Does consumer confidence affect durable goods spending during bad and good economic times equally?*

M Iqbal Ahmed†
Kansas State University

Steven P. Cassou‡
Kansas State University

April 5, 2016

Abstract

Recent econometric analysis shows that consumer confidence innovations have long lasting effects on economic activities like consumption. Using US data, we show this conclusion is more nuanced when considering an economy that has different potential states. In particular, using regime-switching models, we show that the connection between consumer confidence to some types of consumer purchases is important during good economic times, but is relatively unimportant during bad economic times. Our regime-switching models use the unemployment rate as the indicator distinguishing bad and good economic times and investigate impulse responses, Granger Causality and variance decompositions. We consistently find that the impact of consumer confidence is dependent on the state of the economy for durable goods purchases. We also use this type of model to investigate the connection between news and consumer confidence and this connection is also state dependent. These findings have important implications for recent policy debates which consider whether confidence boosting policies, like raising inflation expectations on big-ticket items such as automobiles or business equipment would lead to a faster recovery.

JEL Classification: E21,E32, D12, C32
Keywords: Consumer confidence, durable goods, local projections, regime switching models

*We would like to thank Oscar Jordà for advice as well as making available basic computer programs which we modified for our application.

† Department of Economics, 327 Waters Hall, Kansas State University, Manhattan, KS, 66506 (USA), (785) 313-0668, Fax:(785) 532-6919, email: iahmed@ksu.edu.

‡Corresponding author. Department of Economics, 327 Waters Hall, Kansas State University, Manhattan, KS, 66506 (USA), (785) 532-6342, Fax:(785) 532-6919, email: scassou@k-state.edu.

"It is unfortunate that most economists and business writers apparently do not seem to appreciate this [role of animal spirits] and thus often fall back on the most tortured and artificial interpretations of economic events. They assume that variations in individual feelings, impressions, and passions do not matter in the aggregate and that economic events are driven by inscrutable technical factors or erratic government action." - George A. Akerlof and Robert J. Shiller[1]

1 Introduction

In the wake of the Great Recession of 2008-09, there have been numerous calls that suggest confidence-boosting policies could help speed up the recovery.[2] A recent empirical paper by Barsky and Sims (2012) has reinforced these policy calls by showing that the Michigan Consumer Confidence Index contains important information about "news" on future productivity that has long lasting effects on economic activities like aggregate consumption.[3] In this paper, we investigate the robustness of the Barsky and Sims (2012) results by asking whether there could be differences in the connection between consumer confidence and consumer consumption during bad and good economic times. To explore this issue, we use regime switching models that decompose the connections between confidence and consumption into a part corresponding to the response during bad economic times and a part corresponding to the response during good economic times. We also decompose consumption into its subcomponents to see if different parts of consumption respond differently. Our regime-switching

[1] "Animal spirits: How human psychology drives the economy, and why it matters for global capitalism". 2010. *Princeton University Press.*

[2] "But the hope that monetary and fiscal policies would prevent continued weakness by boosting consumer confidence was derailed by the recent report that consumer confidence in January collapsed to the lowest level since 1992."- Martin Feldstein, *The Wall Street Journal,* February, 20, 2008.

"Economic activity in the United States turned up in the second half of 2009, supported by an improvement in financial conditions, stimulus from monetary and fiscal policies, and a recovery in foreign economies. These factors, along with increased business and household confidence, appear likely to boost spending and sustain the economic expansion." - Ben Bernanke, *Monetary Policy Report to Congress, February 24, 2010.*

[3] Other relevant academic papers include Blanchard (1993), Carroll, Fuhrer and Wilcox (1994) and Ludvigson (2004) who argue that one of the leading causes of the 1990-92 recession was weak household and business confidence. In addition, Bachmann and Sims (2012), and Petev, Pistaferri and Saporta (2012) suggest the slow recovery since the Great Recession of 2008-09 is in part due to weak confidence.

models use the unemployment rate to distinguish bad and good economic times, and we investigate impulse responses, Granger Causality and variance decompositions to address this issue. We consistently find that the impact of consumer confidence on durable goods consumption is dependent on the state of the economy and that the response during bad economic times is small. This may help explain the relatively weak economic recovery since the Great Recession of 2008-09, despite the improvements in consumer confidence, and can be interpreted as supporting the alternative view that is skeptical of whether confidence boosting policies would help the recovery.[4]

The fact that consumers respond differently during bad economic times is not new to the economic literature, with important contributions arising in the investment under uncertainty literature.[5] But so far there has been no work which has used state of the art regime-switching models to investigate whether the link between consumer confidence and consumer spending is the same during different states of the economy. In this paper we use linear projection methods suggested by Jordà (2005) to investigate impulse responses and variance decompositions and more standard switching models to investigate Granger Causality to see if the connections between consumer confidence and consumer spending are equal across the business cycle. The linear projection methods are well suited for the complicated switching structures used in this paper which could not be carried out using standard Vector Autoregression (VAR) methods.[6] However, one complication associated with the linear projection method

[4] "Yale's Bob Shiller argues that confidence is the key to getting the economy back on track. I think a lot of economists would agree with that. [...]. The sad truth is that we economists don't know very much about what drives the animal spirits of economic participants. Until we figure it out, it is best to be suspicious of any policy whose benefits are supposed to work through the amorphous channel of confidence." – N. Gregory Mankiw, Blog, January 27, 2009.

"The stimulus was too small, and it will fade out next year, while high unemployment is undermining both consumer and business confidence."- *The New York Times, The opinion pages, July 28, 2010.*

"Others say that we should have a fiscal stimulus to 'give people confidence,' even if we have neither theory nor evidence that it will work." - *John Cochrane (2009).*

[5] Papers by Bernanke (1985) and Berger and Vavra (2015) show that few households purchase durable goods during recessions. In addition, papers by Katona(1968), Mishkin (1976), Blanchard (1993), Carroll, Furher and Wilcox (1994), Cochrane (1994), Acemoglu and Scott (1994), Matsusaka and Sbordone (1995,) Batchelor and Dua (1998), Howrey (2001), Ludvigson (2004), Berry and Davey (2004) and Starr (2012) have also investigated the connection between confidenc and consumption.

[6] This approach also has several advantages over other new methods such as smooth transition vector autoregressive (STVAR) models used by Auerbach and Gorodnichenko (2013, 2015) which

is serial correlation in the error terms induced by the successive leading of the dependent variable. We address this issue by using methods suggested in Newey and West (1987) to compute standard errors.

Using impulse response, Granger Causality and variance decomposition methods, we find that different subcomponents of consumption respond to consumer confidence differently over the business cycle. In particular, while consumption of nondurable goods behave in ways that are symmetric over the business cycle, consumption of durable goods are state dependent. We find that the connection between consumer confidence and durable goods consumption is quite strong during good economic times, but is small during bad economic times. We further explore the robustness of this result by trying alternative measures of consumer confidence, various subcomponents of the durable goods data, an alternative structural shock identification method (i.e. an alternative Cholesky ordering) and alternative lag structures for our models and find the results continue to hold up. Finally, we explore the "news" origins of consumer confidence by extending the approach used in Barsky and Sims (2012) to include switching structures. Again we find that the connection between news and consumer confidence is state dependent.

The rest of the paper is organized as follows. In Section 2, we spell out the details of the econometric methods used in this paper. Section 3 then discusses the results of the empirical analysis. In Section 4, we describe several alternative exercises we undertook to investigate the robustness of our results. Next, Section 5 explores the connection between news shocks and consumer confidence and finally Section 6 concludes the paper.

---

also incorporate nonlinear features, as the linear projection methods provides greater flexibility in terms of estimation. See, for instance, Stock and Watson (2007) and Owyang, Ramey and Zubairy (2013) and Ramey and Zubairy (2014) for addition discussion on this topic.

2 Econometric method

This section describes the econometric methods used in this paper. For clarity, it is broken into two subsections, with the first describing the linear and threshold empirical models as well as a general approach for finding the impulse response functions. The second subsection gets more specific about the particular methods that were used for identifying the structural shocks in the impulse response function (IRF) calculations.

2.1 Empirical models

All models use four basic variables including one of several measures for consumer confidence, one of several measures for consumption, a measure of income and one of several measures of financial assets, which we will denote generically by cc_t, c_t, y_t and f_t respectively. Here we use mostly standard notations, with the exception of cc_t, which indicates consumer confidence. Inclusion of these variables is motivated by Lettau and Ludvigson (2001, 2004) and Carroll, Fuhrer and Wilcox (1994), who built on a model described in Campbell and Mankiw (1989,1990,1991). The Campbell and Mankiw model includes two types of consumers, one which follows a dynamic consumer optimizing structure and another which follows a rule of thumb. Carroll et el. (1994) extend this model to include consumer confidence, showing that in an economy in which some consumers are not life-cycle optimizers, consumer sentiment will forecast a household's spending on durable and nondurable goods. Their empirical model controls only for household labor income. Inclusion of financial assets can be motivated by Ludvigson (2005) and Leeper (1992).[7]

For now, we will focus on a simple linear model in which there is no threshold behavior, which we will regard as the current frontier for the literature in this area. Because our objective is to show differences in the IRF once thresholds are added, this will be a useful baseline for comparisons. To generate the IRFs we make use

[7] Leeper (1992) finds that consumer sentiment are weekly correlated with other economic variables such as unemployment and industrial production once financial indicators are included in the regression model.

of methods suggested by Jordà (2005), which have the advantage over the more common vector autoregression (VAR) methods, because they only require projecting one period at a time, rather than an increasingly distant horizon as in the VAR methods. This method generates IRFs by running a sequence of forecast equations given by

$$x_{t+s} = \alpha^s + \sum_{i=1}^{p} B_i^{s+1} x_{t-i} + u_{t+s}^s \qquad s = 0, 1, ..., h, \tag{1}$$

where $x_t = [cc_t \quad c_t \quad y_t \quad f_t]'$ is a vector of the model variables which we wish to forecast s steps ahead for h different forecast horizons using a forecasting model consisting of only p lags of the variables in the system. The parameters in the model are straight forward, with α^s denoting a 4×1 vector of constants and B_i^{s+1} denoting 4×4 square matrices of parameters corresponding to the ith lag, x_{t-i}, in the s step ahead forecasting model and u_{t+s}^s is a moving average of the forecast errors from time t to time $t+s$. This method is robust to situations with nonstationary or cointegrated data, so for our application the components of x_t are level data.

Jordà (2005) shows that IRFs generated by the local projections are equivalent to the ones that are calculated from the VAR when the true data generating process (DGP) is a VAR, but that the IRFs for other DGPs that are not true VARs are better estimated using this linear projection method. The IRFs are defined as

$$\widehat{IR}(t, s, d_i) = \widehat{B}_1^s d_i \qquad s = 0, 1, ..., h \tag{2}$$

where $B_1^0 = I$ and d_i is an $n \times 1$ column vector that contains the mapping from the structural shock for the ith element of x_t to the experimental shocks.[8] We construct this mapping matrix using methods suggested in Jordà (2005), which essentially follows methods used in the traditional VAR literature and begins by estimating a linear VAR and applying a Cholesky decomposition to the variance-covariance matrix. We discuss this below in the next subsection.

[8] Here we use Jordà's experimental shock terminology, but the terminology reduced form shock is also appropriate.

One can compute confidence bands using estimates of the standard deviations for the impulses. One issue that needs to be recognized in doing this is that because the DGP is unknown, there could be serial correlation in the error term of (1) induced by the successive leads of the dependent variable. We address this issue by using Newey and West(1987) standard errors which correct for heteroskedasticity and autocorrelation (HAC). Letting, $\widehat{\sum}_s$ be the estimated HAC corrected variance-covariance matrix of the coefficients $\widehat{B}_1^s$, a 68% (or one standard deviation) confidence interval for each element of the IRF at horizon s can be constructed by $\widehat{IR}(t,s,d_i) \pm \sigma(d_i' \widehat{\sum}_s d_i)$, where σ is a $n \times 1$ column vector of ones.

Our extension of this baseline model is to incorporate threshold behavior to the impulse response structure that allows the possibility that the IRF may differ during different phases of the business cycle. We will use the unemployment rate, denoted by w_t, to define the two states of the economy and define our extension to (1) by

$$x_{t+s} = I_t \left[\alpha_R^s + \sum_{i=1}^{p} B_{i,R}^{s+1} x_{t-i} \right] + (1-I_t) \left[\alpha_N^s + \sum_{i=1}^{p} B_{i,G}^{s+1} x_{t-i} \right] + u_{T,t+s}^s \qquad s = 0, 1, ..., h, \tag{3}$$

where most of the notation carries over from above, but subscripts of R or G have been added to the various parameters to indicate bad or good economic times respectively and we use a different notation of $u_{T,t+s}^s$ to denote the error process for this model where the added subscript indicates this is the error for the threshold model.[9] The threshold dummy variable, denoted by I_t, indicates the distinction between bad and good economic times and is defined by the unemployment rate according to

$$I_t = \begin{cases} 1 & \text{for} \quad w_{t-1} \leq w^T, \\ 0 & \text{for} \quad w_{t-1} > w^T, \end{cases} \tag{4}$$

where w^T is the threshold value. We explore two formulations for determining w^T, one is to use a threshold value of 6.5 percent and the other is to allow the threshold to be

[9] We chose the R subscript notation because of the relatively easy mnemonics associated with recession even though the indicator variable does not exactly correspond to that distinction. We did not want to use B for bad times, since B has already been used in the linear projection description. This earlier usage of B was chosen to keep the notation as similar to Jordà (2005) as possible. In our discussion we will typically refer to the R, or below threshold, state as bad economic times.

endogenous and estimate this threshold using methods described in Chan (1993).[10]

By analogy to (2), we define the IRFs for the two states of the economy by

$$\widehat{IR}^{R}(t,s,d_i) = \widehat{B}^{s}_{1,R} d_i \qquad s = 0,1,...,h, \tag{5}$$

and

$$\widehat{IR}^{G}(t,s,d_i) = \widehat{B}^{s}_{1,G} d_i \qquad s = 0,1,...,h, \tag{6}$$

with normalizations $B^0_{1,R} = I$ and $B^0_{1,G} = I$. The confidence bands for the impulse responses of the threshold model are simple extensions of the methodology discussed above.

Now that the local projection approach for computing IRFs has been defined, it may be useful to describe some of its advantages. The primary advantage over the standard VAR approach is its lack of structure from one horizon to the next. This can be understood by reviewing the IRF computation from the typical VAR model. The VAR approach uses the VAR parameters to generate the moving average form from which the IFRs are generated at each horizon, thus the IFRs at all horizons are directly connected to these VAR parameters. On the other hand the local projection method computes the IFRs from a different forecast equation (here (1) or (3)) and thus the structure of the IRFs can vary over the horizon. This allows flexibly when the DGP is nonlinear. So for instance, if the DGP is given by the highly nonlinear structure in (3), the linear VAR structure will not be able to handle this as well as the local projection approach which imposes less structure on the IRF. The local projection method also is attractive relative to methods proposed by Auerbach

[10] We choose the 6.5% threshold because it is often mentioned by the Federal Reserve Bank of the United States as an unemployment rate in which they begin to consider policy changes. See for instance, the Federal Open Market Committee minutes from December 2012 which states, "In addition, all but one member agreed to replace the date-based guidance with economic thresholds indicating that the exceptionally low range for the federal funds rate would remain appropriate at least as long as the unemployment rate remains above $6\frac{1}{2}$ percent, inflation between one and two years ahead is projected to be no more than a half percentage point above the Committee's longer-run goal, and longer-term inflation expectations continue to be well anchored."

and Gorodnichenko (2012).[11] In the STVAR approach suggested by Auerbach and Gorodnichenko (2012), it is assumed that the economy stays in the current state over the horizon in which the impulse responses are calculated. Ramey and Zubairy (2014), for example, argues that this type of assumption is inconsistent with the fact that the average NBER recession period typically last 3.3 quarters, much shorter than the horizons over which one estimates IRFs. On the other hand, the local linear projection approach estimates parameters that are based on data that can be in either state of the world. Thus these parameters have an averaging effect and the projections based on these estimates can be interpreted as weighted averages of the two separate state IRFs.

2.2 Identifying the structural shocks

As suggested in Jordà (2005), the mapping from the structural shocks to the experimental shocks uses the traditional VAR approach described in Sims (1980) which makes use of the Cholesky decomposition. This approach begins with what is called a structural form VAR given by

$$A_0 x_t = \sum_{i=1}^{p} A_i x_{t-i} + \varepsilon_t, \tag{7}$$

where A_i, for $i = 0, ...p$ are 4×4 matrices, p is the lag length for the model and ε_t is a 4×1 vector of structural shocks and we have left out the vector of constant terms to keep things simple. The structural form VAR is not directly estimable without making identification assumptions, so the traditional VAR approach recasts it as a reduced form VAR given by

$$x_t = \sum_{i=1}^{p} A_0^{-1} A_i x_{t-i} + e_t, \tag{8}$$

[11] See Ramey and Zubairy (2014) for details. Also see Auerbach and Gorodnichenko (2013, 2015).

where $e_t = A_0^{-1}\varepsilon_t$ is a 4×1 vector of experimental (or reduced form) shocks.[12] Because the reduced form model has fewer parameters than the structural form model, if one wishes to consider structural model implications, identifying restrictions need to be imposed on the structural parameters and the original suggestion in Sims (1980) was to use the Cholesky decomposition which requires that A_0 be lower (sometimes upper) triangular and this structure implies a contemporaneous causal ordering among the variables, with the variable listed at the top of the vector x_t potentially having contemporaneous causal effects on the remaining variables, the variable listed second from the top potentially having contemporaneous causal effects on all the variables except the first and so on down the list. So, to use this algorithm we must make decisions about how to order the variables in our vector.

We use the ordering that was described earlier in the paper with $x_t = [cc_t \quad c_t \quad y_t \quad f_t]'$. This ordering, for the most part, follows Barsky and Sims (2012) who in a three variable model ordered consumer sentiment first, consumption second and Gross Domestic Product (GDP) third.[13] Here we use labor income rather than GDP, but national income accounts imply the two are similar. To this list, we add financial assets which was ordered last. Ordering financial assets last seems reasonable since impulses in income may have contemporaneous implications for how much people decide to invest, but it is less likely that impulses in financial assets have a contemporaneous impact on income. In other words, because financial assets are highly unpredictable, it is unlikely that agents make short term changes in working decisions as a result of a good or bad year for asset returns.

With these decisions in hand, we can now describe the construction of the d_i vectors used in the impulse response calculations. First note that $e_t = A_0^{-1}\varepsilon_t$ implies that the experimental shock variance is given by

[12] To be consistent with the linear projection approach methods, we use level data for these calculations too. Hamilton (1994) has shown that VAR methods are robust to unknown forms of cointegration. Using level data is quite common and was used by Barsky and Sims (2012) as well.

[13] Barsky and Sims (2012) chose this as their preferred ordering because they identified confidence innovations as "news" on future productivity, which is exogenous to the economy, and has long lasting effects on economic activities like aggregate consumption.

$$e_t e_t^{'} = A_0^{-1} \varepsilon_t \varepsilon_t^{'} \left(A_0^{-1}\right)^{'} = A_0^{-1} \Omega_\varepsilon (A_0^{-1})^{'} \tag{9}$$

where Ω_ε and A_0 are given by

$$\Omega_\varepsilon = \begin{bmatrix} \sigma_{cc}^2 & 0 & 0 & 0 \\ 0 & \sigma_c^2 & 0 & 0 \\ 0 & 0 & \sigma_y^2 & 0 \\ 0 & 0 & 0 & \sigma_f^2 \end{bmatrix} \quad \text{and} \quad A_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \beta_{21} & 1 & 0 & 0 \\ \beta_{31} & \beta_{32} & 1 & 0 \\ \beta_{41} & \beta_{42} & \beta_{43} & 1 \end{bmatrix}.$$

Next note that $e_t = A_0^{-1}\varepsilon_t$ can also be interpreted as showing the mapping from an arbitrary vector of structural shocks given by ε_t into a vector of experimental shocks given by e_t, and that A_0^{-1} provides this mapping. Now, if we define d_i by

$$d_i = A_0^{-1} \Omega_\varepsilon \delta_i, \tag{10}$$

where δ_i is a column vector with a one in the ith position and zeros elsewhere, then d_i has a special interpretation. First note that the term $\Omega_\varepsilon \delta_i$ gives a vector with a one standard error shock for the ith variable in only the ith position, with zeros elsewhere. So by multiplying by A_0^{-1}, d_i can be interpreted as a vector of experimental shocks that arise from a one standard deviation structural shock in the ith variable. This means the impulse response functions given by (2), (5) and (6) show how the vector of variables x_t respond to a one standard deviation shock in the ith structural variable at various forecast horizons.

3 Empirical results

Our empirical analysis uses quarterly data for the US economy from 1960:Q1-2014Q2. We use three different measures of consumer confidence published by the University of Michigan which we denote by ICS, C5Y and C12M where ICS is the Index of Consumer Sentiment and is an overall measure consumer sentiment, C5Y is a measure of consumer confidence that has a five year horizon and C12M is a measure of consumer confidence that has a twelve month horizon.[14] For most of our analysis, we focus

[14] To be more specific, C5Y and C12M are compiled based on a question which askes whether economic conditions would be good or bad over the next five years or twelve months respectively.

on the C12M series, but we look at the others in a robustness investigation later. Our measure of labor income is obtained from the Bureau of Economic Analysis' personal income and its disposition database. We use Line 37 of Table 2.1 that includes wages, salaries, transfer payments and other labor income minus personal contributions for social insurance and personal current taxes. This series is converted to a per capita constant dollar measurement by dividing by the seasonally adjusted personal consumption expenditure chain-type deflator (PCECTPI) obtained from the Federal Reserve Bank of St. Louis (FRED) data base and the population which comes from the BEA personal income and distribution Table 2.1 line 40, reflecting the mid-period total population. Financial assets were tabulated using the Flow of Funds Account of the Federal Reserve Board. Our calculation uses line 9 of Table B.101 for financial assets and subtracted from this line 31 of the same table which are the liabilities to get our measure for financial assets. This series was then adjusted over time by using the price deflator and the population series described above to get a per capital measure in real terms. We use three measures for consumption which are also obtained from the FRED database, including total durable goods (PCDG), non-durable spending (PCND) and motor vehicles (DMOTRC1Q027SBEA). These series are also converted to per capita constant dollar measurements by using the CPI and population series described above. In one robustness exercise we calculate the real per capita spending on "other durable goods" by excluding the real per capita spending on motor vehicles from real per capita spending on durable goods. Finally, our switching variable is the quarterly unemployment rate (UNRATE) which comes from the FRED database. In processing the data, all variables are converted to logarithms and left in log level terms for the computations.

Some preliminary insights into the business cycle aspects this paper investigates can be obtained by plotting the consumer sentiment index and the unemployment rate over time. Figure 1 does this and also provides shaded regions which indicates the

Details on the construction of these indices can be found at http://www.sca.isr.umich.edu/.

exogenous threshold unemployment rate we chose of 6.5%.[15] One important insight from this figure is that it shows an inverse relationship between consumer confidence and unemployment. Furthermore, sharp rises in the unemployment rate are associated with sharp declines in confidence. One of the reasons for this could be that the unemployment rate data is published frequently and unlike some other real time data series, seldom goes through large revisions, which we interpret to mean that consumers regard it as a reliable economic indicator. Also of particular relevance to this investigation is that there can be large increases in confidence during high unemployment phases. We wish to investigate whether impulses to consumer confidence during the high unemployment phases imply the same impact on consumption as impulses to consumer confidence do during the low unemployment phases.

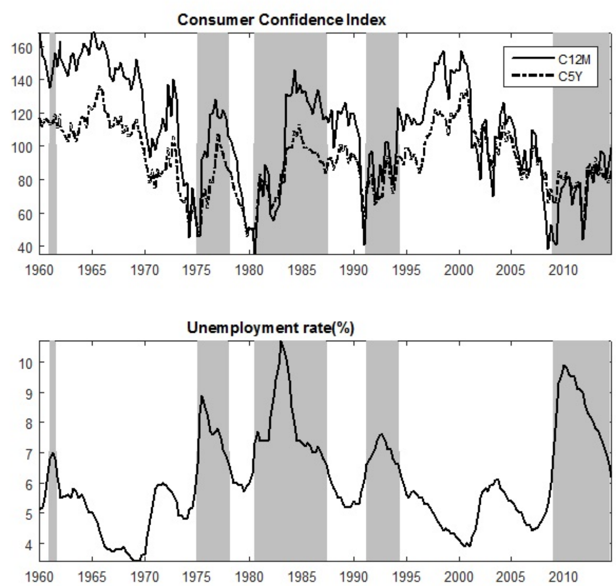

Figure 1: Consumer confidence C12M and C5Y and unemployment over time

Linear model

We begin by looking at the impulse response functions in the linear model given by (1), which we regard as the benchmark model that summarizes the current frontier of the literature. We focus on only the confidence shock results since this is where we

[15] In our robustness section we also estimated models with an endogenous threshold.

want to contribute to the literature by showing that differences arise when considering a threshold model. Using the Akaike Information Criterion (AIC) we find that two lags are appropriate in (1).

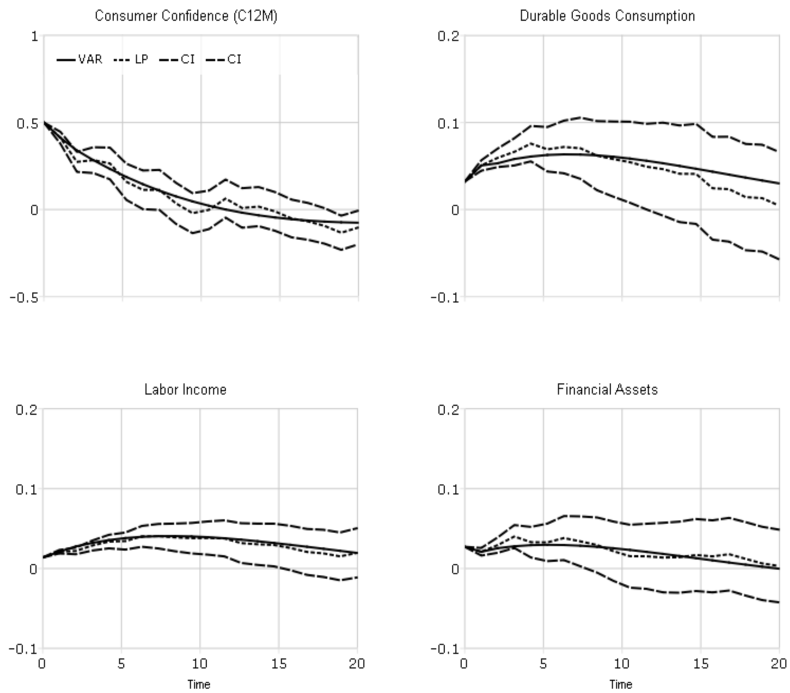

Figure 2a: Impulse responses from one standard deviation confidence shock

Model with durable goods

Figures 2a and 2b show the results for models that use C12M as the measure of confidence with Figure 2a using durable goods as the consumption variable and Figure 2b using non-durable goods as the consumption variable.[16] In each figure four lines are plotted for 20 quarters, or five years. The short dashed line represents the impulse response obtained from the local projection model given by (1), the solid line represents the impulse response based on a two lag VAR model using the same Cholesky ordering to identify the structural shocks as in the local linear method, and the two long dashed lines represent the one standard deviation bands around the local

[16] Our results are unchanged with the measures of confidence as shown in the section of robustness check analysis.

projection impulse response function using the Newey-West method for computing standard errors described earlier.

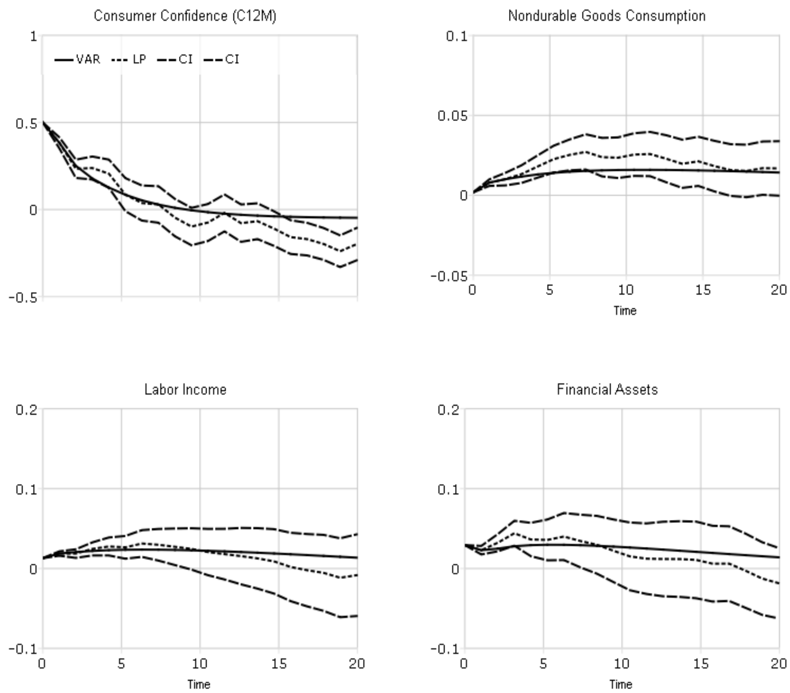

Figure 2b: Impulse responses from one standard deviation confidence shock

Model with nondurable goods

Both figures show that the local projection method and the VAR method yield very similar impulse response patterns with the solid line mostly tracking the short dashed line over the 20 quarter series and, with the exception of the confidence response in Figure 2b, always remaining inside the one standard error bands. One notable difference is that the VAR method impulses are smoother, and this reflects the construction process where the vector moving averages used to find the impulse response functions are always functions of the same estimated VAR coefficients, while the local linear projection method does not impose any construction relationships between impulses at different horizons.

Comparing the different figures we see that confidence responses start near 0.6 and decline slowly toward zero, reaching it after about 10 quarters, the response of consumption goods has a bit of a hump-shaped pattern, with the hump in the durable goods case being more pronounced, the response of income also has a humped pattern increasing for about eight quarters to about 0.3 before starting to decline while financial assets have a more wobbly pattern, with a short decline, then a slow increase which reaches a peak about six quarters latter and then begins to decline. One notable difference is that the response for durable goods is much larger in magnitude.[17] The response for durable goods consumption is roughly three times the size of nondurable goods. We interpret these findings as illustrating the following economic processes. The initial jump in confidence results in an increase in consumption spending which has a multiplier effect on income and thus results in both of these responses exhibiting hump shaped patterns. On the other hand, financial assets has a wobble with an initial decline, because the initial increase in consumption is financed more from existing savings than the rising income. However, as income rises over time, financial assets are rebuilt and start to increase

Our results for these linear models show some differences from that of Barsky and Sims (2012). In particular, Barsky and Sims (2012) showed that confidence innovations have long lasting effects on aggregate consumption and income, while our results show that by disaggregating consumption there are some differences. In particular, while we see long lasting effects for nondurable goods (in both the VAR and the preferred linear projection models), we also see the impulses for durable goods reaching zero near the 20 quarter horizon for the preferred linear projection model. Another difference that the disaggregation reveals is that the short term impact on durable goods consumption is considerable larger than the impact on nondurable goods. Barsky and Sims (2012) did not find this type of result since they simply used aggregate consumption.

Threshold local projection model

Figures 3a and 3b show the impulse response plots for the threshold regression

[17] Please note, this fact can be easily missed because we were forced to use different vertical scales to plot these series.

model (3) using the unemployment rate as the threshold variable and a value of 6.5 as the threshold. Although these figures plot the same four types of responses as in Figures 2a and 2b and use the same scales on the vertical axis, there are a few differences in the plotting notations relative to the plots in Figures 2a and 2b. In particular, we use a convention of plotting the impulse responses for the good state, the state when unemployment is below 6.5%, using a short dashed line and its one standard deviation confidence band using long dashed lines and then for the bad state, the state when unemployment is above 6.5%, we plot it as a solid line without confidence bands. Thus in these figures, the good state takes the previous role used by the linear projection model and the bad state takes the previous role used by the VAR model.

Figures 3a and 3b show some big differences between the good state and bad state impulses. For instance, we see that the stimulative effects of a confidence shock on durable and nondurable consumption goods during good times is considerably higher than the stimulative effects during economic bad times. For durable goods, the bad times confidence shock to durable consumption goods is outside the one standard error bands for roughly ten quarters, whereas for the nondurables goods, it is outside the one standard error bands for quarters two to seven. Similar things can be seen in some of the other plots. So for instance, in the durable goods model, the bad state response for income is outside the one standard error band for about seven quarters while in the nondurable goods model, the bad state response for income tracks the good state response fairly closely. Also, for both the durable and nondurable good models, the bad state response for financial assets stay outside the one standard deviation good state bands for roughly the first seven to nine quarters. Overall, these results show that the stimulative effects of a confidence shock has a considerable smaller effect on consumption, income and financial assets during bad economic times than during good economic times and shows that models that ignore the threshold behavior modeled in (1) are missing something.

Some further insights can be obtained by comparing the magnitudes of the changes between Figures 2a and 2b with the magnitudes of the changes in Figures 3a and 3b. In particular, this comparison shows that the stimulative effect on both types of consumption from a confidence shock during good times is much larger than the average stimulative effect given in Figures 2a and 2b. In addition, the multiplier effect on income in the good state is also larger than the average effect in Figures 2a and 2b. These results also show that the stimulative effects of a confidence shock are different than the simple average effects in Figures 2a and 2b and further show the need for using a threshold model to investigate this issue.

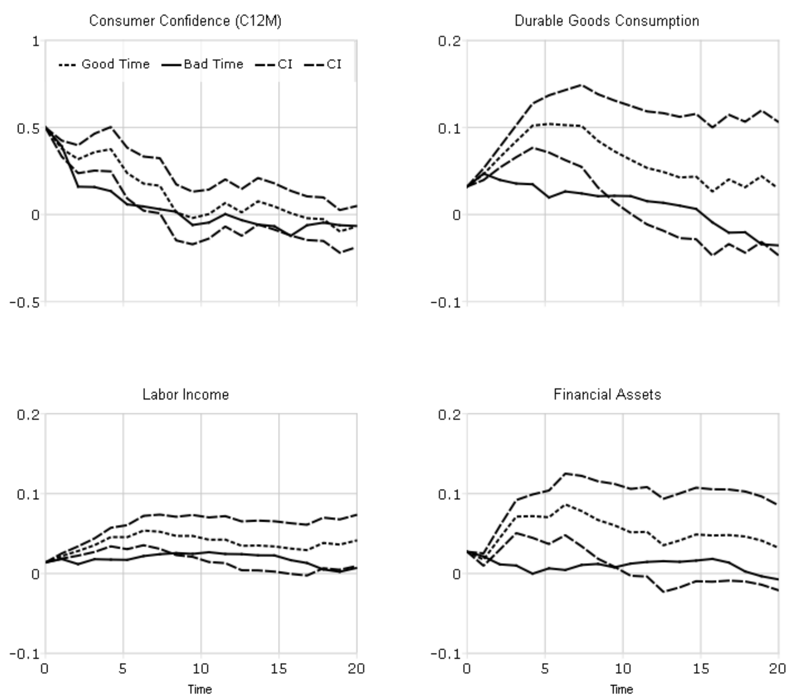

Figure 3a: Impulse responses from one standard deviation confidence shock

Model with durable goods

Economically we can interpret these differences as showing that during good economic times confidence has an amplifying effect on the economic condition, while during poor economic times, confidence shocks are not strong enough to generate lasting improvements in the economic condition. These results are also consistent with findings in Berger and Vavea (2014, 2015) who show that fewer households purchase durable during recessions because of substantial adjustment costs that exist with this type of purchase, leading aggregate durable goods spending to be relatively less sensitive to confidence shocks.

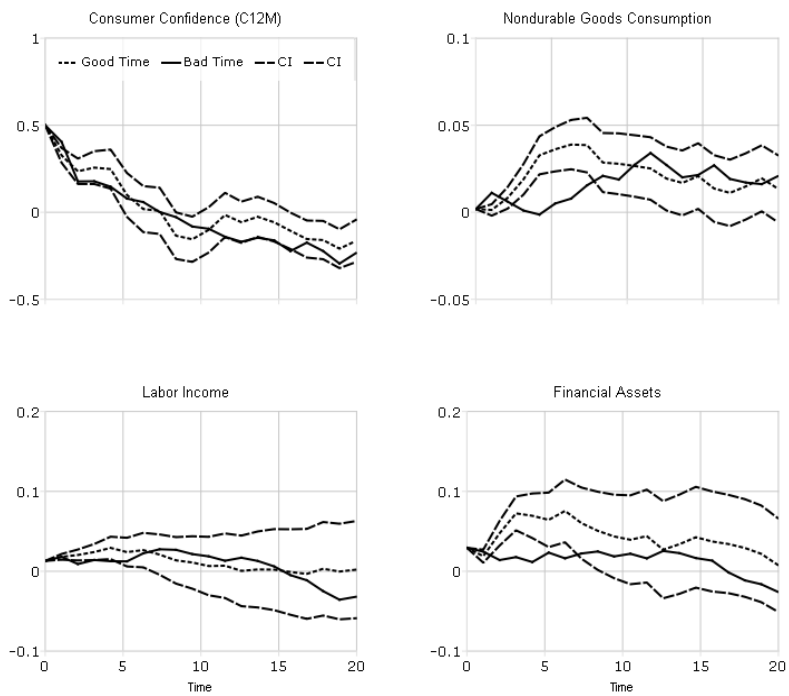

Figure 3b: Impulse responses from one standard deviation confidence shock

Model with nondurable goods

Another useful comparison is to recognize how these results would differ from the STVAR methods such as those used in Auerbach and Gorodnichenko (2012). In the STVAR approach the model assumes that the economy remains in whatever state it begins in. This would have an amplifying effect on the good state results found here. In particular, the results found here can be interpreted as showing coefficients that are weighted averages for the economic outcomes going forward. Although the persistence of a good economy is relatively high, this model does not assume that the economy will remain in a good state for the next 20 quarters, but instead builds in an average transition to the bad state. Because of this averaging effect, the impulse responses are going to be more modest relative to a model that assumes that the economy remains in a good state. By the same token, the STVAR model which assumes remaining in the bad for 20 quarters will have a more negative outcome than this model which assumes an average transition back to the good state.

Granger Causality results

Another way to investigate the necessity of the threshold structure is to run Granger Causality tests. The impulse response results show that the response of the two consumption variables to confidence shocks is different depending on what the initial state of the economy is. Here we investigate whether Granger Causality implications also depend on what the initial state of the economy is.

To conduct this investigation, we ran a simple VAR given by

$$\widetilde{x}_t = \widetilde{D} + \sum_{i=1}^{p} \widetilde{D}_i \widetilde{x}_{t-i} + \zeta_t \tag{11}$$

and a threshold VAR given by

$$\widetilde{x}_t = I_t \left[\widetilde{D}_R + \sum_{i=1}^{p} \widetilde{D}_{Ri} \widetilde{x}_{t-i} \right] + (1 - I_t) \left[\widetilde{D}_G + \sum_{i=1}^{p} \widetilde{D}_{Gi} \widetilde{x}_{t-i} \right] + \zeta_{T,t} \tag{12}$$

where $\widetilde{D}$, $\widetilde{D}_R$ and $\widetilde{D}_G$ are 4×1 vectors of parameters, $\widetilde{D}_i$, $\widetilde{D}_{Ri}$ and $\widetilde{D}_{Gi}$, for $i = 1, ...p$, are 4×4 matrices of parameters, ζ_t and $\zeta_{T,t}$ are white noise errors, p is the lag length for the VAR and I_t is given by (4) with $w^T = .065$ and we use $\widetilde{x}_t = [cc_t \quad \Delta c_t \quad \Delta y_t \quad \Delta f_t]'$[18]. Initial AIC investigations for (11) found that $p = 1$ was optimal. We used this same value for (12).

From these regressions we conducted various Granger causality tests as follows. For the (11) model, the Granger Causality test that confidence does not Granger Cause changes in consumption can be undertaken by testing $H_0 : \widetilde{D}_1(2,1) = ... = \widetilde{D}_p(2,1) = 0$ where the additional (2,1) notation indicates we are looking at the coefficients for the second row and first column of the various matrices. Similarly,

[18] We use $\widetilde{x}_t = [cc_t \quad \Delta c_t \quad \Delta y_t \quad \Delta f_t]'$ rather than $x_t = [cc_t \quad c_t \quad y_t \quad f_t]'$ because the last three variables are nonstationary and Granger Causality tests require stationarity. We also use tilde notations on the matrices, such as $\widetilde{D}$, to indicate that these matrices correspond to the $\widetilde{x}_t$ data vector.

for the (12) model, the Granger Causality test that in good economic times, confidence does not Granger Cause changes in consumption can be undertaken by testing $H_0 : \widetilde{D}_{G1}(2,1) = ... = \widetilde{D}_{Gp}(2,1) = 0$, while the Granger Causality test that in bad economic times, confidence does not Granger Cause changes in consumption can be undertaken by testing $H_0 : \widetilde{D}_{R1}(2,1) = ... = \widetilde{D}_{Rp}(2,1) = 0$. Analogous tests that investigate Granger Causality from changes in consumption to consumer confidence have similar structure.

Table 1 shows these results using the C12M measure for consumer confidence and three consumption good measures. Here we include a third measure of consumption, motor vehicles, which is a subcomponent of durable goods to investigate the robustness of the durable goods results. In this table we use the shorthand $H_0 : cc_t \nRightarrow c_t$ to indicate the null that consumer confidence does not Granger Cause changes in consumption rather than writing out the full set of parameters which we described above. Similarly, we write $H_0 : c_t \nRightarrow cc_t$ to indicate the null that consumption does not Granger Cause changes consumer confidence. In addition, we use the notation "Bad ET" to indicate the above threshold Granger Causality test and "Good ET" to indicate the below threshold tests.

Table 1 is organized into two horizontal panels with the top one exploring whether consumer confidence causes changes in consumption and the bottom one exploring whether changes in consumption causes consumer confidence. Each horizontal panel has three subpanels with each subpanel having two columns that explore the null using (11), which we refer to as the linear model and the other using (12) which we refer to as the threshold model.

Looking across the first row with numbers, we see that the linear model finds that consumer confidence causes all three categories of consumption changes.[19] Next, looking at the second row with numbers, we see that the threshold model never finds

[19] The linear model has F-distribution with 1 degrees of freedom in the numerator and 211 in the denominator, while the threshold model has 1 in the numerator and 206 in the denominator. For the linear model, the 1%, 5% and 10% critical values are 6.757, 3.886 and 2.729 respectively while for the threshold the 1%, 5% and 10% critical values are 6.760, 3.887 and 2.730 respectively, respectively.

that consumer confidence causes changes in consumption during bad economic times while the third row with numbers shows that consumer confidence causes changes in consumption during good economic times. Overall we summarize the findings as showing that the threshold model shows the same causality implications as the linear model only during good economic times, but finds that during bad economic times, consumer confidence has no causal implications for changes in consumption.

Table 1: Granger Causality tests between consumer confidence and consumption

	Durable		Motor Vehicles		Nondurable	
	Linear	Threshold	Linear	Threshold	Linear	Threshold
	Consumer confidence causes consumption					
$H_0 : cc_t \nRightarrow c_t$	10.48***	-	3.60*	-	6.06**	-
$H_0 : cc_t \nRightarrow c_t$ (Bad ET)	-	1.51	-	0.01	-	0.31
$H_0 : cc_t \nRightarrow c_t$ (Good ET)	-	25.44***	-	19.60***	-	3.74*
	Consumption causes consumer confidence					
$H_0 : c_t \nRightarrow cc_t$	1.62		0.83		4.51**	
$H_0 : c_t \nRightarrow cc_t$ (Bad ET)		0.59		0.25		2.95*
$H_0 : c_t \nRightarrow cc_t$ (Good ET)		1.12		0.29		0.01

In Tables 1, 3, and 4, ***, **, and * denote the significance level at 1%, 5% and 10 % respectively.

Next focusing on the lower panel of Table 1, the first row with numbers shows that only nondurable goods rejects the null of no causality from changes in consumption to consumer confidence for the linear model. The next two rows show that the threshold model implies no causality from durable goods and motor vehicles to consumer confidence in either bad or good economic times, while it does show there to be causality from changes in consumption of nondurable goods to consumer confidence during recessions.

Overall these results are consistent with the findings in the impulse response analysis. The Granger Causality results show that there are differences between the linear model and the threshold model for durable goods and motor vehicles. In particular, the threshold model only shows causality from consumer confidence to changes in durable goods consumption and changes in motor vehicle purchases during good economic times and shows no causal implications during weak economic times.

Variance decomposition results

It is also possible to make a case for the threshold models by using variance decomposition analysis, which is a popular tool from the traditional VAR analysis. To understand the variance decomposition method using linear projections, we provide a brief overview of the procedure. For theoretical detail, we refer the reader to Jorda (2005). First note that the mean squared error of the forecast error is given by

$$MSE_u(E(x_{t+s}|X_t)) = E(u^s_{t+s}u^{s'}_{t+s}) \qquad s = 0, 1, ..., h. \tag{13}$$

This can be estimated by using $\widehat{\Sigma}_{u^s} = \frac{1}{T}\sum_{t=1}^{T} \widehat{u}^s_{t+s}\widehat{u}^{s'}_{t+s}$ where $\widehat{u}^s_{t+s} = x_{t+s} - \widehat{\alpha}^s + \sum_{i=1}^{p} \widehat{B}^{s+1}_i x_{t-i}$. The diagonal elements of this will be the variance of the s step ahead forecast errors for each of the elements in x_t. Next, defining the $n \times n$ experimental choice matrix D by the columns d_i defined by (10). Renormalizing MSE_u by the choice matrix D into

$$MSE(E(x_{t+s}|X_t)) = D^{-1}E(u^s_{t+s}u^{s'}_{t+s})D^{'-1} = D^{-1}\Sigma_{u^s}D^{'-1} \qquad s = 0, 1, ..., h. \tag{14}$$

From (14), we can calculate the traditional variance decompositions by directly plugging in the sample-based equivalents from the projections in (1). Extensions of this calculation to the threshold models can be done using a straightforward extension of the vector x_t by putting terms $I_t x_t$ in the upper half of the new vector and $(1 - I_t)x_t$ in the lower half of the new vector.

Table 2 shows the results of this exercise. To save space, only the results showing the percent of the total forecast error variance attributable to confidence innovations are reported. Before summarizing the results of this exercise with general comments, it will be useful to walk through the structure of the table to understand what the numbers mean. The table is organized into two vertical panels, with columns two through four showing the results when using durable goods as the consumption variable and columns five through seven showing the results when using nondurable goods as the consumption variable. The table is also organized into four horizontal panels each of which corresponds to a different forecast horizon. Only the variance decompositions for forecast horizons of four, eight, twelve and twenty quarters are reported. Next, focusing on the top horizontal panel, which summarizes the variance for the four quarter horizon, we see it is organized into three rows, with the first row showing the results for the linear model given by (1) and the next two rows showing the results for bad economic times, which we denote by "Bad ET", and the results for good economic times, which we denote by "Good ET", from the threshold model given by (3). Focusing on the durable goods models, we see that the linear model shows that confidence innovations account for 29.6% of the forecast error variance for durable goods, 24.4% of the forecast error variance for labor income and 7.8% of the forecast error variance for financial assets for the forecast horizon of 4 quarters. Similarly, we see in the durable goods threshold model, during bad economic times, confidence innovations account for 25.1% of the forecast error variance for durable goods, 19.5% of the forecast error variance for labor income and 5% of the forecast error variance for financial assets for the forecast horizon of 4 quarters, while during good economic times, confidence innovations account for 30.1% of the forecast error variance for durable goods, 25.1% of the forecast error variance for labor income and 13% of the forecast error variance for financial assets for the forecast horizon of 4 quarters. The remaining sub panels of the table have similar interpretations.

Now that we understand the structure of the table we can see the following results. The linear model shows that confidence innovations account for 11.4% of the forecast error variance of durable goods at the four quarter horizon and the percentage of the forecast error variance rises quickly to 38.8% at the eight quarter horizon, 43.0% at the twelve quarter horizon and 43.8% at the twenty quarter horizon. This result is consistent with results reported in Barsky and Sims (2012). However, our threshold model shows there are differences in the variance decomposition according to whether the current state corresponds to good or bad economic times. Moving down the table, we see that during bad economic times the portion of the forecast error variance of durable goods attributable to confidence innovations is considerably smaller at the eight, twelve and twenty quarter horizons than the portion of the forecast error variance of durable goods attributable to confidence innovations during good economic times at the eight, twelve and twenty quarter horizons.[20] This shows that a proper modeling structure for this set of variables is the threshold model.

Interestingly, this difference between the linear model and the threshold model in the good and bad economic times is not so important when the consumption variable is nondurable goods. Here we see mostly similar variance decompositions at the four, eight, twelve and twenty quarter horizons for the linear, and two states of the threshold model. This is not surprising given the impulse response results shown earlier.

It is also useful to note the outcomes for labor income and financial assets. Looking at the columns for the variance decompositions for labor income, we see that the linear model and the threshold model in both the good and bad economic times are relatively (i.e. relative to the durable goods results) similar at the different forecast horizons. This is not surprising since the impulse response functions in Figures 3a and 3b showed the labor impulses to be similar in the two regimes. On the other hand, the variance decompositions for financial assets do show greater difference between the linear model and the threshold model results, particularly at the eight, twelve and twenty period forecast horizons. Again this is similar to the impulse response functions in Figures 3a and 3b which showed somewhat large differences for financial assets and this result likely reflects the differences discussed earlier where agents make use of financial assets to make changes in consumption sending patterns due to changes in consumer confidence.

[20] In some results not presented here to keep space down, a model using motor vehicles produced very similar variance decompositions to the durable goods models in Table 2.

Table 2: Percent of total forecast error variance attributable to confidence innovations

	Durable Goods Models			Nondurable Goods Models		
States	Dur Good	Lab Income	Fin Asset	Nondur Good	Lab Income	Fin Asset
			Forecast horizon of 4 quarters			
Linear	29.56	24.41	7.75	3.87	16.80	8.34
Bad ET	25.11	19.49	4.96	2.68	16.08	7.21
Good ET	30.12	25.13	13.00	3.37	14.36	12.58
			Forecast horizon of 8 quarters			
Linear	38.78	41.44	9.82	15.68	21.67	9.88
Bad ET	18.88	30.01	3.43	4.37	24.97	7.18
Good ET	48.17	43.90	25.45	18.64	15.31	20.10
			Forecast horizon of 12 quarters			
Linear	43.02	47.69	9.90	22.88	17.84	9.20
Bad ET	17.03	39.59	3.78	20.67	26.64	8.29
Good ET	52.98	47.77	29.20	23.32	8.30	19.40
			Forecast horizon of 20 quarters			
Linear	43.78	31.76	9.49	26.50	11.42	8.34
Bad ET	17.80	21.60	3.61	28.43	18.67	8.48
Good ET	52.77	39.20	30.62	24.08	5.46	18.21

4 Robustness

The baseline results above suggest that the effects of a confidence shocks on durable goods spending are state dependent with considerably smaller effects during weak economic times. This section discusses some alternative specifications for the model that were investigated in order to determine the robustness of the results. For the most part, the impulse response patterns for these alternative exercises were similar to those above, so rather than provide all the plots, we only describe the exercises and some of the results. An appendix with additional details is available from the authors upon request.

The first exercise was to consider alternative measurements for consumer confidence. For this exercise we used the previously noted indexes ICS and C5Y, thus redefining x_t to $x_t = [AltConf_t \quad c_t \quad y_t \quad f_t]'$, where $AltConf_t$ denotes a vector using either $C5Y_t$ or ICS_t. The results of this exercise produced virtually identical impulse response plots to those in Figures 3a and indicate that during good economic times, the impulse has favorable effects on durable goods, while during bad economic times, the impulse has very little effect. The next exercise was to consider some of the subcomponents of the durable goods index. Here we considered the motor vehicle purchases and the other durable goods subcomponets of the durable goods series. For this exercise we redefine x_t to have these alternative consumption series. The results were unchanged relative to Figure 3a. Next we used an alternative Cholesky ordering in which consumer confidence was ordered last and again the impulse response function for durable goods was unchanged and again showed large effects during good times, and weak effects during bad times. Next an endogenous threshold model was considered and it also resulted in no changes. A one lag model which is optimal by the Schwartz Bayesian lag length selection criterion did produce some changes. But these changes showed even stronger differences between the good times and bad times impulse responses. Finally, we used a broader measure of assets called total assets, which includes housing assets, and this too resulted in no differences.

Overall, the conclusion that consumer confidence innovations lead to increases in durable goods consumption during good economics times, but has small effects during bad economic times proved to be robust.

5 Connections between news and confidence

Another insightful exercise is to extend the analysis connecting news and confidence done in Barsky and Sims. (2012) to include threshold behavior. The previous analysis has shown that there is quite different connections between confidence and economic activity when the economy is booming and when it is not booming. This section shows that the connection between news and confidence found in Barsky and Sims (2012) is also state dependent.

Here we also run three regressions connecting various subindexes of the Michigan Consumer Confidence Survey with confidence. As in Barsky and Sims (2012), these subindexes are interpreted to measure "news" and the question is whether these news shocks impact innovations in consumer confidence.[21] Table 3 shows the results of regressions analogous to those in Barsky and Sims (2012) in columns 2 through 4 while columns 5 through 7 show the results from running a threshold regression with unemployment as the threshold variable and a threshold value of 6.5% unemployment rate. To reduce the amount of space that the table contains, the first 18 rows serve double duty in that they report the coefficient estimates and standard errors for those variables in the nonthreshold regression and they report the coefficient estimates for those variables in the below threshold (i.e. good economic times) case for the threshold model. The above threshold (i.e. bad economic times) values are given in the bottom 18 rows of the table. In addition, we use the notation “GT” for below threshold case rather than the previously used “Good ET”, “BT” for above threshold rather than the previously used “Bad ET” and abbreviated “Favorable” by “Fav” and “Unfavorable” by “Unfav” in order to reduce the width of the table.

The nonthreshold regressions confirm the results in Table 4 of Barsky and Sims (2012) showing that favorable employment news, favorable and unfavorable price movements are significant in the baseline model, adding favorable stock price movements to the models shows it to be significant and further adding unfavorable government spending news and energy crises are also significant. The table also shows unfavorable unemployment is not significant in two of the models and, when added, unfavorable stocks and favorable government are not significant.

[21] We followed an analogous procedure for determining confidence innovations as the one following in Barsky and Sims (2012). In particular, confidence innovations are recovered from a reduced form VAR using the Cholesky ordering with confidence ordered first.

Table 3: Regressions of confidence innovations on news about economic conditions

News categories						
Fav employment (GT)	0.003*	0.001	0.003*	0.007**	0.004	0.003
	(0.002)	(0.002)	(0.002)	(0.003)	(0.004)	(0.004)
Fav price (GT)	0.027***	0.026***	0.019**	0.035***	0.033***	0.031***
	(0.008)	(0.008)	(0.008)	(0.011)	(0.012)	(0.012)
Unfav employment (GT)	-0.001	-0.001*	-0.000	-0.000	-0.001	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Unfav price (GT)	-0.008***	-0.008***	-0.006***	-0.010***	-0.009***	-0.008***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Fav stocks (GT)		0.011*	0.011*		0.008	0.008
		(0.006)	(0.006)		(0.009)	(0.009)
Unfav stocks (GT)		0.002	0.001		0.020	0.010
		(0.003)	(0.003)		(0.014)	(0.014)
Fav government (GT)			0.004			0.06
			(0.009)			(0.014)
Unfav government (GT)			-0.009***			-0.008
			(0.003)			(0.006)
Energy crisis (GT)			-0.008*			-0.008*
			(0.004)			(0.004)
Fav employment (BT)				0.002	0.002	0.004
				(0.002)	(0.002)	(0.003)
Fav price (BT)				0.022*	0.018	0.004
				(0.011)	(0.012)	(0.003)
Unfav employment (BT)				-0.001	-0.000	0.001
				(0.001)	(0.001)	(0.001)
Unfav price (BT)				0.001	0.000	- 0.000
				(0.006)	(0.006)	(0.006)
Fav stocks (BT)					0.006	0.019
					(0.011)	(0.012)
Unfav stocks (BT)					-0.018	-0.009
					(0.013)	(0.013)
Uav govt (BT)						-0.002
						(0.012)
Unfav govt (BT)						-0.011***
						(0.003)
Energy crisis (BT)						-0.003
						(0.012)
R^2	0.14	0.16	0.23	0.18	0.19	0.25

Terms in parenthesis are standard errors.

Shifting over to the threshold models, the table shows that during good economic times, most of the variables that were significant in the nonthreshold model continue to be significant in this state of the economy and none of the variables that were formerly insignificant have become significant. Looking at the bottom of the table we see that only the unfavorable government variable is significant in any of the models.[22] Overall, these results show that except for two variables, news shocks only have significant implications for confidence during good economic times. Furthermore, one of these variables had only marginal significance in the simplest model and this significance disappeared when additional explanatory variables were added to the investigation.

Another insightful exercise is to run Granger Causality tests which investigate whether news has causal implications for confidence. Since including too many unnecessary variables would reduce the power of the Granger Causality tests, we focus on a few simple tests. In particular, we consider tests which investigate whether the four variables in the baseline model from Table 3 above have Granger Causality implications toward consumer confidence. Those variables are favorable employment, favorable price, unfavorable employment and unfavorable price news. So to investigate this we run a simple five variable Granger Causality regression with consumer confidence (C12M) as the dependent variable and explanatory variables including lagged consumer confidence, lagged favorable employment, lagged favorable prices, lagged unfavorable employment and lagged unfavorable price news.[23] All these variables are stationary in levels, so no differencing is necessary. We used two lags for the independent variables based on the AIC. The results of various tests are summarized in Table 4 below.

Table 4 is organized into three columns with the second column showing the results for the nonthreshold model, which is analogous to the model given by (11) and we call the linear model, while the next two columns show the results for good

[22] The unfavorable government variable could reflect other political features not modeled here. In particular, interviewees were more likely to responsed to this question during election years. So perhaps this factor is at work here.

[23] It is probably useful to emphasize that the Granger Causality tests in Table 4 only use lags of the explanatory variables while the regressions in Table 3 followed Barsky and Sims (2012) and used contemporaneous values for the explanatory variables.

economic times and bad economic times cases for a model analogous to (12). The linear model results show that there are highly significant Granger Causal results from unfavorable employment news and unfavorable price news toward consumer confidence while there is insignificant Granger Causal results from favorable employment news and favorable price news to consumer confidence. Looking at the Threshold model results, we see high levels of significance for the below threshold or good economy case for unfavorable employment and unfavorable price new, but for the above threshold or bad economy case we see that the unfavorable price news does not Granger cause consumer confidence and unfavorable employment news has reduced significance. Overall, these results show differences between the linear model and the threshold model which takes into account the state of the economy.

Table 4: Granger Causality tests for news causing consumer confidence

	Linear	Good Economic Times	Bad Economic Times
H_0 : Fav employment$_{t-i} \nRightarrow cc_t$	0.46	0.74	0.27
H_0 : Fav price$_{t-i} \nRightarrow cc_t$	2.05	1.80	1.80
H_0 : Unfav employment$_{t-i} \nRightarrow cc_t$	6.48***	3.85**	2.65*
H_0 : Unfav price$_{t-i} \nRightarrow cc_t$	6.46***	5.27***	1.69

6 Conclusion

This paper investigates whether consumer confidence innovations have long lasting effects on various types of consumption. We find that the connection between consumer confidence and consumption is not robust when considering the state of the economy and some measures of consumption. In particular, during weak economic times, the results of impulse response analysis, Granger Causality tests and variance decomposition investigations show that consumer confidence innovations do not imply that durable goods consumption will change. These results proved to be robust to alternative measurements of consumer confidence, alternative subcomponents of durable goods, alternative Cholesky orderings and alternative measurements for asset holdings. We also investigated the connection between news and consumer confidence and found it is also state dependent.

These results have important implications for recent policy debates which have speculated that improving consumer confidence can lead to a faster economic recovery. Our results show that improving consumer confidence may not produce the economic benefit that has been speculated.

References

[1] Acemoglu, D., and Scott, A. 1994. Consumer confidence and rational expectations: Are agents' beliefs consistent with the theory?. *The Economic Journal* 104, 1-19.

[2] Akerlof, G. A., and Shiller, R. J. 2010. Animal spirits: How human psychology drives the economy, and why it matters for global capitalism. Princeton University Press.

[3] Auerbach, Alan J., and Yuriy Gorodnichenko. 2012. Measuring the output responses to riscal policy. *American Economic Journal: Economic Policy* 4, 1-27.

[4] Auerbach, Alan J., and Yuriy Gorodnichenko. 2013. Fiscal multipliers in recession and expansion, *NBER working paper No.* 18578.

[5] Auerbach, Alan J., and Yuriy Gorodnichenko. 2015 Effects of fiscal shocks in a globalized world, *NBER working paper No.* 21100.

[6] Bachmann, R., and Sims, Eric, R. 2012 Confidence and the transmission of government spending shocks, *Journal of Monetary Economics* 59, 235–249.

[7] Bachmann, Rüdiger, Tim O. Berg, and Sims, Eric, R.. 2015. Inflation expectations and readiness to spend: Cross-sectional evidence." *American Economic Journal: Economic Policy* 7, 1-35.

[8] Barsky, Robert B., and Eric R. Sims. 2012. Information, animal spirits, and the meaning of innovations in consumer confidence. *American Economic Review* 102, 1343-77.

[9] Batchelor, R., and Dua, P. (1998). Improving macroeconomic forecasts: The role of consumer confidence. *International Journal of Forecasting* 14, 71-81.

[10] Berger, D., and Vavra, J. 2014. Measuring how fiscal shocks affect durable dpending in recessions and expansions, *The American Economic Review, Papers and Proceedings* 104, 112-115.

[11] Berger, D., and Vavra, J. 2015. Consumption dynamics during recessions. *Econometrica* 83, 101-154.

[12] Ben Bernanke, Monetary Policy Report to Congress, February 24, 2010.

[13] Ben Bernanke, 1985. Adjustment costs, durables, and aggregate consumption, *Journal of Monetary Economics* 15, 41-68

[14] Berry, S., and Davey, M. 2004. How should we think about consumer confidence?. *Bank of England Quarterly Bulletin*, Autumn.

[15] Blanchard,. O. 1993. Consumption and the Recession of 1990-1991, *The American Economic Review,* Papers and Proceedings. 83, 270-274.

[16] Chan, K.S. 1993. Consistency and limiting distribution of the least squares estimator of a threshold autoregessive model. *The Annals of Statistics* 21, 520-33.

[17] Campbell, John Y., and N. Gregory Mankiw. 1989. Consumption, income and interest rates: reinterpreting the time series evidence. *NBER Macroeconomics Annual* 4, 185-246.

[18] Campbell, J.Y. and Mankiw, N.G. 1990. Permanent income, current income, and consumption. *Journal of Business & Economic Statistics* 8, 265-279.

[19] Campbell, J.Y. and Mankiw, N.G. 1991, The response of consumption to income: A cross-country investigation. *European Economic Review* 35, 723-67.

[20] Carroll, D. Christopher, Fuhrer, C. Jeffrey, and Wilcox, W. David, 1994, *The American Economic Review* 84, 1397-1408.

[21] Cochrane, J. 2009. Fiscal stimulus, fiscal inflation, or fiscal fallacies?. University of Chicago Booth School of Business. Manuscript, February.

[22] Cochrane, J. H. 1994, December. Shocks. *Carnegie-Rochester Conference series on public policy* 41, 295-364.

[23] Feldstein, M. Our Economic Dilemma, *The Wall Street Journal,* February 20, 2008.

[24] Granger, C. W. 1988. Some recent development in a concept of causality. *Journal of Econometrics* 39, 199-211.

[25] Granger, Clive W. J. and Teräsvirta, Timo, 1993. *Modelling Nonlinear Economic Relationships.* Advanced Texts in Econometrics. Oxford University Press

[26] Hamilton, J. D. 1994. *Time series analysis*, Princeton: Princeton University Press.

[27] Howrey, E. P. 2001. The predictive power of the index of consumer sentiment. *Brookings papers on economic activity*, 175-216.

[28] Jordà, O. 2005. Estimation and inference of impulse responses by local projections, *The American Economic Review* 95, 161-182

[29] Katona, G. 1968. Consumer behavior: Theory and findings on expectations and aspirations. *The American Economic Review* 58, 19-30.

[30] Krugman, P. How did we know the stimulus was too small? *The New York Times*, The opinion pages, July 28, 2010

[31] Lettau, M., and Ludvigson, S.C., 2004. Understanding trend and cycle in asset values: Reevaluating the wealth effect on consumption. *American Economic Review* 94, 276-299.

[32] Lettau, M., and Ludvigson, S. 2001. Consumption, aggregate wealth, and expected stock returns. *Journal of Finance 56*, 815-849.

[33] Leeper, Eric M. 1992 Consumer attitudes: King for a day, *Federal Reserve Bank of Atlanta Economic Review* 77, 1-16.

[34] Ludvigson, S.C. 2004. Consumer confidence and consumer spending. *The Journal of Economic Perspectves* 18, 29-50.

[35] Mankiw, G. Shiller on animal spirits, *Greg Mankiw's Blog*, January 27, 2009

[36] Matsusaka, G. John, and Sbordone, M. Argia, 1995. Consumer confidence and economic fluctuations, *Economic Inquiry* 33, 296-318.

[37] Mishkin, F.S., 1978. Consumer sentiment and spending on durable goods. *Brooking Papers on Economic Activities,* 217-232.

[38] Newey, W. K., and West, K. D. 1987. Hypothesis testing with efficient method of moments estimation. *International Economic Review* 28, 777-787.

[39] Owyang, Michael T., Valerie A. Ramey, and Sarah Zubairy. 2013. Are government spending multipliers greater during periods of slack? Evidence from Twentieth-Century historical data. *American Economic Review* 103, 129-34.

[40] Petev, I., Pistaferri, L., and Saporta, I. 2012. Consumption and the Great Recession. The Russell Sage Foundation and the Stanford Center on Poverty and Inequality, 2-3.

[41] Ramey, V. A., and Zubairy, and. 2014. Government spending multipliers in good times and in bad: evidence from US historical data. *NBER working paper No.* 20719

[42] Sims, C. A. 1980. Macroeconomics and reality. *Econometrica* 48, 1-48.

[43] Sims, C. A. 1993. A 9 variable probabilistic macroeconomic forecasting model, in J. H. Stock and M. W. Watson (eds.) Business Cycles, Indicators and Forecasting, *NBER Studies in Business Cycles* 28, 179–214.

[44] Stock, James, and Mark Watson, 2007. Why has U.S. inflation become harder to forecast? *Journal of Money, Banking and Credit* 39, 3-33.

[45] Stock, J. H. and Watson, M. W. 1996. Evidence on structural instability in macroeconomic time series relations, *Journal of Business and Economic Statistics* 14, 11–30.

[46] Starr, M. A. 2012. Consumption, sentiment, and economic news. *Economic Inquiry* 50, 1097-1111.

[47] Thornton, D. L., and Batten, D. S. 1985. Lag-length selection and tests of Granger causality between money and income. *Journal of Money, Credit and Banking* 17, 164-178.

7 Appendix 1: Robustness exercises - Some alternative models (Not intended for publication)

Alternative measures of consumer confidence

In the baseline estimation, we use the confidence index, commonly known as C12M, which is based on consumers' expectations about the economic conditions over a twelve month horizon. However, measures over longer term horizons, such as C5Y, which measures consumer's expectations over a five year horizon, are also available as is the Index of Consumer Sentiment (ICS) which measures confidence based on an average of survey responses. Here we provide results using an alternative definition for x_t given by $x_t = [AltConf_t \quad c_t \quad y_t \quad f_t]'$, where $AltConf_t$ denotes a vector using either $C5Y_t$ or ICS_t.

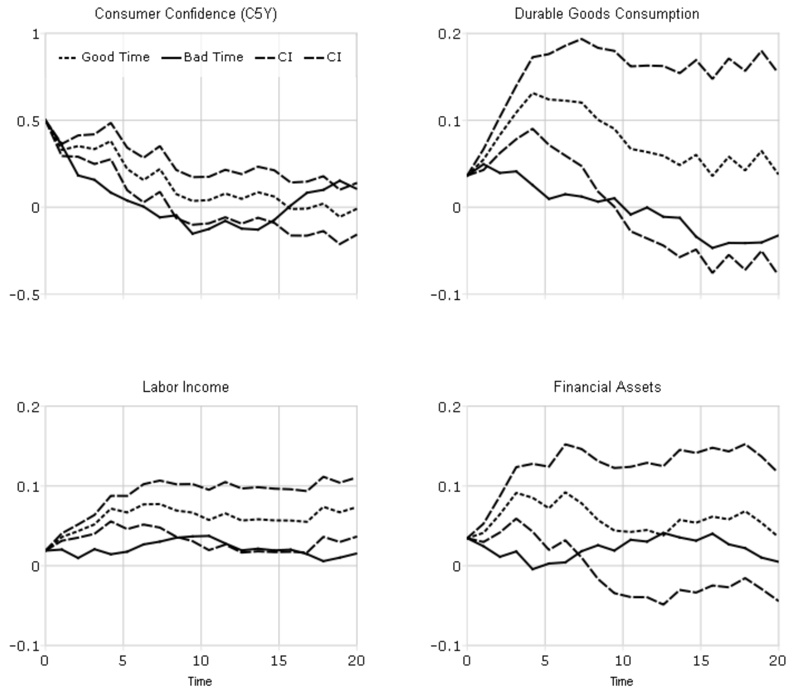

Figure 4a: Impulse responses confidence shocks measured by C5Y

Figures 4a and 4b show the results of this exercise for the durable goods measure for consumption.[24] They are largely consistent with the confidence impulses in Figure 3a, showing that the response of durable goods consumption is highly dependent on what state the economy is in. During good economic times, the impulse has favorable effects on durable goods, while during bad economic times, the impulse has very little effect.

[24] Note, both Figures 4a and 4b are analogous to Figure 3a and be do not present anologues to Figure 3b here to save space.

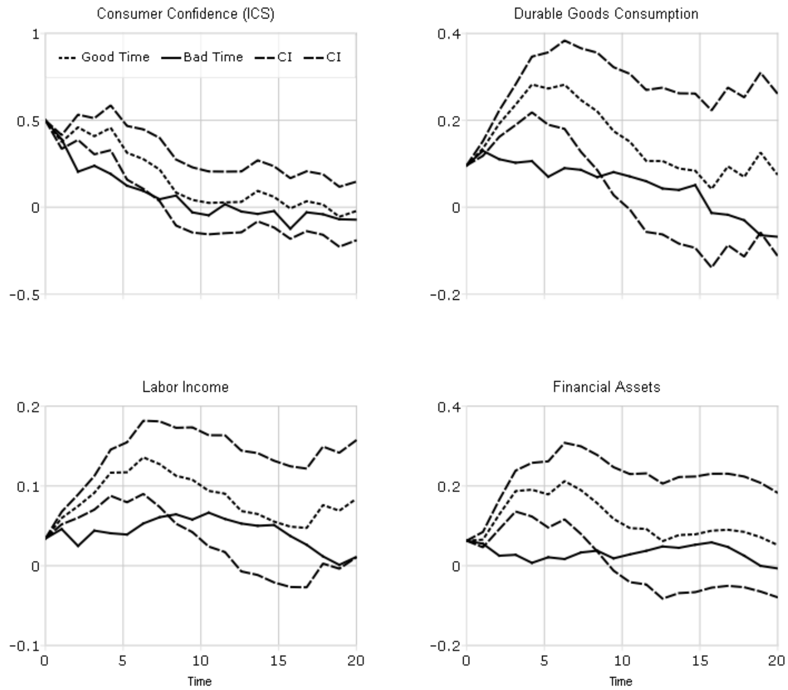

Figure 4b: Impulse responses confidence shocks measured by ICS

Sub-components of durable goods

In our baseline analysis we use aggregate spending on durable spending. Here we investigate two subcomponents, motor vehicles and other durable goods. Figures 5a and 5b show these results. Again they show that confidence produces an effect during good times, but during bad times the effect is hardly noticeable.

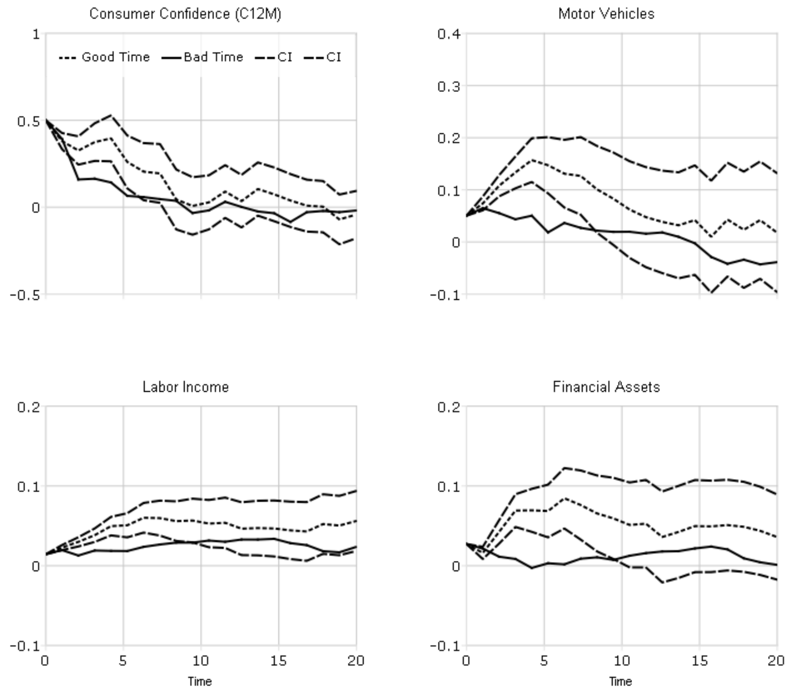

Figure 5a: Substituting motor vehicles for all durable goods

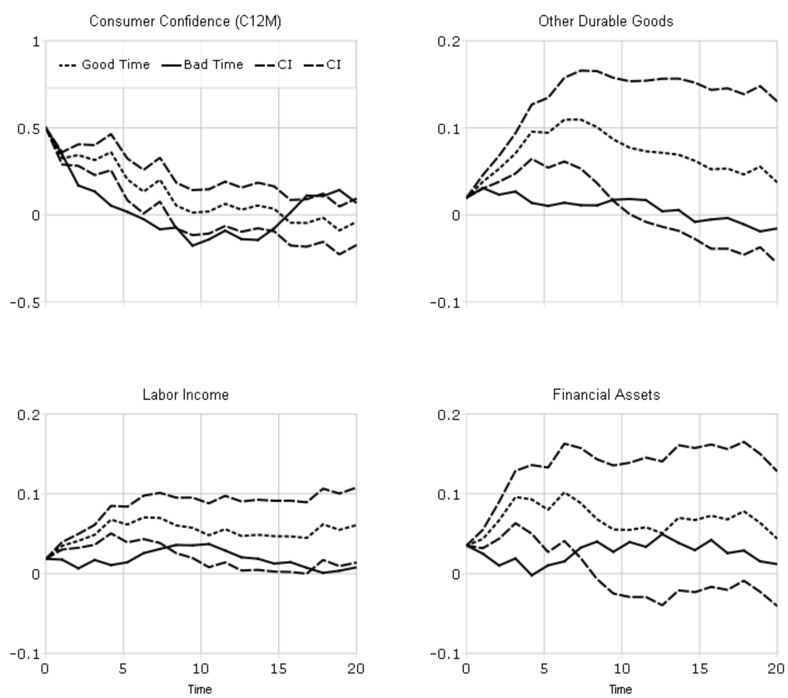

Figure 5b: Substituting other durable goods for all durable goods

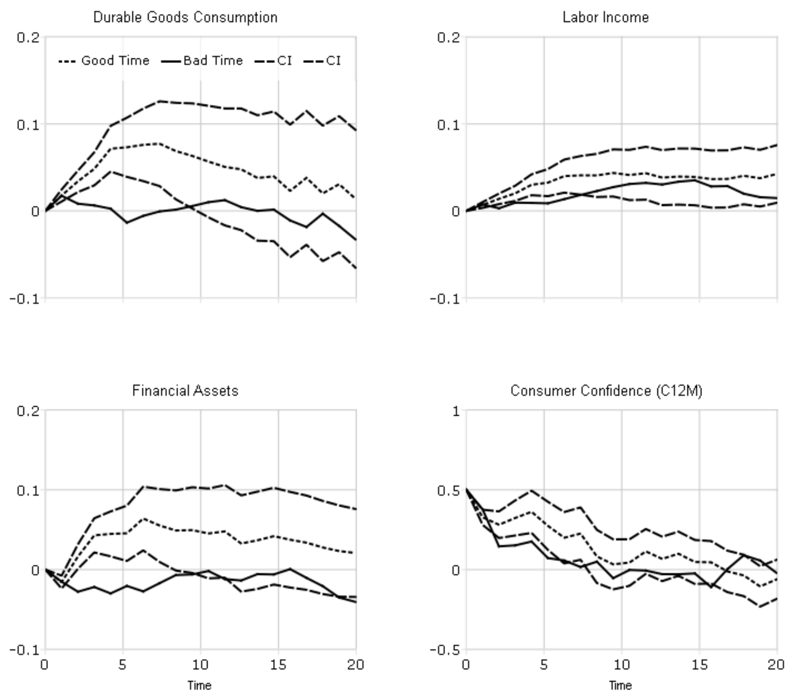

Figure 6: Impulse response when consumer confidence is ordered last

Alternative Cholesky ordering.

In the paper we followed the ordering used in Barsky and Sims (2012) which ordered confidence first and was motivated by the idea that exogenous news shocks cause the innovations in consumer confidence. However, this assumption is questionable as the consumer sentiment may vary in accordance with the personal labor income as well as holding of financial assets. To check the extent to which this ordering affect our results, we reorder the variables in the system as follows $x_t = [c_t \quad y_t \quad f_t \quad cc_t]'$ such that confidence is orthogonalized with respect to financial asset, durable spending and labor income i.e. all the variables in the system are forced to have a zero-on-impact reaction to confidence innovations. The outcomes of the alternative ordering are displayed in Fig 5e. With alternative ordering, the results are comparable to the baseline case. That is even after orthogonalization with respect to financial assets, durable spending and income households responses to confidence innovations are heterogenous across business cycle. However, the point estimates of impulse responses are slightly lower than in case with confidence ordering first. Our results with ordering confidence last hold when we use other measures of confidence indices as well as consider subcomponents of durable goods.

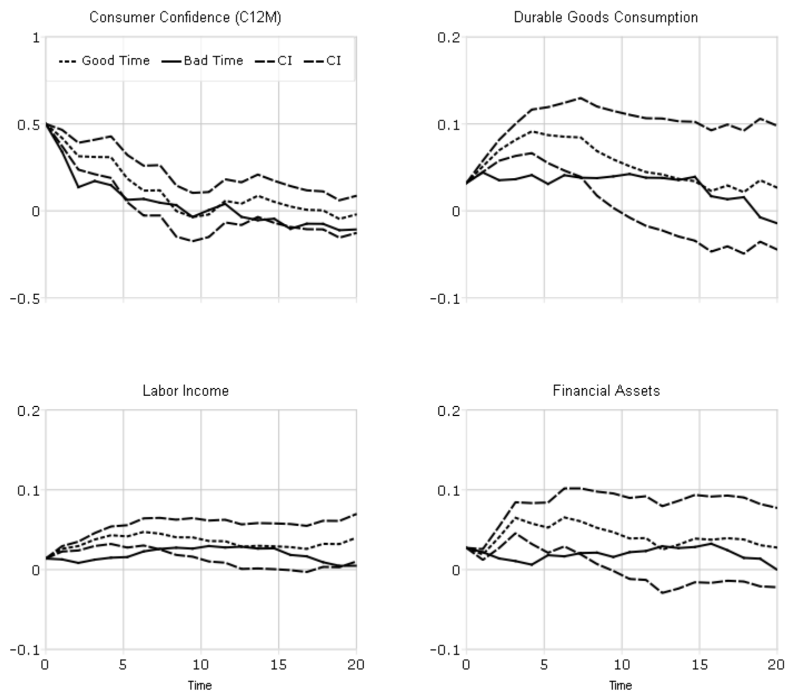

Figure 7: Impulse response in an endogenous threshold model

Endogenous thresholds

In the baseline formulation of the empirical model, we chose an exogenous threshold of unemployment rate as 6.5 percent. However, forcing the threshold exogenously could be a concern as we are not allowing the for the possibility of state-dependance that might arise only for a higher degree of slack in the economy. We use Chan's

(1993) methodology to find the endogenous threshold for the basic TAR model in (3). We find the threshold of unemployment rate at 6.7 percent that differentiates the two states of the economy. As shown in Fig.5f, our main results are still preserved where households'spending on durable goods to one-SD confidence shock vary across the state of the economy. We calculate the endogenous threshold of unemployment rate by considering different measures of confidence indices for different components of durable goods, our results still similar to those of the baseline outcomes. We did not report the impulse responses due to space constraint.

A one lag model

The Bayesian Information Criterion (BIC) implied that only one lag was nessary. As shown in figure 8, at times of slack, impulse responses from one S.D. confidence shock on its own as well as on durable goods consumption and labor income lie outside the confidence bands of normal time at all most entire periods.

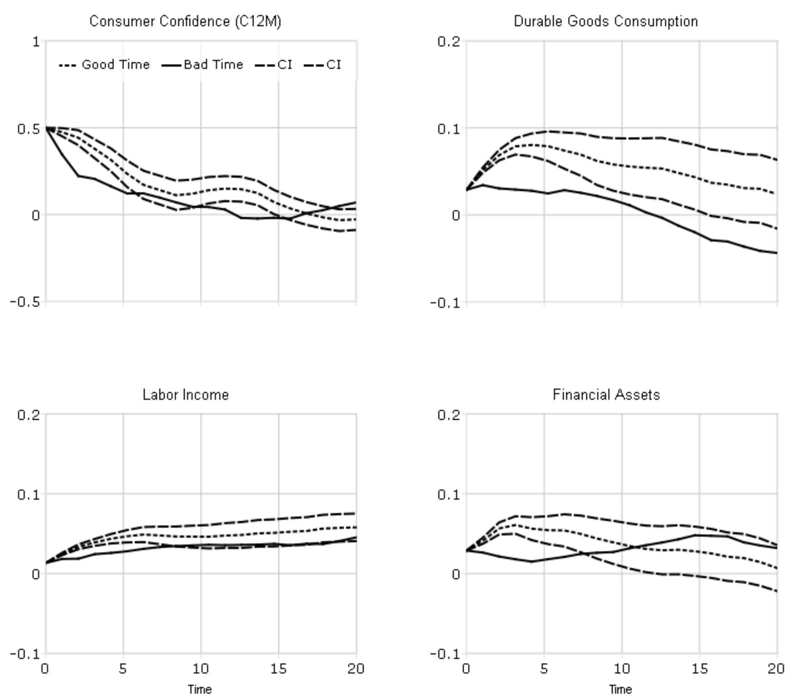

Figure 8: Impulse responses for a one lag model

Using total assets

In the baseline formulation of the empirical model, we chose only the financial assets. When we include both fianancial and non-financial assets (total assets) our results do not change much.

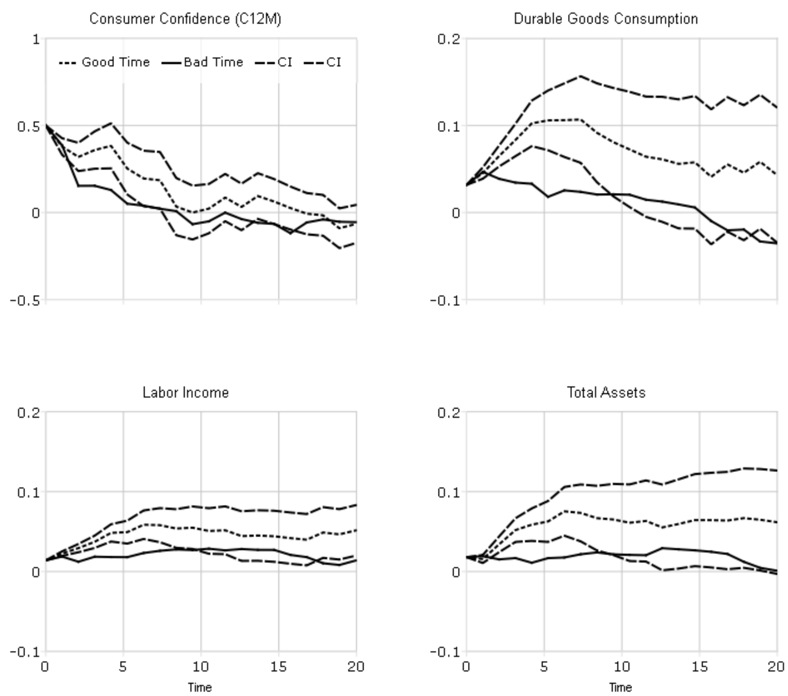

Figure 9: Impulse responses substituting total assets for assets