Consumption and Information: A Study of Consumer Behavior using Daily Data*

Marta Lachowska

Abstract

The availability of high frequency data on expenditure and information provides an opportunity to test models of consumption decisions. This paper studies the joint dynamics of information and consumption using data at a daily frequency. I find that spending reacts sharply to shocks to information, but in contrast to previous research findings, this reaction fades within a very short period of time. Additionally, my data allow me to move beyond representative agent models in studying the response of individuals facing different levels of income stability. Unlike papers using aggregate data, I am able to contrast the reactions of different types of consumers. I find that individuals who face less secure income streams cut back more than those with a secure income. I show that this behavior of consumption cannot be adequately explained by canonical consumption theories, such as the permanent income hypothesis model or the buffer stock model.

*I would like to thank Lars Lefgren for his advice and support. Further, I am grateful to Alan Krueger for kindly enabling me to work with the G1K data and to Jim Harter and Dennis Jacobe from the Gallup Organization for their support. I am especially indebted to Henry Farber, John Hassler, Matthew Lindquist, Andreas Müller, Zhuan Pei, Åsa Rosén, Sam Schulhofer-Wohl and Eskil Wadensjö for their comments. I have benefitted from valuable discussions with Orley Ashenfelter, José Azar, Markus Jäntti, Jan Klingelhöfer and Mark Watson. Further, I thank the seminar participants at Princeton University, Stockholm University and SOFI for their comments. I thank the Jan Wallander and Tom Hedelius' Foundation for generous financial support. All errors are my own.

1 Introduction

In 2007, the United States entered one of the most severe recessions since the 1930s.[1] The following year was marked by the volatility of the stock markets, collapse of several investment banks and major bailouts for troubled financial institutions. Some have dubbed the economic downturn as the Great Recession – a wordplay on the Great Depression.

The severity of the recession has highlighted the ongoing need to understand the underlying causes of macroeconomic distress. A large literature on real business cycles attributes recessions to fluctuations in technology. An older view, associated primarily with Keynes, ascribes the variations in economic activity to fluctuations in consumption demand. In recent years we have witnessed a resurgence in the interest in bridging the divide between the latter view and the real approach. One method has been to model news shocks as generating business cycles. Such shocks have effects on consumer expectations and thus impact consumption decisions and output. As consumption behavior is a key aspect of macroeconomic modeling, it is crucial to understand how spending adapts to changing expectations.

In this paper, I test the predictions of canonical models of consumption theory. In particular, I examine the implications of models in which information shocks affect income. First, I discuss the implications of two models: the permanent income hypothesis model (Friedman (1957), Hall (1978), Deaton 1992, chapter 3) and the buffer stock model, also known as the precautionary savings model (Zeldes (1989), Kimball (1990) and Carroll (1997)). Second, I discuss testable predications of a simple signal-extraction model, where consumers face incomplete information. The latter generates behavior where economic agents overreact to information in the short run.

To do this, I examine the response of individual expenditure levels to the arrival of economic information in a high frequency setting. I regress daily consumption on a proxy for information flows regarding economic prospects. The data on consumption comes at a *daily* frequency, which allows me to observe the immediate reaction of spending to information. Is there a response and if so, how long does it last? I allow for the full endogeneity of my variables by studying the dynamics in a vector autoregressive model (VAR).[2]

My window of observation covers the year 2008. This period is of particular interest because it corresponds to the start of a deep US recession. I can thus study the sensitivity of

[1] This has manifested itself in, for example, one of the highest unemployment rates in recent U.S. history (10.2 percent in October 2009) and the largest contraction (in the last quarter of 2008) of Gross Domestic Product since 1982.

[2] Throughout this paper I use the terms consumption and expenditure interchangeably.

spending to a highly uncertain and fluctuating economic environment. The data come from the Gallup Daily Poll conducted by the Gallup Organization. In addition to information on daily spending on services and goods, the survey collects a rich set of demographic variables. The micro-data allow me to move beyond representative agent models in studying the response of individuals facing different levels of income stability.

In order to capture the impact of information on consumption, I use high frequency data from the world's most popular Internet search engine: Google. These data allow me to study the arrival of economic news by tracking variations in the search activity for particular queries on a daily basis by means of a search volume index. The behavior of searches for certain phrases across time in 2008 allows for a comparable time series reflecting the "zeitgeist" of this turbulent year.[3]

The main findings of the paper are as follows. Consumption drops by 2-4 percent in the days following a one standard deviation shock to adverse economic information. Subsequently, this cutback is reversed within 2-3 weeks. Individuals who face less secure income streams display a stronger reaction. In particular, this is true of those who are jobless or below the typical retirement age. My results indicate that, contrary to the permanent income model, consumption at this frequency does not appear to have an infinite memory. In addition, I argue that the rapid deterioration following the cutback cannot be fully explained by the precautionary saving motive. My findings are, however, in line with the predictions of a signal extraction model of consumption where agents react prudently to incomplete information (Wang 2004).

Most real business cycle models featuring news shocks share some common features. The term news is understood in this literature as a mechanism conveying relevant information about future productivity, orthogonal to the current information set. These models extend the information structure by introducing a signal extraction problem the agents solve. Following a noisy shock, the agents learn to separate the noise from the fundamental. As this proceeds, it may cause a change in expectations about the economy. Such a change in turn affects demand which causes output to fluctuate in the short run. Thus,

[3] Data originating from Internet searches have been used in a few studies. Ginsberg et al. (2009) use Google Trends to estimate the level of flu activity in the United States by correlating search queries for flu with the percentage of flu related visits to a physician. Azar (2009) studies the joint dynamics of an increase in interest in electric cars and oil prices in a Bayesian VAR. Choi and Varian (2009b) study how data from Google Trends forecasts economic activity. They illustrate their approach with queries on specific products, e.g. motorcycles and show that the time series obtained predicts sales. In Choi and Varian (2009a), the authors extend the analysis by forecasting initial unemployment benefit take-up using searches on welfare and unemployment. This approach is also employed by Askitas and Zimmerman (2009) on German data.

the expectation-driven business cycles allow for behavior associated with "animal spirits" in an otherwise rational setting.[4] See Beaudry and Portier (2004), Beaudry and Portier (2006), Jaimovich and Rebelo (2009), Sims (2009a), Lorenzoni (2009) and Barsky and Sims (2009b) for recent examples of news-driven RBC models.

Blanchard et al. (2009) contrast the effects of a shock to information on spending and output. Their paper studies the response of consumption and tries to contrast the paths it will take following a fundamental change as opposed to that of noise. The paper by Blanchard et al. (2009) is similar to a paper by Barsky and Sims (2009a). Whereas the first deals with productivity and expenditures, the paper by Barsky and Sims studies the impulse response functions of income and expenditure following a shock to consumer confidence. The path of consumption following a shock to sentiment can speak in favor of either an "animal spirits" interpretation or an information view. The "animal spirits" view implies that following a shock to consumer confidence, there will be an initial reaction of expenditure followed by decay. Conversely, the information view suggests a gradual reaction of spending. Barsky and Sims find support for the information view in their analysis.

My approach differs in several ways. First, I use a different measure of information. I study the impact of information by using data on individuals' actual searches for economic information. The advantage of using this measure is that it captures the variation of individuals' revealed interest in information.

Second, the information in the micro-data can be used to identify groups facing different risks in their income process. Theory suggests that individuals who face less secure income streams ought to cut back more than those with a secure income. Unlike papers using aggregate data, I contrast the reactions of different types of consumers.

Third, I use data with a much higher frequency. The high frequency of data is paramount, as it captures the quick, often daily variation in the arrival of information. In the process of aggregating to low frequency data, we lose information on high frequency consumption changes. This limits the ability to distinguish between competing explanations for consumption behavior. High frequency changes in consumption and the subsequent reversion to trend are inconsistent with many standard explanations of consumption behavior. This cannot, however, be tested using data at a monthly frequency. My paper sheds light on aspects of consumer behavior that an aggregate analysis omits.

[4] There exists a branch of models, which deal with so-called sunspot theories. These RBC models have microeconomic foundations, where the agents are rational. Fluctuations in output are caused by a multiplicity of equilibria, thus explaining an "animal spirits" like behavior, see for example Azariadis (1981). This paper does not address sunspot models.

The paper is organized as follows. Section two discusses theoretical considerations of modern consumption theory, presents a model with imperfect information and discusses its implications. The following section discusses the implications of theory for the empirical model. In section four, I present the data. Section five outlines the results. The final section concludes.

2 Theoretical Framework

In this section, I begin with a brief discussion on how consumption might react to adverse information regarding economic prospects. I follow up the empirical implications with a formal discussion.

The arrival of adverse information about an agent's economic prospects, e.g. permanent income, should be associated with a change in consumption. If the permanent income hypothesis holds, the full impact of information should be proportional to the change and persist indefinitely. I contrast this to a case where consumption declines following a shock to information, but subsequently reverts to its former level. Transitory movement of consumption might suggest that consumers cut back on spending when their economic prospects are uncertain. As the consumers draw conclusions from the fundamental information, they readjust their spending accordingly.

Consider the reduced form relationship between consumption at date t and a measure of information, for example income:

$$c_t = \gamma c_{t-1} + \theta_0 info_t + \theta_1 info_{t-1} + ... + \theta_p info_{t-p} + e_t.$$

The permanent income hypothesis suggests that c_{t-1} is a sufficient statistic for all the information available up to time $t-1$ (this sufficiency is the essence of Hall's (1978) test). This implies that γ ought to be significant and, if consumption is a random walk, it should equal one. If so, following an arrival of adverse information at date t, we expect $\theta_0 < 0$. If c_{t-1} embodies all the available information up to date $t-1$, the lagged values of $info$ at $t-1, t-2, ..., t-p$ ought not to matter. This is summarized in the following hypothesis:[5]

$$H^{PIH} : \gamma = 1, \theta_0 < 0, \theta_{i \neq 0} = 0.$$

[5]For simplicity, I ignore the intercept of this relationship. If the lagged value of consumption summarizes all the information relevant to current consumption, the intercept should be zero.

The second case considered above violates this null. If a decline in consumption following a shock to information is transitory, then consumption does not follow a random walk. This suggests that γ need not equal unity and that c_{t-1} is not a sufficient statistic for all the events up to $t-1$. This in turn implies that the lagged values of $info$ may be statistically different from zero.

In order to formalize the considerations outlined above, I will present a set of models of consumption theory. I will describe the canonical model of consumption theory, the permanent income model (for a textbook treatment, see Deaton (1992)). Later, I will discuss the implications of a model dropping the certainty equivalence assumption. Finally, I contrast both with a model involving incomplete information.

2.1 A Permanent Income Hypothesis Model

Consider an infinitely-lived consumer with preferences $\sum_{t=0}^{\infty}\beta^t u(c_t)$ where c denotes consumption. The parameter $\beta\in(0,1)$ is the subjective time discount factor. Following Hall (1978), the functional form of utility is quadratic in consumption: $u(c)=\alpha_1 c_t-\frac{\alpha_2}{2}c_t^2$.[6] The consumer maximizes utility subject to a dynamic budget constraint: $k_{t+1}=(1+r)k_t+y_t-c_t$, where r is the interest rate and k_t is the level of assets.[7] There is uncertainty regarding the income process; the objective function is the expectation of the discounted sum of future utility streams. Solving the standard maximization problem yields a Euler equation in consumption. This first order condition characterizes the process of the level of consumption:

$$\mathbb{E}\left(c_{t+1}|\mathbb{I}^t\right)=c_t\Longrightarrow c_{t+1}-c_t=\varepsilon_{t+1} \tag{1}$$

where $\mathbb{I}^t$ is the set of all information available up to time t, $\mathbb{I}^t=\{\mathbb{I}_t,\mathbb{I}_{t-1},...,\mathbb{I}_0\}$ and $\mathbb{E}\left(\varepsilon_{t+1}|\mathbb{I}^t\right)=0$.[8] The empirical implication of this result is that a change in consumption from t to $t+1$ cannot be forecasted with the information at time t, as summarized by $\mathbb{I}^t$. This condition is sometimes referred to as the orthogonality condition: ε_{t+1} is orthogonal to the information set $\mathbb{I}^t$. From this expression we can derive the condition that the level of consumption should change following unexpected changes to permanent income: $\Delta c_{t+1}=\varepsilon_{t+1}$. This error term, ε_{t+1}, is an innovation to consumption.

[6]More precisely, $u(\bullet)$ is an "instantaneous utility" function, sometimes also called the felicity function. The cumulative discounted sum $U=\sum_{t=0}^{\infty}\beta^t u(c_t)$ is the von Neuman-Morgenstern utility.

[7]I assume that $\beta^{-1}=1+r$.

[8]In the paper I use the notation $\mathbb{E}\left(c_{t+1}|\mathbb{I}^t\right)=\mathbb{E}_t\left(c_{t+1}\right)$ interchangeably.

As outlined above, this simple rational expectations model suggests that consumption ought to be a martingale. In the context of an information shock regarding permanent income, this suggests that new information about an agent's permanent income should be associated with a change in consumption which persists indefinitely – this is the infinite memory property of unit root processes. The full impact of information should be proportional to the drop in permanent income. This case corresponds to H^{PIH}.

2.2 A Buffer Stock Model

The assumption of quadratic utility provides an analytic convenience at a cost of a realistic treatment of uncertainty. With this particular assumption on preferences, marginal utility is linear in consumption. In this case, consumers are only interested in the certainty equivalence of future consumption – a direct consequence of the equality of the expected value of marginal utility of consumption, $\mathbb{E}(u'(c))$ with the marginal utility of expected consumption, $u'(\mathbb{E}(c))$. The precautionary saving model[9] relaxes this restrictive assumption by allowing for convexity of marginal utility, $u'''(\bullet) > 0$. With such preferences, the optimal consumption path will take an upward trend across time:

$$\mathbb{E}\left(c_{t+1}|\mathbb{I}^t\right) > c_t.$$

Many important results of the buffer stock model can be derived from the log-linearized approximation of the Euler equation[10]:

$$\mathbb{E}_t\left(\Delta \ln c_{t+1}\right) \simeq \alpha^{-1}\left(r-\beta\right) + \frac{1}{2}\alpha\mathbb{V}_t\left(\Delta \ln c_{t+1}\right) \tag{2}$$

This equation provides insights into the rich dynamics of the optimal consumption path. It is derived from the constant relative risk aversion utility function $u(c) = \frac{c^{1-\alpha}-1}{1-\alpha}$, where $\alpha > 0$ is the coefficient of relative prudence (and risk aversion), see, e.g. Deaton 1992 chapter 6. Here, the parameter β is the subjective discount rate. $\mathbb{V}_t$ denotes the variance at time t of the expected variability in the growth of consumption. Thus, as stressed by Deaton, any

[9] I use the term "buffer stock model" and "precautionary saving model" as short for the full name the buffer stock model of precautionary saving. For an early discussion on the relation between the precautionary motive and the third derivative of utility, see Leland (1968) and Sandmo (1970). For a general treatment, see e.g. Zeldes (1989), Kimball (1990) and Carroll (1997).

[10] An analytic solution of the consumption function is only known in one particular case, see Caballero (1990) and (1991).

variable which predicts future variability of income will also predict consumption growth. The expression $\frac{1}{2}\alpha\mathbb{V}_t\left(\Delta\ln c_{t+1}\right)$ is called the "precautionary saving" term.

An important contribution of the buffer stock model is an explanation of the empirical observation that consumption and income track each other over time. Assets are held as a buffer against income shocks and thus the consumption path differs from the life cycle model. In its simplest form, the model assumes that a prudent economic agent will refrain from borrowing. The liquidity constraints model developed by Deaton (1991) models the inability to borrow directly. One of the similarities to the buffer stock model is that it offers an additional motive for accumulating assets.

As outlined in the permanent income model above, the adjustment of consumption to a change in permanent income is instantaneous and permanent. This is predicted to occur regardless of whether the information shock has affected solely the mean or both the mean and the variance of the income process. In a consumption function derived from the life cycle model, the consumer is solely interested in the first moment of the income process. Changes to future uncertainty, while holding the expected value constant, do not lead to any reaction by the agent. In the context of the precautionary savings model there may, however, be a reversion to a higher consumption level following an initial drop if information has been revealed regarding the uncertainty of future income. This raises two questions:

a) is the shape of the dynamic path of consumption suggested by the precautionary saving model consistent with the predictions of the PIH-hypothesis, H^{PIH}?

b) is the buffer stock model consistent with the magnitude of the change in consumption observed in the data?

As described above, the general answer to question a) is no. Due to the precautionary saving term in equation (2), the consumer will cut back on his consumption following a shock. Once the consumer has acquired his optimal buffer stock, consumption may revert to a higher level. The answer to question b) depends on the magnitude of the drop in consumption and the subsequent speed of the reversion. A reversion to a higher consumption level after an initial drop is possible if information has been revealed regarding the uncertainty of future income. Consumers cut back on consumption in order to acquire a buffer stock. As the buffer stock reaches the optimum level, consumption rises again. Thus, the cumulative change in consumption equals the change in the optimal buffer stock.

The precautionary saving model was developed to account for the close tracking of consumption and income observed in the time series data. Hence, the time required to accumulate the buffer stock should occur over a reasonably long period of time. Deaton (1992) notes in chapter 6 that for simulations, the life-span of the accumulation of the precautionary balances is a couple of years.

In my data, I observe that following a shock to information, daily expenditure declines by about 2-4 percent and is reversed within 2-3 weeks. Such a rapid rate of convergence cannot plausibly be attributed to the precautionary saving motive alone. The buffer stock accumulated following such an abrupt cut would be too small to meet the needs following an increase in an anticipated volatility of future income. It appears that the precautionary saving motive cannot fully explain drops observed in the data. I will develop this line of reasoning further in the section on the results. In order to offer a place to reconcile data with theory, I now turn to a model incorporating incomplete information.

2.3 A Signal Extraction Model of Consumption

I contrast the permanent income model and the precautionary savings model to a case when consumption declines following a shock, but completely reverts to its former level at a fast pace of convergence. Such a quick and unstable movement of consumption would imply that consumers overreact in cutting back on spending when their perception of their economic prospects decline.

What might motivate such behavior? In the buffer stock model, consumers are not only interested in the certainty equivalence of future consumption, but also in the future variability of income. Following an increase in the uncertainty of future income, there may be a reversion of consumption to a higher level following a drop.

Wang (2004) studies a precautionary saving model with incomplete information. In this model, consumers face additional uncertainty about the components of the income process. Income consists of two components: a persistent component and a transitory part. The consumers cannot distinguish between the different components of income; instead they observe its current and past values. In addition, they receive a signal regarding the permanent component of income.

Facing this informational structure, the consumers form their guess about the hidden state – the persistent part of income – by recursively updating their previous guess using what they know about current observables: the total level of income and the signal. Formally, this updating is done by the means of a Kalman filter.

Wang shows that the precautionary saving motive is greater when the individual faces additional uncertainty in the shape of partially observed income. Due to incomplete information, consumers have an additional motive – additional to precautionary saving – for postponing spending.

The intuition for understanding this behavior is as follows. Suppose consumers receive information arrives about economic prospects. Due to the informational structure of this model, consumers receive incomplete news regarding poor economic prospects and it takes time to filter the true information component. This induces consumers to temporarily cut back their spending. As the consumers subsequently learn the meaning of the news, they adjust their consumption accordingly.

Similarly to the precautionary saving model, this model predicts that there can be a reversion to a higher consumption level following an initial drop. In addition to the demand for precautionary savings, consumers in this economy cut back on their spending due to an "estimation risk", i.e. additional uncertainty due to incomplete information about the income process.

The year 2008 was a period of great uncertainty regarding the economic environment. Consider Figure 1 in the Appendix. The plot shows the variations in Internet search activity for the term "recession". The series clearly illustrates upticks in search activity associated with news on the economy. One of the biggest spikes occurs on December 2nd, the day following the retrospective announcement by the National Bureau of Economic Research (NBER) on the U.S. recession. It seems likely from the plot that the variation in the interest in the phrase "recession" captures the relevant variation in the general public's concern about the economic downturn.[11]

If spending correlates with the time series of interest in "recession", it suggests that the searches for this term carry potentially meaningful information. Suppose that an announcement arrives about the recession on any particular day. The consumers perceive the announcement as a signal telling them that a drop in their persistent income has occurred. They do not know with certainty that this has happened – the signal is a sum of news about persistent income and noise. Following the announcement, the consumers will filter out the

[11]The Appendix includes a timeline of notable events related to reports on the economic slowdown during 2008.

information component from the noise component. As the they go forward with extracting the news from the signal, they may cut back on their spending.

Spending may react in the following ways. If the announcement of the recession carries news about the income process, spending should adjust and incorporate this update. If, however, consumption changes, but then is reversed, it would suggest, that the announcement was mostly noise. In the following section, there is an in-depth discussion on how I study the joint dynamics of searches for information and spending.

3 Empirical Methods

The three models outlined in the previous section present a stylized environment of uncertainty. At the beginning of section 2, I outlined a simple statistical model of information and consumption. Up to now I have not specified how I quantify information in the context of an empirical model.

My measure of information comes from the query logs of the Google search engine. My interpretation of Google search activity is the following. The index of Internet searches for the query "recession" is an indication that people have received (noisy) information about the recession. Following a shock to Google searches, consumption may react. By measuring the reaction of consumption, I can conclude whether the information contained in the announcement carried news about a fundamental change in the income process or whether it was mainly noise.

Using my measure of information, I study its joint variations with variation in daily spending throughout 2008. In my empirical analysis I use a vector autoregression to describe the near-real time dynamics of information and expenditure. By shocking the proxy for information, holding everything else constant, I study the path consumption takes. If consumption changes and remains at a new level, I take it to be in favor of a response to a change in fundamentals. This is analogous to H^{PIH} in the simple univariate setting described at the beginning of section 2. Conversely, if there is a change, but rapidly converging to the original level, it is in line with prudent behavior due to incomplete information. Such dynamics violate H^{PIH}.

In the latter part of this section, I utilize the micro data of the Gallup Daily Poll. The micro information contained in my data allows me to calculate consumption for different groups. The micro data enables me to study whether, following a shock, individuals facing high income instability cut back on their spending differently from those with stable incomes.

3.1 Time Series Methods

I model the joint dynamics of consumer expenditure and the proxy for information in a vector autoregressive model. I denote the proxy for information as s; s is a mnemonic for searches. I define $\mathbf{y}_t \equiv \begin{bmatrix} c_t & s_t \end{bmatrix}^T$. I estimate is a system of two variables regressed on their jointly lagged values up to lag p: (the notation here follows from chapter 2 in Lütkepohl (2005))

$$\mathbf{y}_t = \mathbf{A}_1\mathbf{y}_{t-1} + \mathbf{A}_2\mathbf{y}_{t-2} + ... + \mathbf{A}_p\mathbf{y}_{t-p} + \mathbf{u}_t,\ t = 1, ..., T. \tag{3}$$

where each $\mathbf{A}$ is a matrix of coefficients and $\mathbf{u}_t$ is a vector of disturbance terms, $\mathbf{u}_t \sim N(\mathbf{0}, \mathbf{\Sigma}_u)$ with $\mathbb{E}(\mathbf{u}_t\mathbf{u}_s') = 0$ for $s \neq t$.

Re-writing the system in lag notation yields $\mathbf{A}(L)\mathbf{y}_t = \mathbf{u}_t$. Further, defining $\mathbf{\Gamma}(L) \equiv \mathbf{A}(L)^{-1}$ and pre-multiplying this expression gives:

$$\mathbf{y}_t = \mathbf{\Gamma}(L)\,\mathbf{u} = \textstyle\sum_{i=0}^{p} \Gamma_i \mathbf{u}_{t-i} \tag{4}$$

The VAR process can thus be represented by a moving average of the residual times the accumulated responses Γ_i. This representation is useful since given the matrices of estimated coefficients it is possible to visualize the dynamics of the system by plotting out the predicted responses of the variables to a shock. Such an analysis is performed by means of an impulse response function, which plots out the effects of a one-unit increase in an element (say jth) of the $\mathbf{u}_t$ vector on the vector $\mathbf{y}$ for τ periods. In the two-variable case, let there be a one-time shock at $t = 1$, $\mathbf{u}_1 = \begin{bmatrix} 1 & 0 \end{bmatrix}^T$, while at all other times all the elements of $\mathbf{u}$ are zero: $\mathbf{u}_{t\neq 1} = \begin{bmatrix} 0 & 0 \end{bmatrix}^T$. The impulse response function is computed through successive iterations of matrices of estimates and the residuals.

We would like to attribute a causal interpretation to the effect of such a one-unit increase. However, the contemporaneous correlation of the elements of the empirical $\mathbf{u}_t$ precludes the usual analysis where "everything else is held constant". In order to ensure that the errors are contemporaneously uncorrelated, one can impose a recursive structure. This can be done by means of a Cholesky decomposition of the estimated contemporaneous covariance matrix, $\mathbb{E}(\mathbf{u}_t\mathbf{u}_t') = \hat{\mathbf{\Sigma}}_u$. This decomposition will make the error terms uncorrelated within the time dimension, see Lütkepohl (2005) for details.

The Cholesky decomposition of the covariance matrix creates new error terms, called innovations: $\boldsymbol{\varepsilon}_t = \mathbf{P}^{-1}\mathbf{u}_t$. The covariance matrix becomes an identity matrix and the magnitudes of the shocks are scaled to equal standard deviations. Instead of an plotting impulse response function with respect to an increase in $\mathbf{u}_t$, the orthogonalized impulse response function (OIRF) plots out the response of the $\mathbf{y}$ vector through repeated iterations of the estimated matrix and the innovation.[12]

Following a Cholesky decomposition, the first variable, say s_1, responds only to its own shock $\varepsilon_{s,t=1}$, the second variable, c_1, responds to the first variable's and to the second variable's exogenous shocks, etc. In the literature this is called specifying the Wold causal chain (see for example Stock and Watson (2001)) where $s_t \longrightarrow c_t$. This recursive ordering should be motivated by theory as there is no statistical reason for choosing one over the other. Sims (1980), who popularized the use of this method, suggests trying different combinations of Wold causal ordering to check how sensitive the results are. In addition, if the data is of very high frequency, one can expect there to be contemporaneous relationships between variables of interest. The ordering specified here will set searches first, consumption second.

4 Data

The data used in this paper come from two sources.

4.1 The Daily Poll

One part of the data comes from the Gallup Daily Poll conducted by the Gallup Organization. The data, called G1K, are collected *daily* via telephone interviews with a random sample of about 1,000 individuals aged 18 and older living in the United States. Each day a new cross-section is drawn. The survey is conducted seven days a week excluding major holidays.

[12] Sims (2009b) addresses the problem of non-invertibility of structural VARs. Non-invertibility refers to a situation where a "wedge" between the empirical innovations precludes the identification of the theoretical economic shocks (see also Fernández-Villaverde et al. (2006) and Fernández-Villaverde et al. (2007)). Sims (2009b) studies how a structural VAR will perform when compared to an underlying, calibrated RBC model and concludes that the VAR does not disappoint. Although the structural economic innovations cannot be recovered, in practice, the correlations between the innovation identified from the VAR and the model are quite high. This applies even to a simple structure such as a recursive VAR. As Sims (2009b) points out, this situation is likely to occur when studying the effects of information shocks.

The data collects information on about 359,000 individuals living in the United States surveyed across 355 days, from January 2nd 2008 to January 5th 2009. Gallup collects the data using a dual-frame random-digit-dial of both landlines and cellular phones. The interviews are conducted with the head of the household. In order to make the sample representative, Gallup provides survey sampling weights to correspond to the national distribution of age, gender, race, region and educational level. As the data are very recent, they have not yet been applied much in research. A paper by Krueger and Kuziemko (2009) uses the data to estimate the price elasticity of the demand for health insurance. The Daily Poll data were also used by Deaton and Arora (2009) in a study on the benefits of height.

The G1K data cover a variety of demographic measures, a set of questions on health and also evaluations of living and working conditions. G1K also collects economic information posed to a random half-sample of the respondents. A unique feature is that it collects high frequency information on daily expenditure. The expenditure question reads:

> Next, we'd like you to think about your spending yesterday, not counting the purchase of a home, motor vehicle, or your normal household bills. How much money did you spend or charge yesterday on all other types of purchases you may have made, such as at a store, restaurant, gas station, online, or elsewhere?

The answers measure the dollar amount spent on goods, services while excluding some of the biggest durables, such as the purchase of a home and car. The way the question is worded follows the recommendations set out by Browning et al. (2003).[13]

The G1K data is the biggest data set on daily expenditures. In comparison, another daily expenditures dataset, the Diary Survey part of the Consumer Expenditure Survey (CEX) leaves too few cells to be able to construct a time series. The G1K expenditure question was answered by about 185,000 individuals, which averages at about 500 people surveyed per day.

Many purchases made on a daily basis are storable. In the G1K data, on average, 30 percent of the sample report zero daily expenditures. This implies that even goods typically considered as non-durable, e.g. a half a gallon of milk, are not necessarily purchased for immediate consumption. Unlike the CEX, which collects data on specific subcategories of consumption, G1K do not allow me to distinguish between different types of purchases.

[13]When collecting retrospective information on total expenditure on non-durables and services one ought to specify an exact timeframe for the recall and a list of specific sub-items. Browning et al. (2003) argues that such a wording ought to pick up relevant variations in the answers.

G1K, however, allow me to split the sample into different categories of spenders. This division will allow for some suggestive indication whether necessities react differently from other types of purchases.

Another question collected for the same sub-sample as the consumption measure asks to evaluate the present economic conditions in America. The question reads: "Right now, do you think that economic conditions in this country, as a whole, are getting better or getting worse?". The answers are measured from "getting worse", "the same" to "getting better", which I re-code into $\{-1, 0, 1\}$.

4.2 Google Trends

Since it first became available for public use in the early 1990s, the Internet has grown rapidly in the United States. Today, North America has one of the highest personal computer densities (on average, one personal computer per person) in the world ("Getting wired" 2009). In 2001, the U.S. Census reported that over 41 percent of households had access to the Internet. By 2007, the percentage of Internet usage in and outside the home had grown to over 70 percent.[14]

Among the different search engines, Google is by far the most popular engine in the United States. It is difficult to estimate the exact market shares, although, according to an Internet measurement company, `hitwise`[15], `google.com` attracted nearly 70 percent of all searches in the United States in June 2008 (the closest runner-up, `search.yahoo.com`, attracted about 20 percent). As a majority of U.S. households have access to the Internet and since most of the users use Google it is plausible that the variation in the time series of different queries picks up the variation in the general public's demand for information.[16]

The data on internet searches come from a website provided by the owner of the Google search engine, Google Inc. Since May 2006, Google has provided an Internet service called Google Trends (`http://www.google.com/trends`, from hereon Trends). Trends make it possible to track some of the most popular searches and their frequencies by means

[14]http://www.census.gov/population/www/socdemo/computer/2007.html. Gallup asked the respondents of the Daily Poll on selected days whether they had Internet at home. 74 percent answered affirmatively.

[15]http://www.hitwise.com/us/press-center/press-releases/google-searches-oct-09

[16]More anecdotal evidence on the popularity of Google as a search engine comes from the adoption of the verb "to google" as being synonymous with general web searching. In order to avoid the watering down of its trademark, a clarification was posted on the Google blog, asking users to refrain from using the verb "to google" when they were in fact not using Google Inc. products of (Do you "Google?", 10/25/2006 Michael Krantz, Google Blog Team, http://googleblog.blogspot.com/2006/10/do-you-google.html).

of a time series (no specific number is mentioned, but on the Trends help page one can read that only search terms with "a significant amount of traffic" are included.[17])

When creating a Google account, a user can download the data on specific queries freely. The service allows the data to be sorted by time, geographical location and language based on the computer's IP address. The data is available both in a daily and weekly format, and the user can choose to use data normalized to a reference point in time, January 2004, making it comparable across time.

When working with aggregate query logs it is paramount to try to understand why these queries were made. As there is no contextual information on the motives of any particular user, one needs to infer what the purpose of a query might have been. Google provides several online tools aimed at helping to reduce this ambiguity. In August 2008, Google launched a companion site to Trends, called Google Insights for Search (from hereon Insights).[18] The aim of the Insights website is explicitly to help researchers and advertisers understand online search behavior. On the Insights website one can read that Insights uses an algorithm that "takes a look at broad search patterns" in order to determine what queries are related.[19] Unfortunately, neither this algorithm nor the procedure for the scaling of the data is provided. In addition to the information included in Trends, Insights provides a list of words associated with a specific query.[20]

These tools are meant to enable the user to refine a search for a specific item. Once the query specific selection criteria have been made, Insights produces a time series graph of the search volume. The user also has the option to see whether Insights interprets the spikes in interest as associated with particular news events. Trends also includes a time series plot of the number of times a given query was referenced in Google News.[21] Such plots (not available in data form) show how much search interest in a phrase covaries with reports including the same phrase.

[17] http://www.google.com/support/insights/bin/answer.py?hl=en&answer=87276

[18] http://www.google.com/insights/search/#

[19] http://www.google.com/support/insights/bin/answer.py?hl=en&answer=94792

[20] For example, the phrase "foreclosure" is related to queries on "foreclosure homes", "foreclosure sale", "foreclosure listings" etc. Insights also provides a list of items related to one's search called rising searches. These are searches that have experienced a "significant growth in a given time period, with respect to the preceding time period".

[21] Google News is a news aggregator provided by Google Inc. The aggregator references stories from many major online publications, such as the New York Times, Washington Post, Reuters and BBC News.

I concentrate on the web query for "recession" in the United States during 2008.[22] Insights suggests a list of queries associated with this item, namely: "US recession", "recession 2008", "depression recession" and "depression".[23] Turning to the Google News reference index plotted along the Google Trends series, I can infer whether the interest in the query "recession" is correlated with the number of times the word "recession" is mentioned in Google News stories. There appears to be a close correlation between the dates with increased reporting on the word "recession" and the increases in search interest.

4.3 Summary Statistics

Table 1 presents the summary statistics for the G1K variables used in the analysis. Note that this table presents the summary statistics of the unweighted sample.

One of the dependent variables in the regression analysis is consumption. Consumption has a fairly high mean value and is greatly dispersed. On average, the respondents report to spend about $100 on a daily basis. This is higher than the mean daily total expenditure, at $88, reported by Stephens (2003).[24] Stephens (2003) uses data from the daily diary component of the CEX in a paper on the recipients of Social Security checks. The consumption data collected by Gallup are not directly comparable to the measures in CEX. Gallup collects data on the expenditure of individuals as opposed to the consumer units in the CEX. As Gallup asks a recall question, it is also sensitive to measurement errors. In order to

[22] I also analyzed on the following items: "economy" and "crisis". Insights suggests a list of words associated with either one, namely: "recession", "financial crisis", "us economy" and "economic crisis". After relating the search volumes to each other, I limit the times series to "economy", "crisis" and "recession" as these seem to be searched for the most. I also analyze the series of "unemployment". The time series of terms that have the word unemployment as a component, e.g. "unemployment benefits" are dominated by the searches that are simply for "unemployment" and so I delimit the analysis to this latter item. When inspecting the searches for the phrases economy and crisis, it is clear that some of the variation in the series is cyclical and not driven by news alone. Also the context of the searches for "unemployment" is quite ambiguous.

[23] Insights also picks up the interest in a hiphop album entitled "The Recession" by Young Jeezy, released on September 2, 2008 by specifying the increase in phrases like "the recession", "jeezy recession", "jeezy", etc. The uptick in the index on September 2nd is likely to have been generated by some of the interest in the album, but the effect is ambiguous as there are also upticks on September 5th, the day of the release of US unemployment figures. Comparing the series for "recession" and "the recession" shows a clear spike in the latter series on September 2nd. Repeating the Insights search for 2009 lists searches plausibly associated with the current crisis: "2009 recession", "recession proof", "recession jobs", "economy recession", "economy" and "economic recession". To avoid further ambiguity, I do not include the search for depression as it is difficult to disentangle from interest in the illness.

[24] Stephens (2003) calculates the average total expenditure for a consumption unit across the years 1986–1996 to equal $62 in 1995 dollars. This amounts to about $88 in 2008 dollars using the CPI-U.

make the results comparable to a nationally representative sample, I take into account the sampling survey weights provided by Gallup in all of the regressions.

In order to allow for heterogeneity of the consumers, I use demographic data. To allow for different stages of the life cycle of the respondents I utilize the information regarding the age of the consumers. The Gallup data also include information on the number of children in the household and marital status. The sample has an average age of about 53. Relatively few households have children – the average number of children is about 0.6 while about 60 percent of the respondents are married. Table 1 shows that the sample is predominantly white. A high proportion of the sample, 40 percent, claims currently not to hold a job.[25]

Gallup collects information on total monthly household gross income in ten discrete bins. I transform this variable into a continuous measure by replacing the category code with the mid-point of each bin. According to Table 1 the average household income amounts to about \$5,000.[26] A high proportion, over 40 percent, reports to have completed their college education.

Table 2 presents the correlation coefficients between four internet queries and the time series for the survey question on economic conditions. Such correlations reveal whether interest in a phrase like "recession" is typically associated with high or low levels of consumer confidence. The correlations between the Internet searches are reasonably high; "crisis" displays a positive correlation coefficient of above 50 percent with all the queries except for "unemployment". "Recession" correlates highly with "economy", while "economy" correlates with "subprime".

What is noticeable is that the correlation between "unemployment" and the other queries is rather low. The survey question on consumer confidence displays a relatively low negative correlation with the queries for recession. Plotting a time series of the consumer confidence measure and the searches for "recession" for 2008, it appears that the two series tend to vary positively at the beginning of the year. After September, the correlation becomes negative. It suggests that the information contained in the query for "recession" is different from that of consumer confidence.

[25] Gallup has collected the information necessary to deduce whether an individual is unemployed according to the definition of the Bureau of Labor Statistics for a limited number of days of the sample.

[26] According to the U.S. Census, the median annual household income in 2007 was about \$50,000.

In order to create a time series, I collapse the data on expenditure into daily per capita averages, weighted by survey weights.[27] Attanasio and Weber (1993) discuss aggregation problems that arise when estimating Euler equations. When working with aggregate data on consumption, the information is only available in levels. These levels may be transformed into logarithms. The recommendation of Attanasio and Weber for the analysis of averaged micro data on expenditures is to first compute the non-linear transformation (here, the logarithm), then average the data. This approach is complicated with high frequency data. At the daily frequency there are bound to be days when someone does not make any expenditures. In the G1K data, on average, 30 percent of the sample report that they spent zero dollars. A logarithmic transformation of such a sparse variable would incur a huge loss of information. To my knowledge there are no guidelines for how to average such sparse micro-data, thus, I proceed with aggregating consumption to levels and transforming this aggregate.

Figure 1 plots the raw series of the logarithm of mean expenditure and the search volumes for the term recession. The left axis tracks the indices, while the right axis displays the logarithm of consumption. The query series clearly illustrates the upticks in search activity associated with news on the economy, see for example the huge spike on January 23rd, following the biggest cut of interest rates the Federal Reserve had made in 25 years. Other notable activity occurs around March 17th (the date the investment bank Bear Stearns was acquired by JP Morgan Chase). The public's interest in the query takes off after September 15th, when the investment bank Lehman Brothers filed for bankruptcy. A big spike in the recession series occurs on December 2nd. The day before, on December 1st, the National Bureau of Economic Research (NBER) declared the United States to be in a recession. A distinguishing feature of the year covering my observations, 2008, is that the reports about the recession were distinctively negative. This is illustrated in the Appendix, which includes a timeline of notable events related to reporting about the economic slowdown during 2008.

[27] The time series for income has very little time series variation. The coefficient of variation for this series in levels is about 0.037. The series for consumption in levels has a coefficient ten times as great (equal to about 0.31).

5 Results

In this section, I discuss the results from estimating the system of equations in (3). In order to dampen the variance of the time series, I transform all the variables into a logarithmic form. Further, I test for the presence of a unit root by the means of the Dickey-Fuller GLS (DF-GLS) test and the augmented Dickey-Fuller (ADF).

The DF-GLS for the unit root with a trend of the log of mean consumption shows that the null hypothesis of the unit root can be rejected up to lag 12 at a 10 percent significance level, see Table 3. The same test rejects the unit root without a trend at five lags at a 5 percent level of significance. Also, the ADF test for consumption (not shown in the Table) rejects the unit root without a drift or trend at very high lags, e.g. 14. I have also tested for the unit root using the Phillips-Perron test. This test uses Newey-West standard errors to account for serial correlation, whereas the ADF test uses additional lags of the first-differenced variable. I reject the null hypothesis of the unit root at various lags.

For the log of "recession" I reject the null hypothesis of the unit root with a trend at lag 4 at 1 percent significance level, see Table 4. The ADF test rejects the unit root with a drift at considerable lags, such as 10. When including a trend term, it rejects the unit root at lag 4. I include this variable both in levels and in first differences. This turns out to matter little to the results.

I perform a range of time-series diagnostics checking the right specification of the model. I select the number of lags with the help of standard information criteria (Akaike, Schwartz or Hannan-Quinn), see Table 5 for the information criteria test statistics. The top panel of Table 5 shows the lags picked for the simple system of log of consumption and "recession". The bottom panel of the same table presents the three information criteria after including time effects. In the post estimation phase, I check the model's forecasting power by performing Granger causality tests of the coefficients. I also check the stability conditions of the series. According to Proposition 2.1 in Lütkepohl (2005), the stability of a VAR process for $\mathbf{y}_t$ also implies stationarity for that series. All the responses are plotted following standard deviation shocks.

It is hard to interpret the coefficients estimated using a VAR. As discussed previously, I visualize the dynamics of this system by means of a plot of the impulse response functions. The orthogonalization I use allows for a contemporaneous effect of searches on consumption, $s_t \longrightarrow c_t$. The orthogonalization is important when there is substantial intratemporal correlation between the cross-equation residuals. The correlation between the residuals of the consumption and search regressions is negative and quite large, ranging from -0.03 to -0.09, depending on the specification. This implies that the OIRFs following another order will differ. At such high frequency as a day, there is little motivation for allowing consumption to intratemporally affect searches.[28]

Further, a closer scrutiny of Figure 1 reveals that both the consumption and search series have predictable weekend effects. In order to account for them, I include day of the week dummies in (3). In the specifications considered, the F-tests support the inclusion of exogenous effects such as day of the week dummies and a weekly time trend.

Figures 2a and 2b show the orthogonalized impulse response functions for two equations: consumption and the searches for "recession". The response of the log of consumption to a standard deviation shock to "recession" is plotted in the lower left-hand corner of the first figure. The magnitude of one standard deviation shock to "recession" increases "recession" by about 23 percent. Hence, such a 0.23 log point increase in "recession" decreases consumption by about 2 percent (0.02 decrease in the logarithm of consumption). This implies an elasticity of consumption to searches of about -0.1. In other words, a ten percent increase in information searches for "recession" causes a 1 percent drop in daily consumption.

Figure 2a shows that the decay of consumption following a shock to itself is quick, compared to that of "recession" after a shock to itself. Finally, the response of "recession" also reacts following a shock to consumption. I reject the null hypothesis that "recession" does not Granger-cause consumption at the1 percent level – the p-value is about 0.016. On the contrary, the consumption series does not help to predict the searches for "recession". Post-estimation tests of the stability of the VARs show that the series $\begin{bmatrix} c_t & s_t \end{bmatrix}^T$ are stable and subsequently stationary.

Consider the impulse response function of consumption following a shock to the interest in the phrase "recession" – Figure 2b plots this reaction. This is the same graph as the lower left-hand corner of the Figure 2a, but magnified. A standard deviation shock to "recession" results in a drop of about 2 percent in daily expenditure. This drop starts reversing after about 25 days. The shock to searches for recession appears not to have rendered any permanent effects on consumption.[29]

[28]Indeed, when ordering the VAR as $c_t \longrightarrow s_t$, the results change. Also in this specification, consumption drops. This drop is, however, smaller than when $s_t \longrightarrow c_t$.

[29]Including the income time series in a three variable VAR does not generate any new insights for the s and consumption series. This was expected as there is very little time series variation in the series of income: the OIRFs involving income are nearly entirely flat. This holds for different orderings of the VAR.

The signal extraction model predicts a temporary drop in consumption to occur due to an increase in the "estimation risk". In Figure 2b, the effect of the shock dies out within weeks. The drop in consumption only persists for a very short time. It is difficult to attribute this temporary drop solely to a precautionary saving motive. The buffer stock acquired would equal the cumulative change implied by the impulse response function. Given that mean daily consumption is about 100 dollars, a two percent drop, which lasts a few days, would suggest that, on average, the buffer stock equals about 50 dollars. It is, however, plausible that the temporary drop in spending is, in part, due to the noise carried in the news reporting on the recession. Going forward, individuals extract news from noise, and adjust spending accordingly.

In Figures 3a and b, I plot the same system, but extend the number of lags to 6. The impulse response function takes on a zigzag pattern, with the lowest drop of consumption at about 2.5 percent.[30] Again, the Granger causality tests indicate that "recession" does help to forecast the fluctuations in the level of spending. I can reject the null of no effect at 10 percent significance level. The converse is, however, not implied. The system also satisfies the stability condition.

The Granger causality tests indicate that the series with searches for "recession" helps to predict the series for consumption. The low forecasting power of consumption to the searches is not very surprising, as one could expect the feedback from the consumption series to searches to be rather low.

To address concern about the unit root process in the time series for recession, I re-estimate my analysis using the logarithm of the first differences of the Internet searches. Figures 4a and b present the OIRFs from this system. Again, the lower panel presents the magnified plot for consumption. Following a shock to the growth rate of internet searches for recession, consumption drops by about 5 percent. The Granger causality tests indicate that both series help to forecast one another.

In summary, the results point to a sensitivity of consumption to information shocks. This sensitivity is of a magnitude of a 1-5 percent drop in daily spending, depending on which specification is used. Translating this magnitude to elasticity implies that following a 10 percent increase in the searches for "recession", daily consumption drops by 0.5-2 percent. As these cutbacks only last a very limited period of time, they are unlikely to be

[30] Adding more than five lags to this bivariate VAR introduces zigzag, oscillatory behavior of the consumption response.

caused solely by transitory changes in the income process. The results speak in favor of a cutback due to a large noise shock regarding the income process.

5.1 Analysis by groups

The demographic data from the Gallup Daily Poll allow me to break up the analysis by groups. I estimate the VAR separately for different types of consumers and compare differences in consumption sensitivity following a shock to information. Individuals facing high income volatility may cut back on their spending more than those with stable incomes.

I concentrate on the jobless and the employed. I construct this first group using the Gallup question "Do you have a job?". Figure 5 plots the overlaid orthogonalized impulse response functions for the employed and the jobless. The information criteria favor lags of about six-eight for these groups. I plot the response function following a VAR of six lags and controlling for day of the week dummies and a weekly trend. Extending the VAR to eight lags turns out to produce very similar results. The OIRF of the jobless suggests a sharper decline in their daily spending – at most the drop is about 6 percent. As the time series of consumption for both groups is noisy, the confidence intervals are quite wide. The overlaid plot of the reaction of the two groups is suggestive of the differences in sensitivity of the two groups, but does not constitute a formal test. The Granger causality tests suggest that for both groups "recession" is likely to predict the reaction of consumption.

I have also considered other breakdowns. I compare the reaction of consumption of individuals above and below the typical retirement age of 65 years. The two groups consist of individuals aged 50-64 and those 65-85 years. Table 1 shows that the two groups are roughly equal. The Gallup Daily Poll asks the respondents whether they have access to health insurance. Also, a survey question is asked on whether the respondents worry about money. The composition of these groups is quite uneven, as over 80 percent claim to have a health insurance. A similar proportion claims not to worry about money. Additionally, it is hard to disentangle whether the lack of health insurance indicates liquidity constraints or signals something about the respondent's health. Similar problems arise for the survey question on worrying about money. When I estimated this system, the OIRFs did not indicate notable differences between the groups.

To reiterate, the analysis of the reaction of different groups suggests differences in how sensitive spending is to the arrival of information. The point estimate of the reaction path of the jobless is sharper than that of the employed. This reaction is less clear for different age categories and other groups. Economically more vulnerable individuals appear to be more sensitive to noise than other groups.

5.2 Analysis of different levels of expenditure

Following an increase in uncertainty, it is likely that consumers will cut down on expensive items, as opposed to necessities. Data from the National Economic Accounts shows that during 2008, the biggest drops in personal consumption were attributable to durables. Expenditure on household equipment, for example, dropped by 14 percent in the last quarter of 2008. Another subcategory of durables, the expenditures on recreational goods, witnessed a decrease of about 12 percent.[31]

I extend my analysis to include studying whether big expenditure reacts differently from the rest of the sample. I define big expenditure as those purchases above the 90th percentile of the overall expenditure distribution. This equals roughly 150 dollars or more spent on daily expenses. There are about 22,000 respondents in the sample who report such high expenses throughout the year. I also define a "low" consumption group as purchases below this threshold.

Figure 6 graphs the impulse response functions of different types of expenditures. Here, I include the time trend and six lags. The results are very similar when varying the number of lags. Not surprisingly, following an information shock, the largest drops in expenditure occur in the purchases of expensive items. This drop is over 4 percent. In comparison, there appears not to be a reaction for the remaining group. Comparing this with the result in Figure 3b, which plots the reaction of spending of the full sample, suggests that the drop in overall expenditures is driven by a drop at the higher end of the consumption distribution.

6 Conclusion

This paper studies the dynamics of information and consumption using high frequency data. I find that spending reacts rapidly to shocks to information, but in contrast to related research, this reaction fades within weeks. This rapid decay cannot be satisfactorily explained by canonical theories of consumption, such as the permanent income hypoth-

[31] See http://www.bea.gov/national/nipaweb/SelectTable.asp and choose Table “Percent Change from Preceding Period in Real Personal Consumption Expenditures by Major Type of Product” for 2007 and 2008.

esis model or the precautionary saving model. My results are, however, in line with the predictions of a signal extraction model of consumption.

In addition to casting doubt on the permanent income hypothesis, my results allow me to draw some broader implications. First, the data do not support the models of rational inattention. These models predict that individuals do not keep themselves constantly updated about economic information. This inattention manifests itself in a sluggish response of economic variables following shocks.[32] My results imply the contrary: individuals are aware of the information and choose to act upon it immediately. Once they separate the news component from the noise, they adjust their actions accordingly.

Second, viewing daily fluctuations in spending as deviations from the consumption smoothing path of the permanent income hypothesis implies a loss of utility. Cochrane (1989) computes utility losses associated with deviations from the consumption path suggested by the life cycle model. In a similar exercise, Lucas (1987) calculates the welfare costs of eliminating fluctuations in consumption by comparing utility from a deterministic and a stochastic consumption stream. Both papers conclude that small departures from the optimal consumption path (in Lucas' case, from the deterministic trend) of a representative agent incur minor losses in terms of utility.

Aggregate consumption is not a volatile variable. The quickly fading fluctuations in daily consumption I observe in my data ought to have a second order impact on aggregate wellbeing. However, this effect need not be entirely negligible for all consumers. The welfare calculations of Cochrane and Lucas assume a representative agent and hence do not address the utility losses of heterogeneous individuals. In an economy which allows for consumer heterogeneity, Krusell and Smith (1999) show that the poorest consumers may suffer utility losses following fluctuations to consumption.[33] The welfare losses for these individuals are found to be as high as 2 percent of average consumption. The overall welfare gain of eliminating aggregate fluctuations in the economy described in Krusell and Smith is 20 times greater than that of Lucas' original calculations. Their result is still a very small number. Nonetheless, this increase, following the addition of consumer heterogeneity to Lucas' model, is suggestive of welfare implications for the poorest.

My results offer venues for further research. Extending the analysis of consumer heterogeneity to further groups and time periods would be of great interest. Controlling for

[32] See, for example, Sims (2003) and Reis (2006).

[33] Krusell and Smith (1999) argue, however, that such individuals constitute a small fraction of the overall distribution of their calibrated economy.

information about consumption, occupation and labor force attachment of the spouse of the respondent would be one such extension. At present, the analysis is constrained only to individual, not household consumption. Further, my data do not allow me to distinguish between different components of expenditure. It is a fact that durables are the most volatile component of aggregate consumption (Attanasio (1999)). My findings point to the largest drops in total spending occurring at the upper percentiles of the distribution of expenditure. This suggests that fluctuations in consumption may reflect variations in the stock of large durables. With more consumption categories one could extend the analysis to include how specific categories of goods react to information shocks.

References

Askitas, N. and K. F. Zimmerman (2009). Google econometrics and unemployment forecasting. *Applied Economics Quarterly 55*(2), 107–122.

Attanasio, O. P. (1999). Consumption. In J. B. Taylor and M. Woodford (Eds.), *Handbook of Macroeconomics*, Volume 1, Chapter 11, pp. 741–812. Amsterdam: Elsevier.

Attanasio, O. P. and G. Weber (1993). Consumption Growth, the Interest Rate and Aggregation. *Review of Economic Studies 60*(3), 631–49.

Azar, J. (2009). Electric Cars and Oil Prices. Unpublished manuscript. Princeton University.

Azariadis, C. (1981). Self-fulfilling prophecies. *Journal of Economic Theory 25*(3), 380–396.

Barsky, R. B. and E. R. Sims (2009a). Information, Animal Spirits and the Meaning of Innovations in Consumer Confidence. NBER Working Paper 15049.

Barsky, R. B. and E. R. Sims (2009b). News Shocks. NBER Working Paper 15312.

Beaudry, P. and F. Portier (2004). An exploration into Pigou's theory of cycles. *Journal of Monetary Economics 51*(6), 1183–1216.

Beaudry, P. and F. Portier (2006). Stock Prices, News, and Economic Fluctuations. *American Economic Review 96*(4), 1293–1307.

Blanchard, O. J., J.-P. L'Huillier, and G. Lorenzoni (2009). News, Noise, and Fluctuations: An Empirical Exploration. NBER Working Paper 15015.

Browning, M., T. F. Crossley, and G. Weber (2003). Asking consumption questions in general purpose surveys. *Economic Journal 113*(491), 540–567.

Caballero, R. J. (1990). Consumption puzzles and precautionary savings. *Journal of Monetary Economics 25*(1), 113–136.

Caballero, R. J. (1991). Earnings Uncertainty and Aggregate Wealth Accumulation. *American Economic Review 81*(4), 859–871.

Carroll, C. D. (1997). Buffer Stock Saving and the Life Cycle/Permanent Income Hypothesis. *Quarterly Journal of Economics 112*(1), 1–55.

Choi, H. and H. Varian (2009a). Predicting Initial Claims for Unemployment Benefits. Google Inc.

Choi, H. and H. Varian (2009b). Predicting the Present with Google Trends. Google Inc.

Cochrane, J. H. (1989). The Sensitivity of Tests of the Intertemporal Allocation of Consumption to Near-Rational Alternatives. *American Economic Review 79*(3), 319–37.

Deaton, A. (1991). Saving and Liquidity Constraints. *Econometrica 59*(5), 1221–1248.

Deaton, A. (1992). *Understanding Consumption*. New York: Oxford University Press.

Deaton, A. and R. Arora (2009). Life at the top: the benefits of height. *Economics and Human Biology 7*(2), 133–136.

Fernández-Villaverde, J. and J. F. Rubio-Ramírez (2006). Economic and VAR Shocks: What Can Go Wrong? *Journal of the European Economic Association 4*(2-3), 466–474.

Fernández-Villaverde, J., J. F. Rubio-Ramírez, T. J. Sargent, and M. W. Watson (2007). ABCs (and Ds) of Understanding VARs. *American Economic Review 97*(3), 1021–1026.

Friedman, M. (1957). *A Theory of the Consumption Function*. Princeton: Princeton University Press.

"Getting wired" (2009, December 19). The Economist.

Ginsberg, J., M. H. Mohebbi, R. S. Patel, L. Brammer, M. S. Smolinski, and L. Brilliant (2009). Detecting influenza epidemics using search engine query data. *Nature 457*, 1012–1014.

Hall, R. E. (1978). Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence. *Journal of Political Economy 96*(6), 971–987.

Jaimovich, N. and S. Rebelo (2009). Can News About the Future Drive the Business Cycle? *American Economic Review 99*(4), 1097–1118.

Kimball, M. S. (1990). Precautionary Saving in the Small and in the Large. *Econometrica 58*(1), 53–73.

Krueger, A. B. and I. Kuziemko (2009, March 21). The Demand for Health Insurance among Uninsured Americans: Results of a Survey Experiment and Implications for Policy. Unpublished manuscript. Princeton University.

Krusell, P. and J. Anthony A. Smith (1999). On the Welfare Effects of Eliminating Business Cycles. *Review of Economic Dynamics 2*(1), 245–272.

Leland, H. E. (1968). Saving and Uncertainty: The Precautionary Demand for Saving. *Quarterly Journal of Economics 82*(3), 465–473.

Lorenzoni, G. (2009). A Theory of Demand Shocks. *American Economic Review 99*(5), 2050–84.

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. New York: Springer.

Lucas, Robert E., J. (1987). *Models of Business Cycles*. New York: Basil Blackwell.

Reis, R. (2006). Inattentive Consumers. *Journal of Monetary Economics 53*(8), 1761–1800.

Sandmo, A. (1970). The Effect of Uncertainty on Saving Decisions. *Review of Economic Studies 37*(3), 353–360.

Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica 48*(1), 1–48.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics 50*(3), 665–690.

Sims, E. R. (2009a). Expectations Driven Business Cycles: An Empirical Evaluation. Unpublished manuscript. University of Notre Dame.

Sims, E. R. (2009b). Non-Invertibilities and Structural VARs. Unpublished manuscript. University of Notre Dame.

Stephens, M. (2003). "3rd of tha Month": Do Social Security Recipients Smooth Consumption Between Checks? *American Economic Review 93*(1), 406–422.

Stock, J. H. and M. W. Watson (2001). Vector Autoregressions. *Journal of Economic Perspectives 15*(4), 101–115.

Wang, N. (2004). Precautionary saving and partially observed income. *Journal of Monetary Economics 51*(8), 1645–1681.

Zeldes, S. P. (1989). Optimal Consumption with Stochastic Income: Deviations from Certainty Equivalence. *Quarterly Journal of Economics 104*(2), 275–98.

A A Timeline of the Recession[34]

- 9 January 2008: The World Bank predicts a global economic slowdown in 2008.
- 21-22 January: Stock markets suffer big falls. The US Federal Reserve cuts rates by three quarters of a percentage point to 3.5% - its biggest cut in 25 years - to try and prevent the economy from slumping into recession..
- 31 January: A major bond insurer MBIA, announces a loss of $2.3bn - its biggest to date for a three-month period – blaming its exposure to the US sub-prime mortgage crisis.
- 7 February: US Federal Reserve chairman Ben Bernanke voices concerns about financial markets and the economy
- 17 March: Wall Street's fifth largest bank, Bear Stearns, is acquired by larger rival JP Morgan Chase for $240m in a deal backed by $30bn of central bank loans.
- 8 April: The International Monetary Fund (IMF) warns that potential losses from the credit crunch could reach $1 trillion.
- 5-7 September: The US labor market figures show the unemployment rate rising to 6.1%. Mortgage lenders Fannie Mae and Freddie Mac are rescued by the US government in one of the largest bailouts in US history.
- 15-16 September: Lehman Brothers files for Chapter 11 bankruptcy protection. Merrill Lynch agrees to be taken over by Bank of America for $50bn. The US Federal Reserve announces an $85bn rescue package for AIG.
- 25 September: In the largest bank failure yet in the United States, Washington Mutual is closed down by regulators and sold to JPMorgan Chase.
- 28-29 September: The lawmakers announce they have reached an agreement on a rescue plan for the American financial system. The US House of Representatives rejects a $700bn rescue plan for the US financial system - sending shock waves

[34]The information is taken from the BBC Recession Timeline found at: http://news.bbc.co.uk/2/hi/business/7521250.stm and lists events that correspond to increases in the search activity for the query "recession".

around the world. Wall Street shares plunge, with the Dow Jones index slumping 7%, a record one-day point fall.

- 3 October: The US House of Representatives passes a $700bn government plan to rescue the US financial sector.
- 8 October: The US Federal Reserve, European Central Bank (ECB), Bank of England, and the central banks of Canada, Sweden and Switzerland make emergency interest rate cuts of half a percentage point. The Fed cuts its base lending rate to 1.5%, the ECB to 3.75%, and the Bank of England to 4.5%.
- 11 October: Finance ministers from leading industrialized nations pledge action to tackle the financial crisis.
- 15 October: Figures for US retail sales in September show a fall of 1.2%, the biggest monthly decline in more than three years. The figures underscore fears that the wider US economy is now being hit by the financial crisis. The Dow Jones index falls 7.87% - its biggest percentage fall since 26 October 1987.
- 24 October: The UK is on the brink of a recession.
- 30 October: The Federal Reserve cuts its key interest rate from 1.5% to 1%.The Commerce Department issues figures showing the US economy shrank at an annualized rate of 0.3% between July and September.
- 14 November: Leaders of the G20 developed and emerging economies gather in Washington to discuss ways to contain the financial crisis and agree on longer-term reforms.
- 1 December: The US recession is officially declared by the National Bureau of Economic Research (NBER). NBER concludes that the US economy started to contract in December 2007.

B Results

Table 1: Sample Summary Statistics

Variable	Mean	Sd	Min	Median	Max	N
Consumption	103.41	513.84	0	20	10,000	185,629
Age	53.54	16.85	18	54	94	183,559
Age $\in [50, 65)$	0.32	0.47	0	0	1	183,559
Age $\in [65, 85]$	0.26	0.44	0	0	1	183,559
No. of Children < 18 years	0.59	1.07	0	0	15	185,379
Married	0.57	0.5	0	1	1	184,590
Woman	0.49	0.5	0	0	1	185,628
White	0.84	0.37	0	1	1	183,077
Black	0.07	0.25	0	0	1	183,077
Hispanic	0.03	0.16	0	0	1	183,077
Has a Job?	0.6	0.49	0	1	1	185,331
Has a College Degree	0.42	0.49	0	0	1	184,327
Monthly Gross Income	4,943.38	3,126.25	0	4,499.5	10,000	143,076
Economic Conditions	-0.71	0.67	-1	-1	1	182,301

Source: Gallup Daily Poll

Table 2: Correlation Matrix between Internet Searches and Survey Question on Economic Conditions

Variables						
Searches for:	Unemployment	Recession	Economy	Subprime	Crisis	Economic Conditions
Unemployment	1.000					
Recession	0.398	1.000				
Economy	0.280	0.6061	1.000			
Subprime	-0.014	0.430	0.626	1.000		
Crisis	0.287	0.517	0.877	0.637	1.000	
Economic Conditions	-0.201	-0.135	0.316	0.205	0.335	1.000

Correlations between Trends Searches and Gallup's question on economic conditions.
The condition question is collapsed to a daily time series weighted by survey weights.

Table 3: Unit Root Tests for the Log of Consumption

Lags	**DF-GLS τ Test Statistic***	**1% Critical Value**	**5% Critical Value**	**10% Critical Value**
12	**-2.886**	-3.480	-2.890	**-2.570**
11	**-3.314**	-3.480	**-2.890**	-2.570
10	-3.461	-3.480	-2.890	-2.570
9	-3.539	-3.480	-2.890	-2.570
8	-3.619	-3.480	-2.890	-2.570
7	-3.474	-3.480	-2.890	-2.570
6	**-3.878**	**-3.480**	-2.890	-2.570
5	-4.782	-3.480	-2.890	-2.570
4	-6.125	-3.480	-2.890	-2.570
3	-6.910	-3.480	-2.890	-2.570
2	-8.234	-3.480	-2.890	-2.570
1	-10.273	-3.480	-2.890	-2.570

*This specification assumes a simple random walk.

Lags	**DF-GLS μ Test Statistic♣**	**1% Critical Value**	**5% Critical Value**	**10% Critical Value**
6	**-1.643**	-2.580	-1.950	**-1.620**
5	**-2.276**	-2.580	**-1.950**	-1.620
4	**-3.166**	**-2.580**	-1.950	-1.620
3	-3.746	-2.580	-1.950	-1.620
2	-4.732	-2.580	-1.950	-1.620
1	-6.247	-2.580	-1.950	-1.620

♣This specification allows for a non-zero mean, no linear-trend random walk.

Table 4: Unit Root Tests for the Log of "Recession"

Lags	DF-GLS τ Test Statistic*	1% Critical Value	5% Critical Value	10% Critical Value
11	**-3.107**	-3.480	**-2.890**	-2.570
10	-3.239	-3.480	-2.890	-2.570
9	-3.316	-3.480	-2.890	-2.570
8	-3.288	-3.480	-2.890	-2.570
7	-3.258	-3.480	-2.890	-2.570
6	-3.007	-3.480	-2.890	-2.570
5	-3.113	-3.480	-2.890	-2.570
4	**-3.581**	**-3.480**	-2.890	-2.570
3	-3.952	-3.480	-2.890	-2.570
2	-4.283	-3.480	-2.890	-2.570
1	-5.403	-3.480	-2.890	-2.570

*This specification assumes a simple random walk.

Lags	DF-GLS μ Test Statistic♣	1% Critical Value	5% Critical Value	10% Critical Value
8	-1.983	-2.580	-1.950	-1.620
7	-1.944	-2.580	-1.950	-1.620
6	-1.735	-2.580	-1.950	-1.620
5	-1.821	-2.580	-1.950	-1.620
4	**-2.188**	-2.580	**-1.950**	-1.620
3	-2.467	-2.580	-1.950	-1.620
2	**-2.742**	**-2.580**	-1.950	-1.620
1	-3.656	-2.580	-1.950	-1.620

♣This specification allows for a non-zero mean, no linear-trend random walk.

Table 5: Lag Length Selection Information Criteria for the log of Consumption and log of "Recession"

Lags	AIC*	HQIC*	SBIC*
0	-3.71578	-3.71578	-3.71578
1	-5.25289	-5.23534	**-5.2088***
2	-5.24171	-5.20662	-5.15353
3	-5.28988	-5.23723	-5.15761
4	-5.30004	-5.22985	-5.12368
5	-5.30625	-5.21851	-5.0858
6	-5.38218	**-5.27688***	-5.11763
7	-5.39759	-5.27474	-5.08895
8	-5.41718	-5.27679	-5.06446
9	-5.40364	-5.24569	-5.00682
10	-5.40079	-5.22529	-4.95988
11	-5.39087	-5.19782	-4.90587
12	-5.39584	-5.18524	-4.86675
13	-5.4224	-5.19426	-4.84922
14	**-5.42941***	-5.18371	-4.81214

*Simple specification (no exogenous variables)

Lags	AIC♣	HQIC♣	SBIC♣
0	-2.4511	-2.4511	-2.4511
1	-4.0813	**-4.0637***	**-4.0372***
2	-4.0768	-4.0417	-3.9887
3	**-4.1126***	-4.0599	-3.9803

♣Specification with linear weekly trend,
and day of the week dummies.

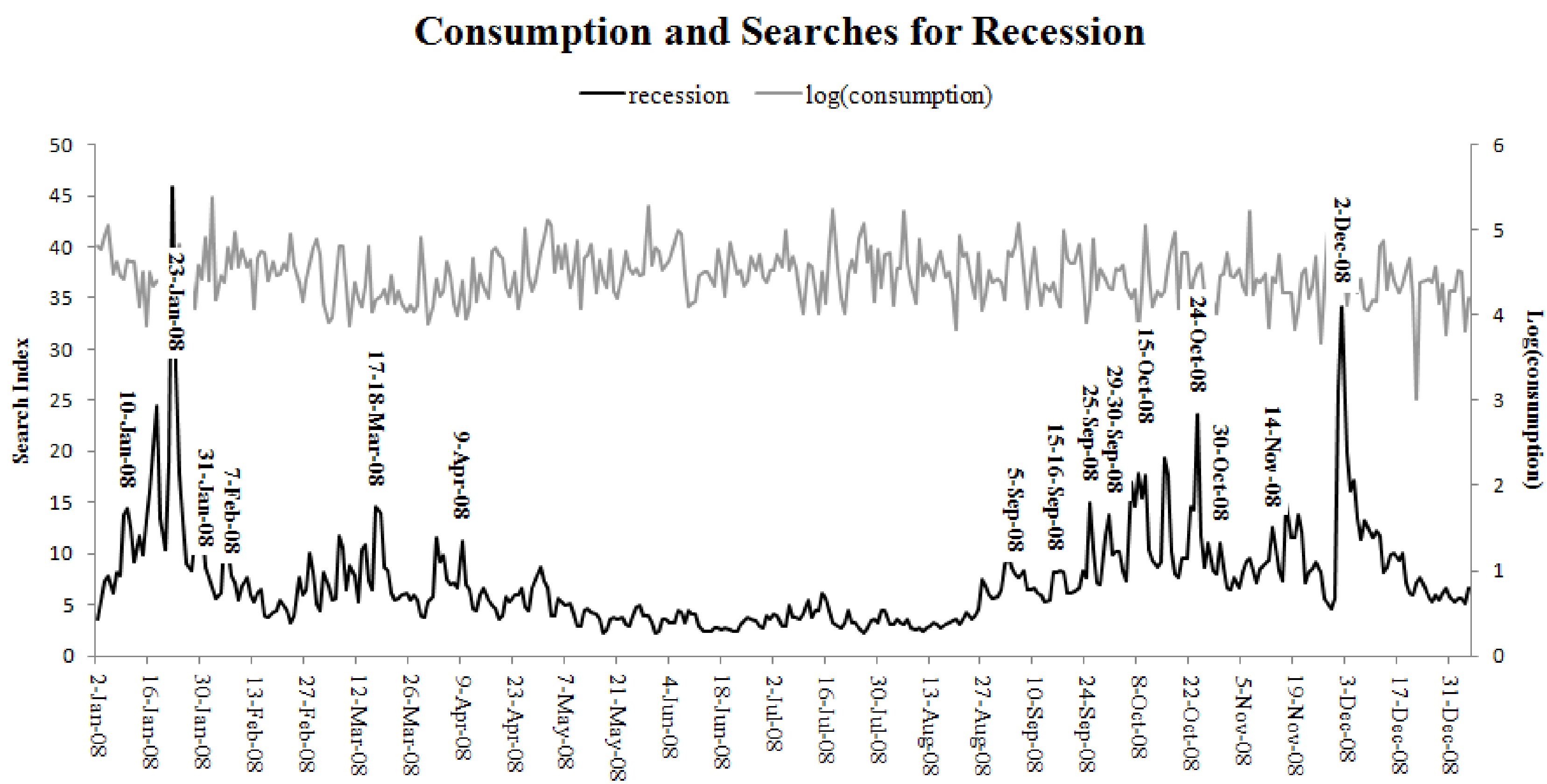

Figure 1: Plot of the Logarithm of Consumption and Searches for "Recession" for 2008. Major events of the year 2008 are marked in the plot. See the time line in the Appendix for descriptions of the events.

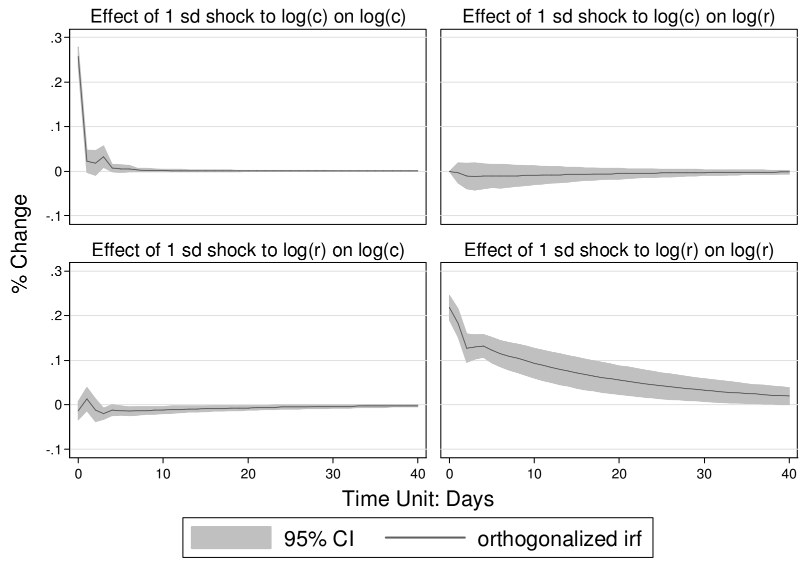

Figure 2a: Orthogonalized Impulse Response Functions. System with two variables: Logarithm of Consumption and Logarithm of Recession. Exogenous variables included: day of the week dummies and a weekly trend. Confidence Intervals are computed using bootstrapping method with 200 draws. Lags: 3. Recursive restriction: Recession, Consumption.

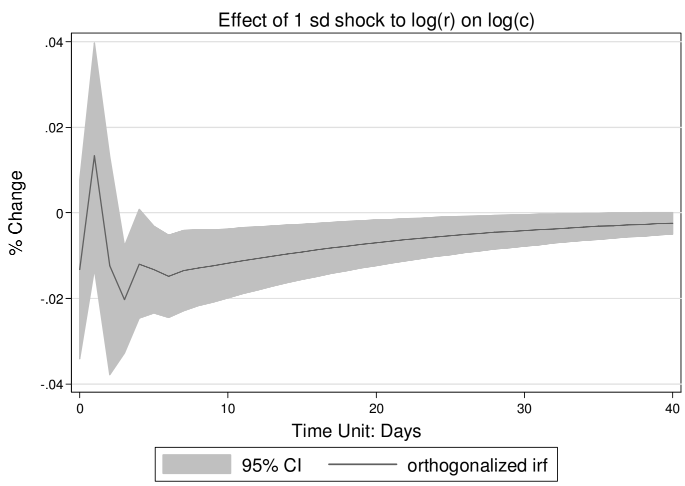

Figure 2b: Orthogonalized Impulse Response Function of Consumption following a one unit shock to the searches for Recession. System with two variables: Logarithm of Consumption and Logarithm of Recession. Exogenous variables included: day of the week dummies and a weekly trend. Confidence Intervals are computed using bootstrapping method with 200 draws. Lags: 3. Recursive restriction: Recession, Consumption.

Granger-tests: p-values

H_0: Consumption does not Granger-cause Searches for Recession: 0.884

H_0: Searches for Recession does not Granger-cause Consumption: 0.016

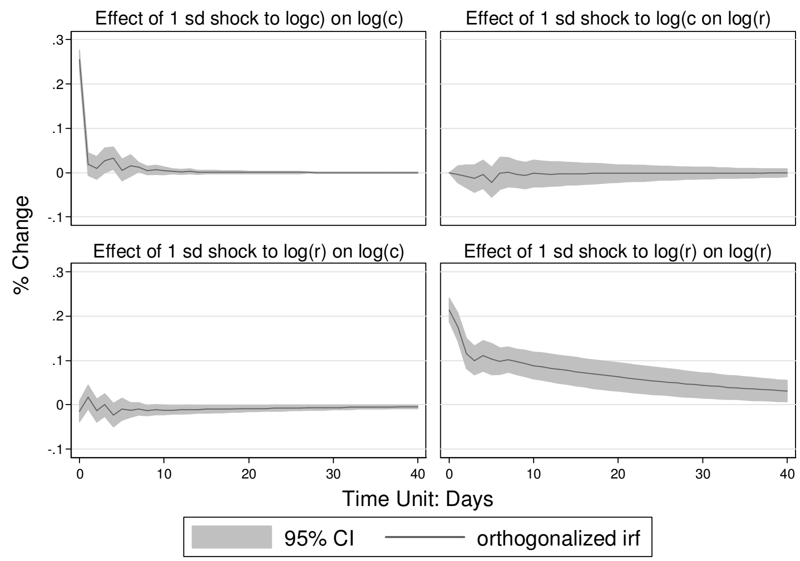

Figure 3a: Orthogonalized Impulse Response Functions. System with two variables: Logarithm of Consumption and Logarithm of Recession. Exogenous variables included day of the week dummies and a weekly trend. Confidence Intervals are computed using bootstrapping method with 200 draws. Lags: 6. Recursive restriction: Recession, Consumption.

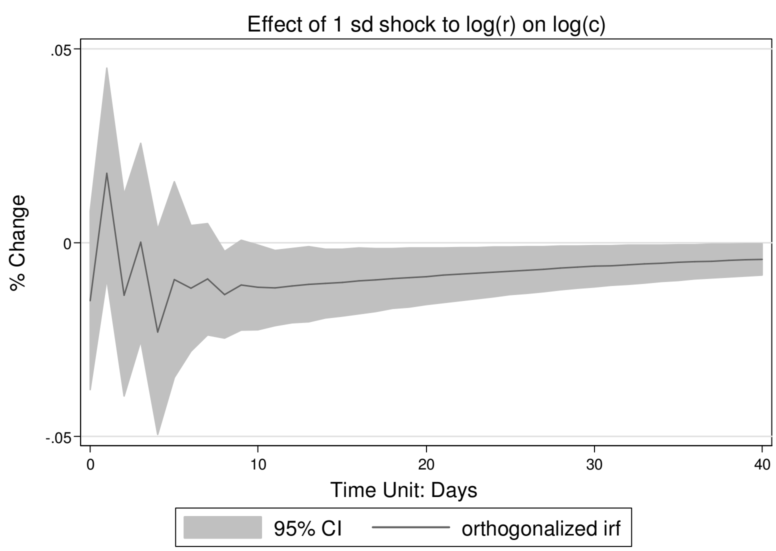

Figure 3b: Orthogonalized Impulse Response Function of Consumption following a one unit shock to the searches for Recession. System with two variables: Logarithm of Consumption and Logarithm of Recession. Exogenous variables included: day of the week dummies and a weekly trend. Confidence Intervals are computed using bootstrapping method with 200 draws. Lags: 6. Recursive restriction: Recession, Consumption.

Granger-tests: p-values
H_0: Consumption does not Granger-cause Searches for Recession: 0.361
H_0: Searches for Recession does not Granger-cause Consumption: 0.072

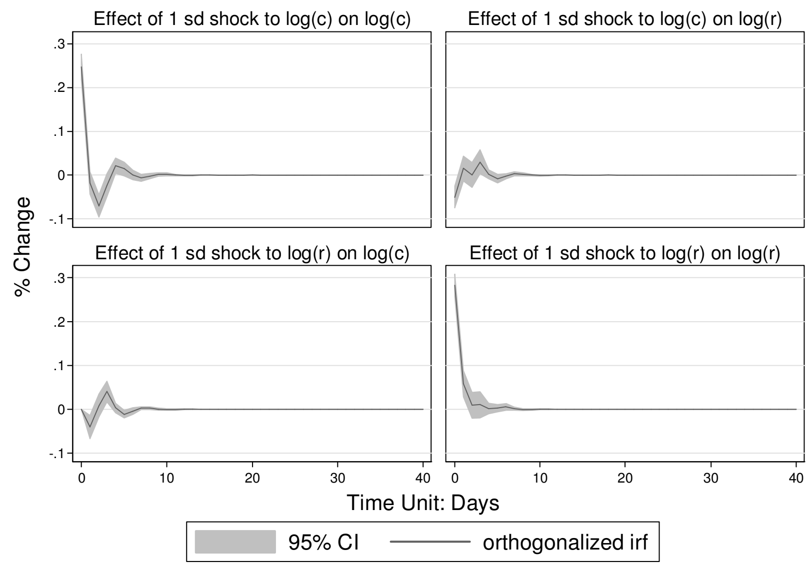

Figure 4a: Orthogonalized Impulse Response Functions. System with two variables: Logarithm of Consumption and Logarithm of first difference of Recession. Confidence Intervals are computed using bootstrapping method with 200 draws. Lags: 3. Recursive restriction: Recession, Consumption.

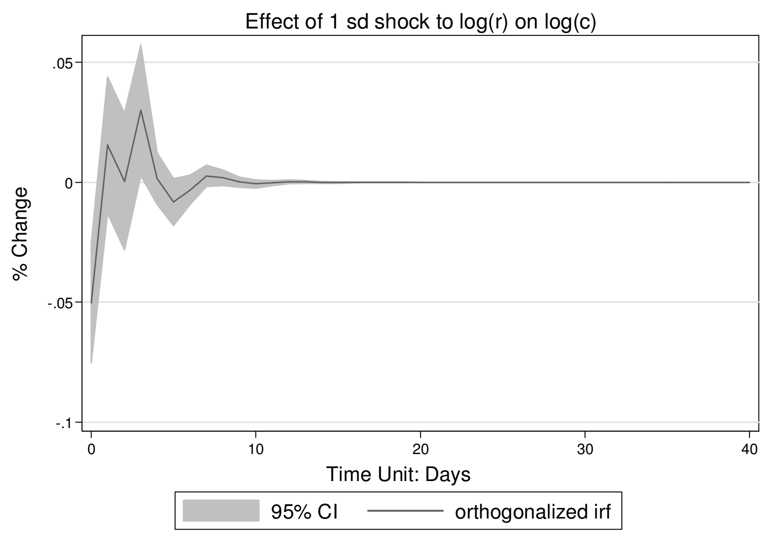

Figure 4b: Orthogonalized Impulse Response Function of Consumption following a one unit shock to the searches for Recession. System with two variables: Logarithm of Consumption and Logarithm of the fist difference of Recession. Confidence Intervals are computed using bootstrapping method with 200 draws. Lags: 3. Recursive restriction: Recession, Consumption.

Granger-tests: p-values
H_0: Consumption does not Granger-cause Searches for Recession: 0.002
H_0: Searches for Recession does not Granger-cause Consumption: 0.048

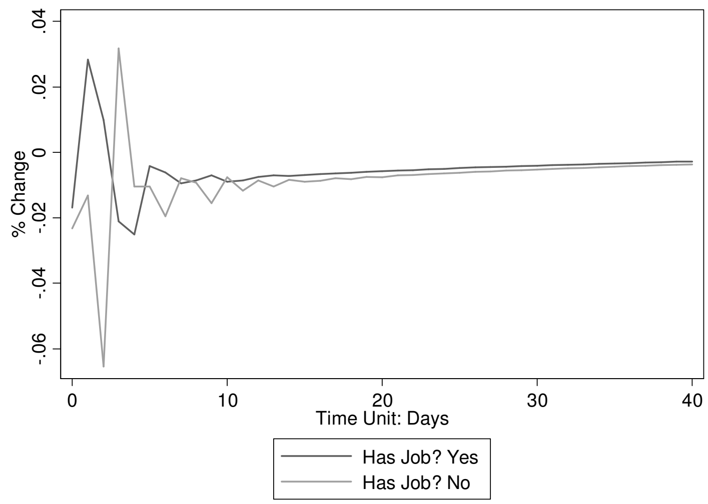

Figure 5: Overlaid Orthogonalized Impulse Response Function of Consumption following a one unit shock to the searches for Recession. The graph shows two response functions: one group consists of the employed. The other group consists of the jobless. The system has two variables: Logarithm of Consumption and Logarithm of Recession. Exogenous variables included: day of the week dummies and a weekly trend. Lags: 6. Recursive restriction: Recession, Consumption.

Has Job? Yes
Granger-tests: p-values
H_0: Consumption does not Granger-cause Searches for Recession 0.801
H_0: Searches for Recession does not Granger-cause Consumption 0.057

Has Job? No
Granger-tests: p-values
H_0: Consumption does not Granger-cause Searches for Recession 0.003
H_0: Searches for Recession does not Granger-cause Consumption 0.006

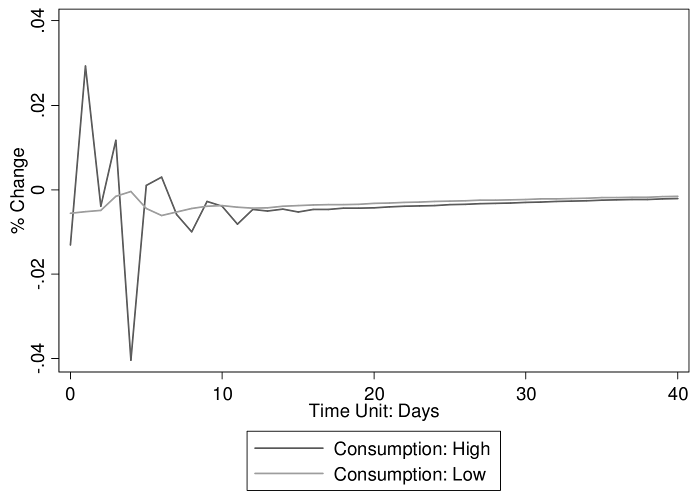

Figure 6: Overlaid Orthogonalized Impulse Response Function of Consumption following a one unit shock to the searches for Recession. The graph shows response functions of two groups. One group, "Consumption: High", is created from those individuals who report consumption above 90th percentile of the distribution. "Consumption: Low" consists of those who report consumption below this threshold. The system has two variables: Logarithm of Consumption and Logarithm of Recession. Exogenous variables included: day of the week dummies and a weekly trend. Lags: 6. Recursive restriction: Recession, Consumption.

Consumption: High
Granger-tests: p-values
H_0: Consumption does not Granger-cause Searches for Recession: 0.186
H_0: Searches for Recession does not Granger-cause Consumption: 0.023

Consumption: Low
Granger-tests: p-values
H_0: Consumption does not Granger-cause Searches for Recession: 0.307
H_0: Searches for Recession does not Granger-cause Consumption: 0.019