x_t}{k_t}\right)} \\ P_{k,t} &= \frac{1-\delta-\Phi\left(\frac{x_t}{k_t}\right)+\Phi'\left(\frac{x_t}{k_t}\right)\frac{x_t}{k_t}}{1-\Phi'\left(\frac{x_t}{k_t}\right)}. \end{aligned}$$