ESSAYS IN TIME SERIES ANALYSIS AND APPLIED MACROECONOMICS

A Dissertation

Presented to

The Faculty of the Department

of Economics

University of Houston

In Partial Fulfillment

Of the Requirements for the Degree of

Doctor of Philosophy

By

Bocong Du

December 2016

ESSAYS IN TIME SERIES ANALYSIS AND APPLIED MACROECONOMICS

Bocong Du

APPROVED:

Bent E. Sørensen, Ph.D.
Committee Chair

Christian J. Murray, Ph.D.

German Cubas, Ph.D.

Alex Nikolsko-Rzhevskyy, Ph.D.
Lehigh University
Department of Economics

Steven G. Craig, Ph.D.
Interim Dean, College of Liberal Arts and Social Sciences
Department of Economics

ESSAYS IN TIME SERIES ANALYSIS AND APPLIED MACROECONOMICS

An Abstract of a Dissertation

Presented to

The Faculty of the Department

of Economics

University of Houston

In Partial Fulfillment

Of the Requirements for the Degree of

Doctor of Philosophy

By

Bocong Du

December 2016

Abstract

This dissertation is composed of two essays on time series analysis and applied macroeconomics.

The first essay extends the Zero-Information-Limit Condition (ZILC) theory to the Generalized ZILC, and shows how the Generalized ZILC applies in the GARCH(1,1) model theoretically and empirically. In the theoretical part, under the Generalized ZILC theory proposed in the essay, the estimated information of the GARCH coefficient in the GARCH(1,1) model is overestimated; the estimated variance and estimated standard error of the GARCH coefficient is too small relative to the true value. Therefore, the actual size of the t-statistics of the GARCH estimate is too large. When sample size increases, this problem still exists. Because of the underestimated variance, it would be too often to reject the true null hypotheses. This essay proposes an empirical application strategy, by constructing the ZILC zone and safe zone for the GARCH(1,1) model. In the application part, this paper uses Value-at-Risk analysis in the risk management, to show that, if we fail to pay attention to the Generalized ZILC issue, the risk calculated by Value-at-Risk methodology using the GARCH(1,1) model, would be underestimated. At last, this paper proposes a Parametric Bootstrapping strategy, to generate a ratio and correct the underestimated variance of the GARCH coefficient in the GARCH(1,1) model.

In the second essay, I estimate the extent to which shocks to "animal spirits" can have an effect on real economic outcomes at business cycle frequencies. Recent advances in rational expectations models that formalize a role for animal spirits shocks (or "sentiments" shocks) motivate an empirical examination of this question. I use monthly data on consumer confidence and coincident economic activity indexes at

the level of U.S. states in a structural Vector AutoRegression (SVAR) model with long run restrictions to identify shocks to animal spirits and to economic fundamentals (which we refer to as "news" shocks). Specifically, I assume that animal spirits shocks cannot have an effect on the level of output in the long run. I find that, although most variation in the level of output (in the short run and in the long run) can be explained by innovations in news, animal spirits do have statistically and economically significant effects at business cycle frequencies. Two years after a positive innovation in animal spirits, the level of output is about three percent higher than it was before the shock. Significant effects can also be observed on retail sales, non-farm payrolls, the unemployment rate, and aggregate wages and salaries.

Acknowledgements

I consider myself extreme lucky to have Professor Bent Sorensen, Professor Christian Murray, and Professor German Cubas to serve as my committee, and guide me through my Ph.D. study. I cannot thank Professor Bent Sorensen enough for his exceptional advice, support, and encouragement for me during these years. I'm so lucky and it is such a great honor for me to have him as my academic advisor. I would like to thank Professor Christian Murray and Professor German Cubas for guiding me and supporting me in the research. Professor Christian Murray led me to the ZILC issue, so I can come up the first essay; Professor German Cubas taught me the advanced Macroeconomics topics and led me to the consumer confidence topic in the applied macroeconomics research. I would also like to thank Professor Alex Nikolsko-Rzhevskyy in Lehigh University. He gave me great comments in the conference where I presented my paper. I thank him to be one of my committee members and guided me in my research.

It would not have been as great for the life at UH had it not been with my fantastic friends. I want to thank Yi Zhu and Hui Du for their guidance, support, and endless love for me. They are my family in U.S.. I want to thank Chon-Kit Ao and Somdeep Chatterjee for being such great friends for me.

Finally, I want to thank my parents Zetian Du and Xiuhua Liu, who deserve my deepest gratitude for their never-ending belief and love.

to my Dad and Mom

Contents

1 Spurious Inference in the GARCH(1,1) Model when Generalized ZILC Holds: Theory and Application **1**

1.1 Introduction 1

1.2 The Generalized ZILC in the GARCH(1,1) Model 6

1.2.1 Zero-Information-Limit Condition (ZILC) 6

1.2.2 Generalized ZILC in the GARCH(1,1) Model 7

1.3 Evidence of Spurious Inference from Monte Carlo Experiments 10

1.3.1 Inference when there is no GARCH effect 11

1.3.2 Inference when there is moderate GARCH effect 12

1.3.3 ZILC Zone and Safe Zone in the GARCH(1,1) Model 13

1.4 Application: Distorted t-statistics 14

1.4.1 Different Tests Compare 14

1.4.2 Strategy on the Significance of β 15

1.5 Application: Distorted Variance 16

1.5.1 Value-at-Risk Analysis in GARCH(1,1) Model 16

1.5.2 Parametric Bootstrapping 17

1.6 Conclusion 18

1.7 Appendix 20

1.7.1 The Closed Form Information Matrix 20

1.7.2 The Asymptotic Variance of β 21

2 Do Shocks to Animal Spirits Cause Output Fluctuations? 28

2.1 Introduction 28

2.2 Data 32

2.2.1 Consumer Confidence 32

2.2.2 State Coincident Indexes 34

2.2.3 Retail Sales 36

2.3 Econometric Methodology: Structural VAR 37

2.4 Results 45

2.5 Conclusion 53

List of Figures

1.1 Frequency to Reject t-test for β: no GARCH effect ($\beta = 0$) 24
1.2 Frequency to Reject t-test for β: moderate GARCH effect ($\beta = 0.5$) . 25
1.3 The Actual Size of the t-test ZILC Zone and Safe Zone in the GARCH(1,1) Model 26

2.1 Consumer Confidence 58
2.2 State Economic Indexes 59
2.3 Consumer Confidence and Economic Activity Growth in Connecticut 60
2.4 State Retail Sales 61
2.5 Impulse Responses of Output and Consumer Confidence 62
2.6 Impulse Responses of Nonfarm Payrolls 63
2.7 Impulse Responses of Unemployment Rate 64
2.8 Impulse Responses of Average Weekly Hours of Manufacturing Workers 65
2.9 Impulse Responses of Real Wages and Salaries 66
2.10 Impulse Responses of Real Wages and Salaries per Worker 67
2.11 Impulse Responses of Retail Sales 68

List of Tables

1.1 Size of t-test for β at 5% Level 22

1.2 Size of Various Tests for β at 5% in GARCH(1,1) Model when $\beta = 0$ 23

1.3 Size of Various Tests for β at 5% in GARCH(1,1) Model when $\beta = 0.5$ 27

2.1 Conference Board Categorization of States into Regions 55

2.2 Summary Statistics of Key Variables 56

2.3 Variance Decomposition of Effects on State Coincident Economic Index and Consumer Confidence 57

Chapter 1

Spurious Inference in the GARCH(1,1) Model when Generalized ZILC Holds: Theory and Application

1.1 Introduction

The GARCH(1,1) model is popular in modeling the changing-volatility of the time series data. It has become one of the benchmarks in modeling time-varying volatility since introduced by Bollerslev (1986). Capturing time-varying volatility is one of the most important issues in modeling time series. The GARCH(1,1) model is originally proposed as ARCH(AutoRegressive Conditional Heteroskedasticity) in 1982 by Robert Engle. The ARCH/GARCH family has been widely used to extract a potential volatility-changing process. Bera and Higgins (1993) remarked that "a major contribution of the ARCH literature is the finding that apparent changes in the volatility of economic time series may be predictable and result from a specific type of nonlinear dependent rather than exogenous structural changes in variables."

Zero-Information-Limit Condition (ZILC) is first proposed by Nelson and Startz (2007). Their paper has shown that when identification of the target parameter is conditional on another structural parameters in the model, the inference for the target parameters will be misleading by the realization of the structural parameters, if ZILC holds in the structure of the information matrix. In models when ZILC applies, the standard errors tend to be underestimated when the identifying structural parameters are close to some specific values, no matter how large the sample size would be. Examples include the Weak Instrument problem, ARMA(1,1) models, and some certain nonlinear regression models. Ma, Nelson, and Startz (2007) shows ZILC also holds in the GARCH(1,1) model.

To illustrate how the ZILC applies to the ARCH-GARCH family model, I propose a simulation-based experiment as motivation. I first set up an ARCH(1) model:

$$\epsilon_t \sim i.i.d. N(0, \sigma_t^2) \tag{1.1}$$

$$\sigma_t^2 = \omega + \alpha \cdot \epsilon_{t-1}^2 \tag{1.2}$$

Note that σ^2 is the conditional variance and is driven by past realizations of ϵ. α is called the ARCH coefficient.

The typical GARCH(1,1) model is written as:

$$\epsilon_t \sim i.i.d.N(0, \sigma_t^2) \tag{1.3}$$

$$\sigma_t^2 = \omega + \alpha \cdot \epsilon_{t-1}^2 + \beta \cdot \sigma_{t-1}^2 \tag{1.4}$$

Note that σ^2 is still the conditional variance and is driven by past realizations of ϵ, but also with added persistence determined by β. Here, α is called ARCH coefficient, and β is called GARCH coefficient. In the case $\beta = 0$, the model reduces to the pure ARCH(1) model, and in the case $\alpha = 0$, the GARCH effect β cannot be identified.

The GARCH coefficient β represents the relationship between the current volatility with the past volatility. Intuitively, the GARCH model depicts the volatility clustering phenomenon, which implies "the large changes tend to be followed by large changes, and small changes tend to be followed by small changes." This phenomenon exists among many time series data, especially for stock price, exchange rates, and oil price. That is why GARCH model is very popular and GARCH coefficient β is on the focus of the model research.

The purpose of the experiment is to investigate the spurious inference of GARCH estimate β in the GARCH(1,1) model. I implement a series of Monte Carlo experiments. In the sequence of Monte Carlo experiments, I simulate the data from ARCH(1) process defined by equation (1) and (2) with three sets of parameters values:

$$\begin{pmatrix} \omega \\ \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 1 \\ 0.01 \\ 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0.05 \\ 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0.1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0.9 \\ 0 \end{pmatrix}$$

Then I use the ARCH(1) simulated data and GARCH(1,1) model setting to estimate the GARCH coefficient β. The estimated variance of β will be calculated, then the t-statistics of the β. Run the Monte Carlo simulation for 1000 times, and record the chances that the t-test of β rejects the true null hypothesis of $\beta = 0$. Table 1 gives the empirical sizes of t-test at the nominal 5% level for GARCH coefficient β.

The result is very interesting. The data is generated from ARCH(1) model (which the GARCH coefficient $\beta = 0$). If I estimate β from the ARCH(1) simulated data, the probability of rejecting the null hypothesis of $\beta = 0$ should be around the nominal 5%. However, the real Monte Carlo experiment shows, we would have around 50% chance that we falsely reject the null hypothesis of $\beta = 0$, and falsely reckon that the data fits GARCH(1,1) model setting. This also implies that, we might falsely reckon the data contains the significant clustering volatility phenomena. In Table 1.1, even the sample size is large as increased to 5000, the spurious inference issue still exists: the estimated variance of the GARCH coefficient β is underestimated, and the t-statistics of β is overestimated. The explanation of this phenomena is ZILC.

Ma, Nelson, and Startz (2007) first proves that ZILC applies in the GARCH(1,1) model. In this paper, I extend the ZILC theory to the Generalized ZILC theory, and depict the ZILC zone by constructing a complete Monte Carlo Simulation in the GARCH(1,1) model. In the ZILC zone, the estimated standard errors of β are too small; as a result, the actual size of the t-test for the GARCH coefficient is far too large, which leads the rejection of the true null hypotheses occurring too often. Thus, researchers unaware of this spurious effect may be tempted to conclude that the persistence due to the GARCH effect is significant, while in fact the persistence of the volatility is absent. For people who care about the significance of the GARCH coefficient, this paper provides a test strategy for empirical application. In addition, for people who aim to use estimated variance of the GARCH coefficient for further empirical application, I propose a simulation-based strategy to correct the under-estimated standard error, and eventually resolve the Generalized ZILC issue. This solution can be used in the Value-at-Risk analysis in the risk management. Value-at-Risk analysis uses estimated variance of the target parameter to construct the Value-at-Risk value. In my paper, I propose an investment hedging strategy with a virtual data, to show that how the spurious inference led by Generalized ZILC would influence the Value-at-Risk analysis.

This essay is organized in the following way: Section 2 introduces what is the Generalized ZILC, and how the Generalized ZILC applies in the GARCH(1,1) model theoretically. Section 3 presents evidence by Monte Carlo experiments, to elaborate what is the actual size of the t-test for the GARCH coefficient, and how the Generalized ZILC distorts the t-test results. Section 3 also illustrates the ZILC zone and safe zone of the estimated α and β in the GARCH(1,1) model. Section 4 presents various tests under Generalized ZILC issue, and proposes a test strategy. This strategy can be applied for research interests on the significance of GARCH coefficient β. Section 5 design a volatility hedging investment hedging strategy, and show how the Generalized ZILC applies in the Value-at-Risk analysis in the risk management. Section 5 proposes the solution for ZILC: a simulation-based ratio to correct the underestimated variance. Section 6 concludes this essay.

1.2 The Generalized ZILC in the GARCH(1,1) Model

1.2.1 Zero-Information-Limit Condition (ZILC)

ZILC identifies how the weak identification issue leads to the spurious inference on the target parameter estimates. In many Econometrics models, the asymptotic variance of a parameter estimate depends on the value of other structural parameters. If the data contains little information about the target parameter when the structural parameters are close to some critical values, Nelson and Startz (2007) identified this phenomena as ZILC.

Consider an Econometrics model with parameters γ and β. β is of the interest for hypothesis testing. Sample size is T, and exogenous data is X. $V_{\hat{\beta}}(\beta, \gamma, X)$ is the asymptotic variance of β. The estimated information matrix of β is $I_{\hat{\beta}}(\beta, \gamma, T, X) = [V_{\hat{\beta}}(\beta, \gamma, T, X)]^{-1} \geq 0$.

When ZILC holds:

$$\lim_{\gamma \to \gamma_0} I_{\hat{\beta}}(\beta, \gamma, T, X) = 0 \tag{1.5}$$

Thus, we can get:

$$\lim_{\gamma \to \gamma_0} \left[Pr \left(\frac{I_{\hat{\beta}}(\hat{\beta}, \hat{\gamma}, T, X)}{I_{\hat{\beta}}(\beta, \gamma, T, X)} \right) > M \right] = 1, \ \forall M \tag{1.6}$$

The t-statistic is $t_{\hat{\beta}}^2 = (\hat{\beta} - \beta_0)^2 \cdot I_{\hat{\beta}}(\hat{\beta}, \hat{\gamma}, T, X)$. The implication of the equation is, when γ is close to a critical value, true information of β is close to 0. Under this scenario, the estimated information of β will be always larger than the asymptotic true variance of β. Information is overestimated; variance would be underestimated; standard error would be underestimated; thus, t-statistics will be overestimated. Spurious inference occurs.

1.2.2 Generalized ZILC in the GARCH(1,1) Model

Nelson and Startz (2007) illustrates that under ZILC, the standard errors tend to be underestimated; the size of the asymptotic t-test is distorted (either be oversized or undersized). Ma, Nelson, and Startz (2007) shows that ZILC applies to the GARCH(1,1) model when ARCH coefficient α goes to 0. The estimated standard error of GARCH coefficient β is underestimated; the t-test is oversized.

To illustrate how ZILC applies to the GARCH(1,1) Model, I rewrite the archetypal GARCH(1,1) model as:

$$\epsilon_t = \sqrt{\sigma_t^2} \cdot \phi_t \tag{1.7}$$

$$\sigma_t^2 = \omega + \alpha \cdot \epsilon_{t-1}^2 + \beta \cdot \sigma_{t-1}^2 \tag{1.8}$$

Following the previous discussion, σ^2 is the conditional variance and is driven by past realizations of ϵ, with added persistence determined by β. Here, the mean of equation (7) is set to be zero without loss of generality (see Bollerslev (1986)). ϕ_t is independently and identically distributed with zero mean and unit variance, i.e., $i.i.d.(0, 1)$, with finite higher moments (see Lumsdaine(1996) for details).

Write up the log-likelihood function:

$$L_T(\theta) = T^{-1} \cdot \sum_{t=1}^{T} l_t(\theta) \tag{1.9}$$

$$\sum_{t=1}^{T} l_t(\theta) = -\frac{1}{2}\ \log 2\pi - -\frac{1}{2}\ \log h_t - \frac{1}{2}\frac{\epsilon_t^2}{h_t} \tag{1.10}$$

Where, $\theta = (\omega, \alpha, \beta)'$ and $\hat{\theta_T}$ maximizes the quasi log-likelihood function for a given sample data ϵ_1,ϵ_2, ϵ_3, ... ϵ_T, and therefore is the QMLE. Fiorentini, Calzolari, and Panattoni (1996) derives the first and second derivatives in GARCH(1,1) models. Ma (2008) takes a local approximation of each element in the neighborhood of small α to avoid taking the expectation of a non-linear form. Based upon Ma (2008)'s results, it is straight forward to show that the inverse of the asymptotic variance of

$\hat{\beta}$, the information measure of Nelson and Startz (2007), goes to 0 as α approaches 0, i.e., ZILC holds:

$$\lim_{\alpha \to 0} I_{\hat{\beta}}(\omega, \alpha, \beta) = 0 \tag{1.11}$$

Appendix gives a formal proof of (1.11) based upon Ma (2008)'s analytic results.

Ma, Nelson, and Startz (2007) proves that when α is close to 0, information of β is close to 0. In this paper, as proved in Appendix, when β goes to 0, the information of β will also go to 0, which is presented as:

$$\lim_{\beta \to 0} I_{\hat{\beta}}(\omega, \alpha, \beta) = 0 \tag{1.12}$$

Appendix gives a formal proof of (1.12). Furthermore, the following Monte Carlo experiments investigate that the spurious inference occurs when both α and β go 0. That is how this paper extends ZILC to the Generalized ZILC in the GARCH(1,1) model.

In the GARCH(1,1) model, the estimated variance of $\hat{\beta}$ is $V_{\hat{\beta}}(\omega, \alpha, \beta, T, X)$. The inverse of the variance is a natural measure of information associated with $\hat{\beta}$, so the estimated information of $\hat{\beta}$ is $I_{\hat{\beta}}(\omega, \alpha, \beta, T, X) = V_{\hat{\beta}}^{-1}(\omega, \alpha, \beta, T, X)$. Under the Generalized ZILC issue, either α or β close to 0 would both lead the information measure of β close to 0. The estimated information measure will be always larger than the true information measure. Proved by the series of Monte Carlo experiments, the estimated variance of the GARCH estimate β is too small under both of the scenarios; t-statistics of β is overestimated.

Generally speaking, consider a model with scalar parameters γ. β is the parameter of interest for hypothesis testing. The asymptotic variance of estimator $\hat{\beta}$ is assumed to have a representation as a function of β, other parameters γ, and exogenous data X. The Generalized ZILC issue can be illustrated as:

Definition: The Generalized Zero-Information-Limit Condition (Generalized ZILC) holds for estimator $\hat{\beta}$ if any value in the domain Θ makes the following equation: [1]

$$\lim_{\Theta} I_{\hat{\beta}}(\beta, \gamma, X) = 0 \tag{1.13}$$

So

$$\lim_{\Theta} \left[Pr\left(\frac{I_{\hat{\beta}}(\hat{\beta}, \hat{\gamma}, X)}{I_{\hat{\beta}}(\beta, \gamma, X)} \right) > M \right] = 1,\ \forall M \tag{1.14}$$

1.3 Evidence of Spurious Inference from Monte Carlo Experiments

In the GARCH(1,1) model, the problem of Generalized ZILC can be represented as:

[1] The domain Θ represents any values to make $I_{\hat{\beta}}(\beta, \gamma, X)$ goes to 0, even including the condition of β itself.

$$\lim_{\Theta}\left[Pr\left(\frac{I_{\hat{\beta}}(\hat{\omega},\hat{\alpha},\hat{\beta},T,X)}{I_{\hat{\beta}}(\hat{\omega},\hat{\alpha},\hat{\beta},T,X)}\right)>M\right]=1,\forall M \tag{1.15}$$

Estimated information of β will be always larger than the true information of β under Generalized ZILC. Since $V_{\hat{\beta}}^{-1}=I_{\hat{\beta}}$, estimated variance will be underestimated, and $t_{\hat{\beta}}^2=(\hat{\beta}-\beta_0)^2\cdot V_{\hat{\beta}}^{-1}$ will be oversized.

A series of Monte Carlo experiments is implemented to investigate how worse the spurious inference occurs when the GARCH(1,1) model is under the Generalized ZILC issue.

1.3.1 Inference when there is no GARCH effect

In this sequence of Monte Carlo experiments, data is simulated from the GARCH(1,1) defined by equation (1.7) and (1.8) with the following sets of parameter values:

$$\begin{pmatrix}\omega\\ \alpha\\ \beta\end{pmatrix}=\begin{pmatrix}1\\ 0.01\\ 0\end{pmatrix},\begin{pmatrix}1\\ 0.02\\ 0\end{pmatrix}\cdots\begin{pmatrix}1\\ 0.1\\ 0\end{pmatrix},\begin{pmatrix}1\\ 0.2\\ 0\end{pmatrix}\cdots\begin{pmatrix}1\\ 0.9\\ 0\end{pmatrix}$$

Ma (2008) sets the local approximation for each element in the neighborhood of small α to avoid taking the expectation of non-linear form. This paper extends the Monte Carlo experiment range to a complete area of α and β. Note that in the GARCH(1,1) model $\alpha+\beta$ should be smaller than 1.

Since β is 0, there is no GARCH effect and the process is actually an ARCH(1) process. The scale parameter ω is normalized to be unity. For each set of parameter values, I set up the sample size $T = 1000$ to be consistent with the experiments in Nelson and Startz (2007), and Ma, Nelson, and Startz (2007). For all experiments, 1000 simulated paths of sample data of length T are generated. Figure 1.1 gives the empirical size of t-test at the nominal 5% level for estimated parameter $\hat{\beta}$.

Ma, Nelson, and Startz (2007) shows the similar results that when ARCH coefficient α is close to 0, the actual size of t-test for β is around 50% even for a large sample size as shown in Table 1. Their paper also mentions that, for sufficiently large α, and for sufficiently large sample size, the size distortion of β is greatly reduced. As Figure 1.1 shows, the size distortion of β is reduced as α increases at the beginning. However, when α continues to become larger, the actual size of β goes worse. This is due to the Generalized ZILC proved in Section 2: in the GARCH(1,1) model, when β is small, the estimated information of β is always close to 0, no matter how the α changes.

1.3.2 Inference when there is moderate GARCH effect

Ma, Nelson, and Startz (2007) also proposes the moderate GARCH effect Monte Carlo experiment. It is important to note that ZILC holds whenever the ARCH coefficient α is small, regardless of the magnitude of true β. In the moderate GARCH model, the β is set as 0.5. Here I extend the Monte Carlo experiment to the complete rang of α and β, which β is set as 0.5, and α is increased from 0.01 to 0.49 ($\alpha+\beta<1$). Data is simulated from the GARCH(1,1) process defined by the following equation:

$$\begin{pmatrix} \omega \\ \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 1 \\ 0.01 \\ 0.5 \end{pmatrix} \cdots \begin{pmatrix} 1 \\ 0.1 \\ 0.5 \end{pmatrix}, \begin{pmatrix} 1 \\ 0.2 \\ 0.5 \end{pmatrix} \cdots \begin{pmatrix} 1 \\ 0.49 \\ 0.5 \end{pmatrix}$$

The sample size is fixed at 1000 and the number of simulation is also 1000. Graph 2 presents the empirical size of t-test of β at the nominal 5% level.

As proved in Section 2, for the moderate GARCH(1,1) model, when α is small, ZILC applies to the GARCH(1,1) model as the estimated information matrix of β is close to 0. So the estimated standard error of β is underestimated, t-statistics is overestimated. Figure 1.2 shows the actual size of t-test for β is around 40% when α is small in the moderate GARCH(1,1) model, while the t-test goes back to the normal size level when α increases, which implies the ZILC issue goes away.

1.3.3 ZILC Zone and Safe Zone in the GARCH(1,1) Model

As Section 2 proves, and Figure 1.1 and Figure 1.2 show, ZILC issue suffers either α is small or β is small, which is the Generalized ZILC. Thus, information of β close to 0 leads the estimated information of β would be always larger than the true information of β. Therefore, estimated standard error of β is underestimated, t-statistics of β is overestimated, actual size of the t-test for β is too large. When α and β increase, estimated information of β does not suffer the ZILC issue. It would be interesting to know that, regarding the estimated α and β, under what range, the t-test of β is not distorted. Figure 1.3 shows the ZILC zone and safe zone of the

GARCH(1,1) model. To get Figure 1.3, I simulate the data from the GARCH(1,1) process starting from $\alpha = 0.01$ and $\beta = 0$, run the Monte Carlo simulation for 1000 times to calculate the actual size of t-test of β. Then I increase α and β each with 0.01 value, run the Monte Carlo simulation again, and record the actual size of t-test of β. When the actual size of the GARCH coefficient β falls into the nominal 5%, the estimated α and β combine a boundary, which forms the ZILC zone and safe zone, as shown in Figure 1.3.

Figure 1.3 provides a useful way for future empirical work. If a series of data is estimated in GARCH(1,1) model, and the estimated α and β fall into the safe zone, then we can trust the t-test for β. However, if the estimated α and β fall into the ZILC zone, we should be careful of using t-statistics and the estimated standard error of β, since the estimated standard error of β might be underestimated and therefore the t-statistics might be oversized.

1.4 Application: Distorted t-statistics

1.4.1 Different Tests Compare

Assume research focus of empirical work is on the significance of β, t-test is distorted and the standard error is underestimated due to the overestimated information of β under Generalized ZILC. Although t-test is distorted, Likelihood Ratio(LR) test and lagrange Multiplier(LM) test do not suffer from ZILC. This will be proved by the Monte Carlo experiment with the same modeling setting in Section 3.1 and Section 3.2. Table 1.2 gives the empirical sizes of LR test and LM test when the model is setting as no GARCH effect.

It's good for practitioners that LR test and LM test perform much better than the t-test for model with no GARCH effect. The better performance of the LR test and LM test can be traced to the fact that both of these two tests are calculated without the restriction on the weakly identified parameters; see Zivot, Startz, and Nelson (1998) for discussion of this in the weak instrument case.

Table 1.3 also proves the LR and LM tests perform better than t-test when there is moderate GARCH effect. I run the complete Monte Carlo experiment under the domain with $\alpha >0$, $\beta >0$, and $\alpha + \beta<1$. The conclusion is, LR test and LM test perform the actual size of the test, which proves LR test and LM test are valid no matter the ARCH and GARCH estimates fall into ZILC zone or not.

1.4.2 Strategy on the Significance of β

Based on the analysis in Section 1.4.1, an empirical strategy could be proposed when the application focuses on the significance of β:

- ZILC holds in GARCH(1,1) model when $\alpha \to 0$ and $\beta \to 0$. t-test is distorted (t-statistics is overestimated) proved by a series of Monte Carlo experiments. So when we use GARCH(1,1) to model the data, we should be careful about the t-test of β.

- LR test and LM test are not distorted, so a practice strategy is proposed as:

  - If t-test rejects the null, but LR test and LM test don't reject null, we go for LR and LM test and don't reject null.

– If t-test, LR test, and LM test all reject the null, then we reject the null.

1.5 Application: Distorted Variance

If the empirical work is emphasized on the estimated variance of β, I propose a Monte Carlo based experiment to correct the underestimated variance. The methodology is Parametric Bootstrapping. This methodology is followed by an empirical application of Value-at-Risk analysis in the risk management.

1.5.1 Value-at-Risk Analysis in GARCH(1,1) Model

Value-at-Risk (VaR) is a statistical technique term in risk management. It is used to measure and quantify the level of financial risk within a firm or investment portfolio over a specific time frame. Value-at-Risk is measured through three fundamental variables: the amount of potential loss, the probability of that amount of loss, and the time frame. For example, a financial firm may determine that it has a 1% one month value at risk of $100 million. This means there is a 1% chance that the firm could lose more than $100 million in any given month. So the Value-at-Risk for 1% and one month is $100 million.

Assume an investment strategy to hedge the persistence volatility index β constructed by the GARCH(1,1) model. The benchmark model is MA(1)-GARCH(1,1), and estimated daily Value-at-Risk is at 99% confidence level. $\hat{\omega} = 1.061$, $\hat{\alpha} = 0.077$ and $\hat{\beta} = 0.773$. Generalized ZILC applies in this case as estimated α is close to 0.

The estimated variance of β is 0.05246. If the capital exposure for this investment strategy is \$10 million, within one day, there is 1% chance that the bank would lose more than $0.05246 \times \$10\text{m} \times 2.2 = \$1.16\text{m}$. Here, 2.2 is calculated by the Quantile Loss Function proposed by Angelidis, Benos, and Degiannakis(2004). Because of Generalized ZILC, estimated variance of *beta* is underestimated; the Value-at-Risk is underestimated. The risk of this strategy tends to be underestimated.

1.5.2 Parametric Bootstrapping

The following steps propose how a ratio is calculated to correct the underestimated variance based on the Parametric Bootstrapping methodology:

- Set a sequence of Monte Carlo experiment that $\omega = 1.061$, $\alpha = 0.077$ and $\beta = 0.773$. Also set the sample size T = 1000, and 1000 simulated paths of sample data of length T are generated.
- For each of the simulated sample data, use GARCH(1,1) model to estimate and record the 1000 of the estimated variance of β.
- The asymptotic variance of β can be evaluated numerically through the equation in Appendix.
- Take the average of the 1000 estimated variance of β, and divide the value of the numerical asymptotic variance of β. The ratio is 7.1.

This ratio is the corrected ratio for the GARCH(1,1) model when $\omega = 1.061$,

$\alpha = 0.077$ and $\beta = 0.773$. The underestimated variance of β is 0.05246. By multiplying the corrected ratio 7.1, the corrected estimated variance of $\hat{\beta}$ is $0.05246 \times 7.1 = 0.372$. Therefore, the corrected Value-at-Risk is $0.372 \times \$10\text{m} \times 2.2 = \$8.19\text{m}$. Previously, the Value-at-Risk is \$1.16m, which means there is a 1% that this investment strategy could lose \$1.16m in any given business day. However, because of the Generalized ZILC, the variance is underestimated, thus the risk is underestimated. By conducting the parametric bootstrapping, the Monte Carlo based corrected ratio is generated. By multiplying the corrected variance ratio, the amount that under 1% this investment strategy might lose in any given business day could increase from \$1.16m to \$8.19m. We might underestimate the risk if we fail to consider Generalized ZILC in the GARCH(1,1) model in the application of the Value-at-Risk analysis.

1.6 Conclusion

This paper extends the ZILC issue in GARCH(1,1) model. This paper shows that Generalized ZILC holds in the GARCH(1,1) model. ZILC is formulated by Nelson and Startz(2007). Ma, Nelson, and Startz (2007) proves when α is close to 0, ZILC holds in the GARCH(1,1) model. By constructing a seris of Monte Carlo simulation, this paper proves when β is close to 0, ZILC still applies in the GARCH(1,1) model. Therefore, the Generalized ZILC theory is proposed. This paper depicts the ZILC zone and safe zone for detecting ZILC issue, and illustrates two empirical strategies to deal with the Generalized ZILC issue. One strategy is the test choice when researchers focus on the significance of β. The other strategy is to generate a Monte Carlo based ratio by parametric bootstrapping, to correct the underestimated variance of β. This strategy proves its value in the Value-at-Risk analysis, since the risk would be underestimated if we failed to consider the Generalized ZILC issue in the GARCH(1,1) model application.

1.7 Appendix

1.7.1 The Closed Form Information Matrix

The information matrix of the GARCH(1,1) model given by Ma (2008):

$$I = \begin{pmatrix} A & B & B \\ B & C & D \\ B & D & E \end{pmatrix}$$

Where A, B, C, D, E are all the functions of estimates ω, α, and β.

$A = \frac{(1-\alpha-\beta)^2}{2\omega^2(1-\beta)^2}$,

$B = \frac{1-\alpha-\beta}{2\omega(1-\beta)^2}$,

$C = \frac{1-\alpha-\beta}{2(1-2\alpha\beta-\beta^2)} \cdot [\frac{3(1+\alpha+\beta)}{1-3\alpha^2-2\alpha\beta-\beta^2} + \frac{2\beta}{(1-\beta)^2}]$

$D = \frac{(1+\alpha+\beta)\cdot(1-\alpha-\beta)}{2(1-3\alpha^2-2\alpha\beta-\beta^2)\cdot(1-\beta^2)}(\frac{1}{1-\alpha\beta-\beta^2} + \frac{3\alpha\beta}{1-2\alpha\beta-\beta^2}) + \frac{\beta}{2(1-\beta^2)}(\frac{2}{1-\beta} - \frac{\alpha+\beta}{1-\alpha\beta-\beta^2} - \frac{\alpha}{1-2\alpha\beta-\beta^2})$

$E = \frac{1-\alpha-\beta}{2(1-\beta^2)\cdot(1-\alpha\beta-\beta^2)}[\frac{(1+\alpha\beta+\beta^2)\cdot(1+\alpha+\beta)}{1-3\alpha^2-\alpha\beta-\beta^2} + \frac{2\beta}{1-\beta}]$

The information measure for $\hat{\beta}$ is the inverse of $\hat{\beta}$'s variance, which is:

$$I_{\hat{\beta}} = \frac{T}{I^{-1}(3,3)} = T \cdot \frac{[1-(\alpha+\beta)]^2}{2\omega^2} \cdot \frac{B^2(2D-C-E)+A(CE-D^2)}{AC-B^2}$$

Ma, Nelson, and Startz (2007) proves that,

$T \cdot \frac{(1-\alpha-\beta)^2}{2\omega^2} \neq 0 \quad as \quad \alpha \to 0$

$$AC - B^2 \neq 0 \quad as \quad \alpha \to 0$$

$$B^2(2D - C - E) + A(CE - D^2) \;\to\; 0 \quad as \quad \alpha \to 0$$

This completes the equation (1.11):

$$\lim_{\alpha \to 0} I_{\hat{\beta}}(\omega, \alpha, \beta) = 0$$

In addition,

$$B^2(2D - C - E) + A(CE - D^2) \;\to\; 0 \quad as \quad \beta \to 0$$

Which leads,

$$\lim_{\beta \to 0} I_{\hat{\beta}}(\omega, \alpha, \beta) = 0$$

Equation (1.12) is proved, which is the Generalized ZILC case.

1.7.2 The Asymptotic Variance of β

The information matrix of parameter $(\omega, \alpha, \beta)'$ is

$$I = \begin{pmatrix} A & B & B \\ -- & C & D \\ -- & -- & E \end{pmatrix}$$

The information measure of $\hat{\beta}$ is: $I_{\hat{\beta}} = \frac{T}{I^{-1}(3,3)}$

The asymptotic variance of β is $V_{\hat{\beta}} = I_{\hat{\beta}}^{-1}$

Table 1.1: Size of t-test for β at 5% Level

	T=500	T=1000	T=5000
$\alpha = 0.01$	55.2%	51.7%	48.9%
$\alpha = 0.05$	41.5%	36.9%	25.9%
$\alpha = 0.1$	30.5%	21.9%	10.8%
$\alpha = 0.9$	36.7%	24.8%	21.7%

Notes: Each row represents a different parameter setting of the GARCH(1,1) Monte Carlo simulation. The sample size increase as 500, 1000, and 5000. The repeated time of the Monte Carlo simulation is 1000.

Table 1.2: Size of Various Tests for β at 5% in GARCH(1,1) Model when $\beta = 0$

	T=500	T=1000	T=5000
	$\omega = 1,\ \alpha = 0.01,\ \beta = 0$		
t-test	55.2%	51.7%	48.9%
LR test	13.3%	10.7%	8.3%
LM test	4.7%	5.2%	4.6%
	$\omega = 1,\ \alpha = 0.05,\ \beta = 0$		
t-test	41.5%	36.9%	25.9%
LR test	11.0%	9.9%	7.3%
LM test	4.7%	6.2%	4.2%
	$\omega = 1,\ \alpha = 0.10,\ \beta = 0$		
t-test	30.5%	21.9%	10.8%
LR test	4.5%	6.1%	4.5%
LM test	3.7%	6.0%	4.0%
	$\omega = 1,\ \alpha = 0.9,\ \beta = 0$		
t-test	36.7%	24.8%	21.7%
LR test	5.5%	5.2%	4.5%
LM test	3.9%	4.0%	3.7%

Notes: Each column represents the size of various tests when α changes and samples size is the same. $\beta = 0$ implies no GARCH effect. Each row represents, for the same α, when sample size increases, whether the t-test, LR test, and LM test become more valid. The repeated time of each Monte Carlo simulation is 1000. The table results is consistent with Ma, Nelson, and Startz (2007).

Figure 1.1: Frequency to Reject t-test for β: no GARCH effect ($\beta = 0$)

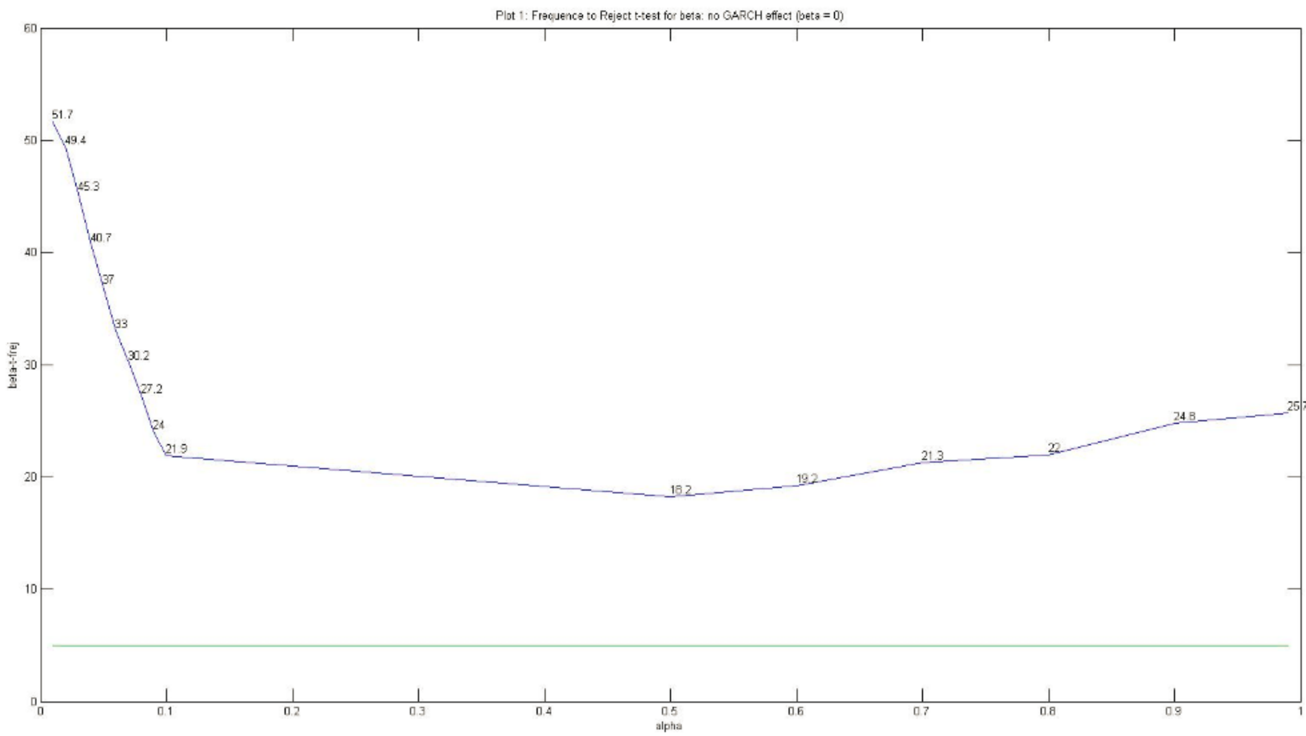

Figure 1.2: Frequency to Reject t-test for β: moderate GARCH effect ($\beta = 0.5$)

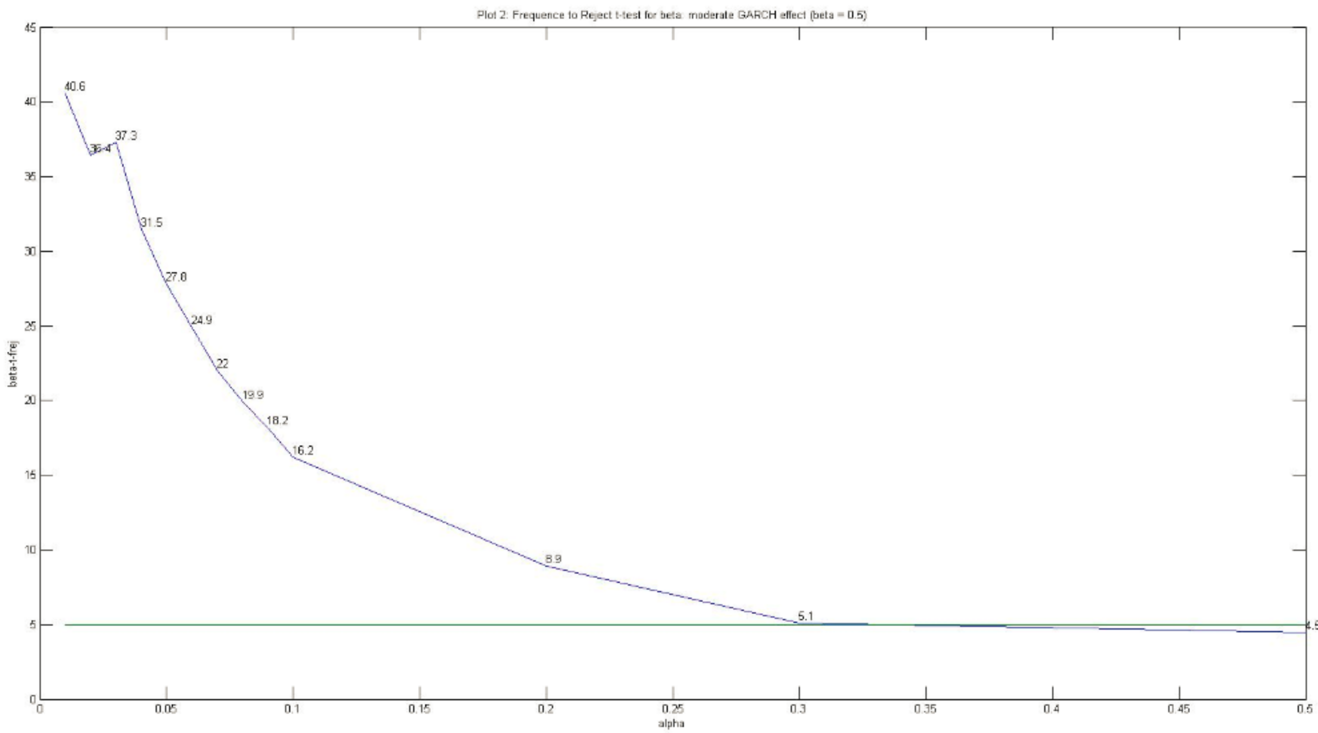

Figure 1.3: The Actual Size of the t-test ZILC Zone and Safe Zone in the GARCH(1,1) Model

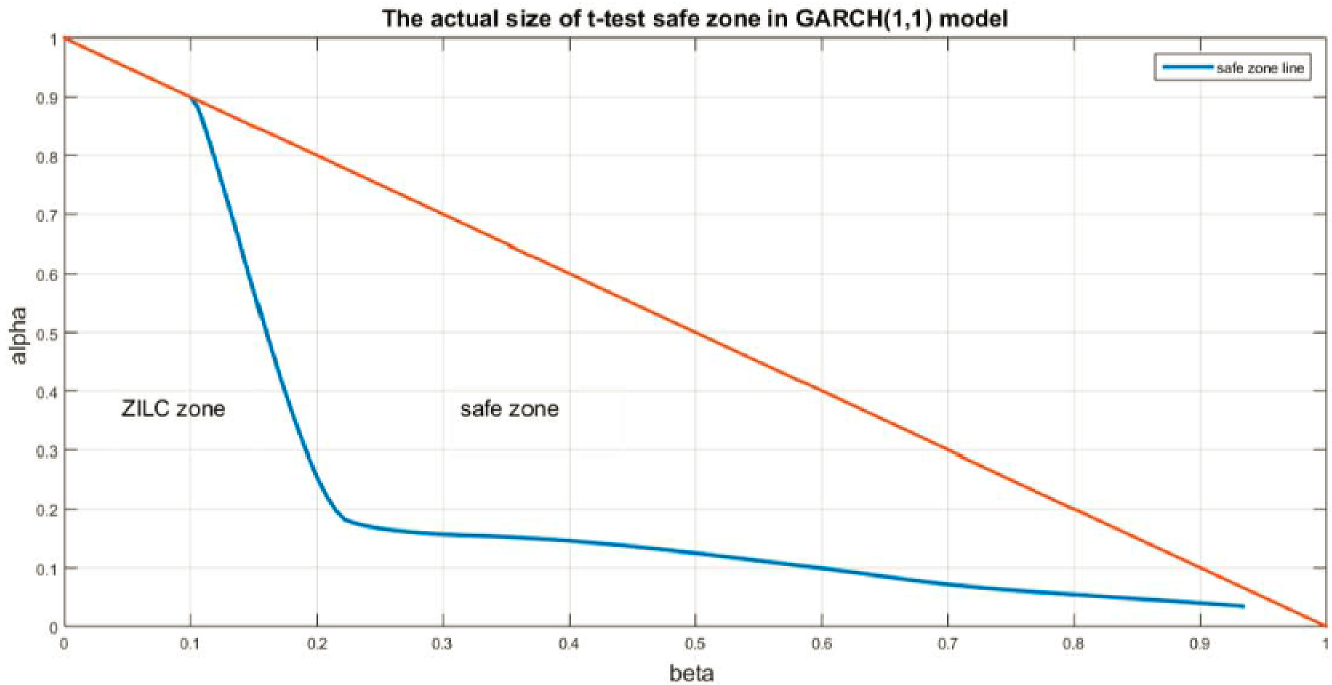

Table 1.3: Size of Various Tests for β at 5% in GARCH(1,1) Model when $\beta = 0.5$

	T=500	T=1000	T=5000
	$\omega = 1$, $\alpha = 0.01$, $\beta = 0.5$		
t-test	43.5%	40.6%	37.9%
LR test	10.0%	8.8%	9.1%
LM test	5.4%	4.9%	4.1%
	$\omega = 1$, $\alpha = 0.05$, $\beta = 0.5$		
t-test	30.1%	27.0%	19.1%
LR test	7.1%	9.8%	6.5%
LM test	3.6%	3.2%	5.2%
	$\omega = 1$, $\alpha = 0.10$, $\beta = 0.5$		
t-test	19.9%	16.3%	9.1%
LR test	4.1%	5.1%	4.2%
LM test	3.3%	4.3%	3.8%
	$\omega = 1$, $\alpha = 0.49$, $\beta = 0.5$		
t-test	6.3%	4.5%	4.9%
LR test	5.4%	6.2%	4.7%
LM test	3.7%	4.0%	3.0%

Notes: Each column represents the size of various tests when α changes and samples size is the same. $\beta = 0.5$ implies the moderate GARCH effect. Each row represents, for the same α, when sample size increases, whether the t-test, LR test, and LM test become more valid. The repeated time of each Monte Carlo simulation is 1000. The table results is consistent with Ma, Nelson, and Startz (2007).

Chapter 2

Do Shocks to Animal Spirits Cause Output Fluctuations?

2.1 Introduction

The notion that confidence on the part of economic agents, be they consumers or firms, is an important source of economic fluctuations is an old one in economics. In *The General Theory of Employment, Interest, and Money*, Keynes (1936) writes, "Most, probably, of our decisions to do something positive, the full consequences of which will be drawn out over many days to come, can only be taken as a result of animal spirits - of a spontaneous urge to action rather than inaction, and not as the outcome of a weighted average of quantitative benefits multiplied by quantitative probabilities." Despite the venerable nature of "animal spirits", however, it is not easy to precisely pinpoint their role in the economy's dynamics. Largely, this is because they are obviously unobservable and deeply entangled with other factors that are likely also to play a strong role, such as the arrival of news about expected future events.

Recently, the macroeconomic literature has revisited this question in both empirical and theoretical studies. Much of the theoretical literature has attempted to formalize animal spirits in the context of otherwise fairly standard models. The transmission mechanisms in these models differ, but they all seek to describe how agents' decision making can be influenced by forces orthogonal to economic fundamentals. These decisions, in turn, generate fluctuations in economic aggregates that may have appeared intuitive to Keynes. Alternatively, it has still proven difficult empirically to convincingly identify the effects of these innovations in animal spirits.

In this paper, we attempt to quantify the importance of animal spirits for changes in economic activity. Using monthly data on economic activity and consumer confidence at the level of U.S. states, we will formally identify shocks to animal spirits that are orthogonal to shocks to economic fundamentals. We will do so in the context of a structural vector autoregression (SVAR), making use of long run and sign restrictions, just as in the seminal work of Blanchard and Quah (1989). Our SVAR will include consumer confidence and a state-level measure of output, and we will impose that innovations in animal spirits cannot permanently affect output and that, contemporaneously, their effect ought to be positive.. We demonstrate that animal spirits shocks do have a statistically significant effect on output and consumption (proxied by retail sales) in the medium term, but that the bulk of fluctuations is still accounted for by innovations in fundamentals (which we will refer to as news shocks).

Our paper is among the first to empirically evaluate the effects of animal spirits shocks on real macroeconomic variables. An important earlier work is that of Barsky and Sims (2012), who also use VAR analysis to identify shocks to consumer confidence. They employ short run restrictions in a VAR that includes output and consumption as well as consumer confidence. By incorporating both news shocks and animal spirits shocks in a standard New Keynesian dynamic stochastic general equilibrium (DSGE) model, they conclude that the innovations they identify in the data behave more like fundamental shocks.[1]

Still, the literature that seeks to develop a formal theoretical link between changes in confidence or sentiment and fluctuations in real variables is rapidly expanding. We discuss a couple of recent examples. Aangeletos and La'O (2013) develop a model in which agents engage in pair-wise trading and the parties to a given pairing receive exogenous sunspot-like signals. An agent's signal contains noisy information about her trading partner's productivity as well as information about her trading partner's signal about her own productivity. With production decisions made before all information is revealed, such information frictions can then lead to aggregate fluctuations. Aangeletos and La'O (2013) define "sentiments," which we interpret as being fundamentally equivalent to animal spirits, as the cross-sectional average of expectations of agents about the productivity levels of their trading partners.[2] In a similar setting (and one that is also reminiscent of the classic island model of Lucas

[1]Interestingly, Bachmann and Sims (2012) find that consumer confidence is an important transmission mechanism of shocks to government purchases, although only in recessions.

[2]It is possible to extend the model of Aangeletos and La'O (2013) to take account of higher order beliefs of the agents in the economy about their trading partner's productivity levels as well as their beliefs about their trading partners' beliefs about their beliefs and so on. This is the approach taken by Huo and Takayama (2015a) and Huo and Takayama (2015b).

(1972), Benhabib, Wang, and Wen (2015) build an economy in which households receive shocks to their expectations about the aggregate economy (the sentiments shock) and to their preferences for differentiated goods. Island firms then must consider a signal extraction problem, in which fluctuations in the demand for their output are in part due to households' changing tastes over goods and in part due to their changing views of the economy as a whole. Like in the case of Aangeletos and La'O (2013), the result is that changes in optimism or pessimism of agents can generate business cycles.[3] In neither case is there a departure from rationality required, a common approach taken to formalize animal spirits in the past. What is interesting is how these model predictions are, in general, not confirmed by the empirical work of Barsky and Sims (2012).

We will provide evidence that affirms a role for animal spirits in real macroeconomic fluctuations. In particular, we show that a positive innovation in animal spirits causes the level of economic activity to be about three percent higher after two years, and it takes nearly five years for the effect of the animal spirits shock on output to wear off. In contrast, however, an innovation in fundamental economic activity itself leads to a ten percent increase in the level of output over a similar time horizon. The effect on consumption, as proxied by retail sales, of a shock to animal spirits is also significantly positive for about 18 months, but is less persistent than the effect of a fundamental shock and is less persistent than the effect of animal spirits on overall activity.

The rest of this paper proceeds as follows. Section 2.2 discusses the data that we

[3] Other important works include Benhabib, Wang, and Wen (2013), Lorenzoni (2009), Schmitt-Grohe and Uribe (2012), and Blanchard, L'Huillier, and Lorenzoni (2013).

use for our analysis, in particular our measure of consumer confidence provided by The Conference Board. Section 2.3 explains our econometric approach. Section 2.4 reports and interprets the results of our analysis, and Section 2.5 concludes.

2.2 Data

Our data consists primarily of three primary variables, which we collect at the U.S. state level. These are an index of consumer confidence, the state coincident indexes compiled by the Federal Reserve Bank of Philadelphia, and a measure of retail sales at the state level. We discuss each of these in turn.

2.2.1 Consumer Confidence

We obtain data on consumer confidence from the Conference Board, who construct their index using their Consumer Confidence Survey. Each month, respondents are asked to assess current business conditions, current employment conditions, expectations for future business conditions (meaning business conditions six months in the future), expectations for future employment conditions (with the future defined in the same way), and expectations for future family income. For each category, they must evaluate their perspective as positive, negative, or neutral. Then, the proportion responding positively is divided by the proportion responding positively or negatively for each category, and this ratio is averaged across the five categories to make up the consumer confidence index. For further details, see The Conference

Board (2011).[4]

Importantly for our purposes, the Conference Board does not only construct their consumer confidence index at the level of the United States, but also at the regional level and stratified by age. Table 2.1 reports the states that each region comprises. We have observations on region-level consumer confidence from January 1981 to December 2014. In order to turn these regional consumer confidence indexes into state consumer confidence indexes, we make use of the separate indexes created for three age groups. Specifically, the Conference Board builds indexes for households where the household head is under the age of 35, is between the ages of 35 and 54 , and is 55 and older. Let $Conf_i$ denote the consumer confidence index for age group $i \in \{young,\ middle,\ old\}$, with *young* referring to the under-35 age group, *middle* referring to the 35-54 age group, and *old* referring to the over-54 age group.[5] We combine these age-specific indexes with information from the U.S. Statistical Abstract on the age profile of each state's population to create a state index that is a weighted average of the age-specific indexes, where the weights are determined by the individual state's age profile.[6] That is, the age-based consumer confidence index for state j is:

$$Conf_{age,j} = \rho_{young} Conf_{young} + \rho_{middle} Conf_{middle} + \rho_{old} Conf_{old}\ , \tag{2.1}$$

[4]Additionally, the Conference Board constructs subindexes based on the appraisals for the first two categories, which make up the "Present Situation" index and on the answers for the last three categories, which form the "Expectations" index. We focus on the main Consumer Confidence Index which derives from the answers to all five questions.

[5]In this exposition, for notational convenience, we suppress the time subscripts.

[6]We note that information on the age profile of each state is available for the years 1980, 1990, 2000, 2005, and 2010. We impose that the age profile remains the same each year until an update is available. That is, the age profile for 2003 is assumed to be the same as it was in 2000, but it updates in 2005.

where the ρ's are the share of state j's population belonging to each age group. We then take a population-weighted average of these state-level, age-based consumer confidence indexes by region. If there are J states in region k, this measure is:

$$Conf_{age,k} = \sum_{j=1}^{J} \beta_j Conf_{age,j} \ . \tag{2.2}$$

Let $Conf_{CB,k}$ denote the published regional level consumer confidence index produced by the Conference Board. We construct a scaling factor, θ_k, such that:

$$\theta_k = \frac{Conf_{CB,k}}{Conf_{age,k}} \ , \tag{2.3}$$

and the final state-level consumer confidence index for state j is

$$Conf_j = \theta_k Conf_{age,j} \ . \tag{2.4}$$

This is the measure that we will be employing in our analysis.[7] Figure 2.1 produces the computed state-level consumer confidence index for nine states, one from each of the regions for which the Conference Board produces a consumer confidence index, with each region represented by the state listed first alphabetically. Table 2.2 contains the summary statistics for our consumer confidence measure.

2.2.2 State Coincident Indexes

Our chosen measure of economic activity at the level of U.S. states is the Coincident Economic Index compiled by the Federal Reserve Bank of Philadelphia.[8] The

[7] In practice, the state-level consumer confidence indexes generally do not differ that much from the coincident region-level consumer confidence index. This is because, within a region, differences in states' age profiles tend to be fairly modest.

[8] We will use the term "economic activity" index also to refer to this measure.

interested reader should consult the Philadelphia Fed's website for details, but this measure applies the latent dynamic factor model of Stock and Watson (1989) to state-level economic data. Such an application was popularized in Crone and Clayton-Matthews (2005). The model assumes that the state of economic activity in a given U.S. state can be described by a latent factor that summarizes the comovement of four variables, which are nonfarm payroll employment, the unemployment rate, average manufacturing labor hours, and real wages and salaries. Clearly, this index is dominated by labor market variables,[9] but these indexes have a number of attractive features for us. In particular, they are collected at monthly frequency (matching the frequency of our consumer confidence data) and the data is available from January of 1979, encompassing the time dimension of our confidence data. Also, labor markets variables tend to be highly correlated with general economic activity and are of great interest to policymakers. For all of these reasons, we believe that this is a good measure to use for our state-level analysis.

Figure 2.2 plots the annualized growth rate in this index for the same nine states for which we plotted the consumer confidence index. This figure helps demonstrate the advantages inherent in exploiting data at the U.S. state level, as opposed to only considering aggregate U.S. macroeconomic data. The figures reveal substantial variation in annualized growth in economic activity cross-sectionally. For example, while most of the states displayed in the figure experienced a very deep recession along with the rest of the United States in the early 1980s, Alaska doesn't fall into

[9] Using a modified form of this technique, Arias, Gascon, and Rapach (2015) construct similar monthly indexes for large metropolitan areas in the United States, and they make use of a wider range of indicators than merely labor market figures.

recession until the mid-1980s (when oil prices began to fall). In some states, the recent recession associated with the global financial crisis was the most severe in the sample period, but in others, it is eclipsed by the early 1980s recessions. Arizona witnesses a tremendous spike in growth in the mid-1980s, while, in the early 1990s, growth in Arkansas does not turn negative as it does in many other states. Such patterns have also been noted by, among others, Owyang, Piger, and Wall (2005) and Owyang, Rapach, and Wall (2009).

Summary statistics for the annualized growth rate of the state coincident economic index are also found in Table 2.2. Figure 2.3 simultaneously displays the consumer confidence index and annualized growth in the coincident economic index for one state, Connecticut. The two series apparently track each other fairly closely, although consumer confidence lags economic activity somewhat.

2.2.3 Retail Sales

We also collect data on retail sales at the state level and at quarterly frequency. We consider retail sales as a useful proxy for consumption, especially since consumption data is not readily available for states at such frequency.[10] In order to construct series on retail sales for U.S. states, we follow the work of Garrett, Hernandez-Murillo, and Owyang (2004), who use quarterly data on state sales tax revenue, which comes from the U.S. Census Bureau. We also collect data on state sales tax rates from a

[10] We note that the Bureau of Economic Analysis has recently introduced data on Personal Consumption Expenditures for U.S. states, but these data are at annual frequency, which makes them an imperfect fit for our analysis, which attempts to exploit relatively high frequency variation in consumer confidence and real economic variables.

variety of sources, including the Census Bureau, the Book of the States, and the Tax Foundation. We compute state retail sales as the sales tax revenue divided by the sales tax rate for a given state in a given quarter. Our retail sales data extends from 1994:Q1 to 2014:Q4, and we have data for 42 states.[11] We deflate the retail sales series by the national-level consumer price index. Garrett, Hernandez-Murillo, and Owyang (2004) show that a when the inferred state retail sales data are summed up cross-sectionally, they have a correlation coefficient of 0.975 with national retail sales data collected by the U.S. Census.

Summary statistics for annualized growth in retail sales can be found in the third line of Table 2.2, and Figure 2.4 contains plots of the time series of annualized retail sales growth for those states for which we have data on retail sales and for which we have already produced plots for the consumer confidence and state coincident economic index series.

2.3 Econometric Methodology: Structural VAR

Our analysis begins with the identification of shocks to animal spirits and to fundamental economic variables, which we will refer to as "news." This is motivated by the work of Barsky and Sims (2012), who also distinguish between innovations to expectations about future economic activity ("news") and innovations to consumer confidence that are otherwise unrelated to real economic variables. Unlike

[11] As noted by Garrett, Hernandez-Murillo, and Owyang (2004), there are five states that do not have a sales tax, so we cannot compute an estimate of state retail sales for these. They include Alaska, Delaware, Montana, New Hampshire, and Oregon. Further, data for Nevada, Utah, and Wyoming was not of high quality, so we drop these states as well.

Barsky and Sims (2012), who use VARs with short-run restrictions to inform their DSGE model, we will make use of long-run restrictions in our VAR, following the model of Blanchard and Quah (1989). Specifically, we will restrict the animal sprits innovations to have only transitory effects on the level of output.

Our structural VAR will include two variables, the state-level consumer confidence index that we modify from the region-level confidence index collected by the Conference Board (with $Conf_{i,t}$ denoting the consumer confidence index in state i at time t) and the annualized growth rate in the state coincident economic index constructed by the Federal Reserve Bank of Philadelphia (with $\Delta y_{i,t}$ denoting the annualized growth rate in the economic activity index in state i at time t), each observed at a monthly frequency. Consider the following reduced form VAR representation of these two variables:[12]

$$\begin{bmatrix} Conf_{i,t} \\ \Delta y_{i,t} \end{bmatrix} = \begin{bmatrix} A_{11}(L) & A_{12}(L) \\ A_{21}(L) & A_{22}(L) \end{bmatrix} \begin{bmatrix} Conf_{i,t-1} \\ \Delta y_{i,t-1} \end{bmatrix} + \begin{bmatrix} e_{1,i,t} \\ e_{2,i,t} \end{bmatrix} \tag{2.5}$$

We assume that the reduced form residuals $[e_{1,i,t}, e_{2,i,t}]'$ are in fact linear combinations of two underlying structural shocks, which we interpret as shocks to animal spirits and to economic fundamentals or expectations about economic fundamentals. We call this latter innovation "news." Our objective is to uncover from these reduced form VAR residuals the structural innovations in animal spirits and in news. To do so, we will apply the following critical identification assumption:

Assumption 1: The structural innovation in animal spirits must not have an effect on the level of economic activity in the long run. That is, fluctuations in

[12]Much of the notation to follow is borrowed from Enders (2010).

economic activity caused by animal spirits shocks must be transitory in nature.

The motivation for this assumption draws heavily from Blanchard and Quah (1989).[13] In that paper, the authors assumed that, in the long run, shocks to aggregate demand could not have an impact on the level of output. Only shocks altering the productive capacity of the economy could do that. Similarly, we assert that shocks to animal spirits, the "spontaneous urge to action," in the parlance of Keynes (1936), cannot permanently change the long run level of output. Econometrically, it is more or less trivial to identify innovations that have only transitory effects on output and call them "animal spirits." Intuitively, it may be worth delving a little deeper into this assumption.

We argue that such flights of optimism that we call animal spirits are likely to wear off with time and as information about the true state of the economy is revealed. That is, if a particular household suddenly (irrationally) believed that the level of output would be higher in the future than their given information set would imply, then they ought to eventually revise their belief as new information showed it to be ill-formed. Consider even the model of Aangeletos and La'O (2013), which relies on communicational frictions to provide a role for animal spirits in economic fluctuations. While the sunspot-like extrinsic shock of their model generates boom-and-bust cycles, output eventually returns to its long run level as information about the true productivity levels of the various island economies was revealed.[14]

It might be argued that such animal spirits could have a long run effect on the

[13] See also Basu, Fernald, and Kimball (2006) and Christiano, Eichenbaum, and Vigfusson (2004).
[14] See Figure 1 in Aangeletos and La'O (2013).

level of output if they spurred firms to invest in greater quantities, thus raising aggregate supply and subsequently output. Here,we would assert that our econometric approach to identifying shocks to animal spirits does not allow for this possibility. For one, we use a measure of consumer (as opposed to business or CEO) confidence to identify animal spirits. Thus, the shocks we identify are far more likely to impact consumption in the short run than investment. Secondly, if consumer confidence were to rise in such a way as to generally reflect a broader optimism about future productivity levels in the economy, this sort of innovation would be captured in our identified news shock. Such increases in confidence can impact the economy in the long run, because they may be associated with expectations for faster future consumption growth due to higher output.

More broadly, it may seem ambitious to suppose that these economies that we consider are driven by only two shocks. It is certainly the case that we are assuming that a very large number of disparate types of innovations that each affect the economy in very different ways (i.e., some can have permanent effects and some may have only transitory effects) can be rolled up into just one "fundamental" shock, with only the effects of animal spirits shocks excluded and separately identified. Also, we note the critique of Blanchard, L'Huillier, and Lorenzoni (2013), who argue that structural vector autoregression can not be used to disentangle noise shocks (which may be considered roughly analogous to the animal spirits shocks that we are trying to identify), because, when consumers face a signal extraction problem, the reduced form VAR is not invertible.[15] That is, the econometrician can not observe noise

[15] Fernande-Villaverde et al. (2007) show that when the VAR is not invertible, one cannot interpret the innovations in VARs as being equivalent to the economic shocks posited by theory.

shocks when she has at most the same data available to the consumer.

In this respect, we note the following. First, we note that the exercise that we undertake in this paper is subtly different from that described in Blanchard, L'Huillier, and Lorenzoni (2013). They assume that "noise" shocks cannot influence productivity, and the problem that they describe is that the econometrician can not use long run VAR restrictions to separate the permanent shock to productivity from the combination of the temporary shock to productivity and the noise shock. In our paper, we seek to separate the animal spirits shock from any shock that can have a long run effect on output (like their permanent productivity shock). The risk then is that our animal spirits shock might be an amalgam of the true animal spirits shock and the temporary shock to productivity (again in the framework of Blanchard, L'Huillier, and Lorenzoni (2013)). We believe that our results are not consistent with this interpretation, which will be discussed further below. Further, their assumption that noise shocks do not have an effect on productivity is one that we are inherently questioning in this paper. As in Aangeletos and La'O (2013), Benhabib, Wang, and Wen (2015), and Huo and Takayama (2015a), our model proposes that animal spirits shocks can in fact induce output fluctuations.

Second, we draw on our use of panel data. Although we run our VARs state by state, we estimate the impulse responses in a panel setting (as will be discussed further below). By incorporating data on 50 small, open economies (the 50 U.S. states), this paper believes that, with the use of time fixed effects, we can control for a substantial number of economic shocks that might plausibly be affecting all states simultaneously, such as monetary policy shocks, fiscal policy shocks, and transitory and permanent shocks to aggregate productivity. Of course, there could still be idiosyncratic shocks to productivity in each state. Again, we will argue below that our results are not consistent with the notion that our identified shock is a transitory shock to productivity, but they are consistent with the idea that it is a sentiments shock.

Blanchard, L'Huillier, and Lorenzoni (2013) demonstrate that when long run restrictions are used to identify shocks that have permanent and transitory effects on productivity in an environment where noise shocks force the consumer to solve a signal extraction problem in order to estimate the permanent component of productivity, long run restrictions in a SVAR, such as those used in Blanchard and Quah (1989) and that we propose to use here, will systematically attribute too much of the variance in output fluctuations to the permanent component and underestimate the contribution of transitory components. In this sense, then, it is likely that our estimates of the role for animal spirits in generating business cycle fluctuations are biased toward zero. While this is a problem, we note that the predominant study in the empirical literature (Barsky and Sims (2012)) finds a very small role for animal spirits, so the evidence that we will present for a moderate contribution of animal spirits is fairly novel.

As Enders (2010) points out, one unappealing aspect of using long run restrictions in a VAR to uncover structural shocks is that the econometrician must impose some restriction on the signs of the contemporaneous effects of the structural innovations on the observed variables in order to identify them. That leads us to our second key identifying assumption.

Assumption 2: The structural innovation in animal spirits must have a positive contemporaneous effect on both consumer confidence and growth in economic activity.

We assume that a shock to animal spirits must be, in part, captured by a positive change in the consumer confidence index. That is, if households are feeling optimistic about the state of the economy in a fashion that is orthogonal to economic fundamentals, then their answers to the questions posed in the Conference Board's Consumer Confidence Survey must, on balance, be positive. Similarly, a positive "spontaneous urge to action" is more likely to lead to an uptick in consumption on the part of the household (and thus higher output) than a decline.

The imposition of the assumption that animal spirits cannot affect the long run level of economic activity, combined with the contemporaneous sign restriction, can be used to decompose the reduced form VAR residuals into structural innovations in animal spirits and news. Enders (2010), pages 338 to 342, provides a step-by-step guide.

We estimate the structural VAR with long run restrictions individually, state-by-state. Once we have uncovered the structural innovations in animal spirits and in news, we run the following panel regressions:

$$\Delta y_{i,t} = \alpha_{1,i} + \delta_{1,t} + \sum_{j=0}^{T} \beta_{1,j} \epsilon_{NEWS,i,t-j} + \sum_{j=0}^{T} \beta_{2,j} \epsilon_{SPIRITS,i,t-j} + v_{1,i,t} \tag{2.6}$$

$$Conf_{i,t} = \alpha_{2,i} + \delta_{2,t} + \sum_{j=0}^{T} \beta_{3,j} \epsilon_{NEWS,i,t-j} + \sum_{j=0}^{T} \beta_{4,j} \epsilon_{SPIRITS,i,t-j} + v_{2,i,t}\ . \tag{2.7}$$

In these equations, $\epsilon_{NEWS,i,t}$ and $\epsilon_{SPIRITS,i,t}$ denote the structural innovations in news and animal spirits, respectively, for state i at time t. $\alpha_{k,i}$ for $k = 1, 2$ represents a set of state fixed effects in either equation, and $\delta_{k,t}$ for $k = 1, 2$ is a set of time fixed effects. What we are interested in are the sequences of coefficients $\{\beta_{1,j}, \beta_{2,j}, \beta_{3,j}, \beta_{4,j}\}$ that make up the impulse responses of state economic activity and consumer confidence to the structural innovations in news and animal spirits. We set T equal to 60 months, or five years.

In addition to tracking the impulse responses of output (proxied by the state economic activity index) and consumer confidence to innovations in news and animal spirits, we also study the response of consumption, proxied by state-level retail sales. Our data on retail sales is quarterly, as opposed to the monthly frequency in our consumer confidence and economic activity series. Therefore, in order to estimate the impulse response of retail sales, we must identify a quarterly innovation in animal spirits and in news. Thus, we re-estimate the structural VAR with long run and sign restrictions at the quarterly frequency and use the quarterly structural innovations to estimate the impulse response of retail sales in a similar regression as Equations 2.6 and 2.7.

There is a tension in identifying structural innovations in news and animal spirits at the monthly frequency and then aggregated up to the quarterly frequency. Indeed, it is very likely that quarterly innovations are not nearly as well identified, as they probably contain an element that is already known to agents. This is a caveat that we bear in mind when examining the impulse responses of retail sales to these quarterly innovations.

2.4 Results

This section contains our empirical results. Figure 2.5 contains the benchmark estimates of the impulse responses of output (top row) and consumer confidence (bottom row) to structural innovations in news (left column) and animal spirits (right column).[16]

Consider first the response of consumer confidence to an innovation in animal spirits, in the bottom right hand quadrant of Figure 2.5. This response is a key test of the validity of our approach to identifying shocks to animal spirits. Intuitively, if our conceptualization of animal spirits is correct, and they are changes in optimism or pessimism orthogonal to economic fundamentals, then they should have a significant effect on consumer confidence in the short run. Indeed, Barsky and Sims (2012) identify animal spirits as the orthogonalized shock to consumer confidence when it is ordered first in a Cholesky decomposition of reduced form residuals from a VAR with short-run restrictions. Reassuringly, our identified shock to animal spirits does have a significant contemporaneous positive effect on consumer confidence. A one standard deviation shock to animal spirits pushes the consumer confidence index up by 6 index points on impact. The effect quickly reverts to 2 index points, and then it slowly erodes over the next five years, until the consumer confidence index is back to its pre-shock level.

[16] We report results using conventional panel OLS standard errors. Of course, because our right hand side variables are estimated, and not directly measured, we could potentially have a generated regressor problem. See Pagan (1984). We attempt to address this concern by estimating a specification with bootstrapped standard errors. The results are nearly identical. We also cluster standard errors at the state and at the region level. Similarly, the results are barely changed.

Indeed, it is interesting that an animal spirits shock has such a persistent effect on consumer confidence, even though more than half of the initial response decays quite quickly. This persistence, as we will see, is also reflected in the response of output to an animal spirits innovation.

In the top left hand quadrant of Figure 2.5, one can see the response of output (measured by the state-level coincident economic index) to an innovation to "news," or economic fundamentals. Here again, the response is intuitive and helps validate our identification strategy. The level of output rises slowly in response to a news shock, before it plateaus at a level about 10% higher than the pre-shock level of output. Five years after the shock, output is still higher, suggesting that news shocks (or shocks to fundamentals) have a permanent effect on the level of output. This is consistent with the notion that what we refer to as news shocks carry information to agents reflective of changes in the level of productivity in the economy.

The panel in the bottom left of the figure shows the response of the consumer confidence index to a news shock. Interestingly, the immediate effect is negative, with the consumer confidence index falling more than two points on impact when a positive news shock hits the economy. In subsequent periods, consumer confidence recovers strongly and remains about one point above its pre-shock level for the duration of the estimation window. We can compare this result to those of Basu, Fernald, and Kimball (2006), who famously found that, following a shock to total factor productivity in the economy, employment, hours, and utilization all decline significantly. While we do not observe the same decline in real activity (which is an aggregate of four labor market variables, including employment and hours worked) after an innovation in news, it is notable that consumer confidence, which is likely to be heavily influenced by labor market variables like employment and hours, exhibits a similar pattern in response to a similar kind of shock. Again, speaking broadly, we find such a result to be supportive of our interpretation of this innovation as a news shock.

Finally, the top right hand panel in the figure shows the effect of an innovation in animal spirits on real economic activity. By construction, animal spirits have no long term effect on the level of activity, and, after five years, the influence of the shock has completely eroded. In the interim, though, animal spirits shocks do have a significant positive effect on output, generating a hump-shaped path for our measure of fundamentals. The state coincident economic index rises slowly, but steadily, in response to an innovation in animal spirits, peaking after about a year and a half, with the level of output rising 3% above its pre-shock level (on an annualized basis), after which the effect deteriorates.

If the question is whether or not an innovation in animal spirits that is unconnected to any fundamental change in real driving variables can have a real effect on output, this plot demonstrates that the answer is yes. Thus, our results do not accord with the results of Barsky and Sims (2012), but are more consistent with the theoretical findings of Angeletos and La'O (2013), Benhabib, Wang and Wen (2015), and Huo and Takayama (2015a). A change in the annual level of output of about three percentage points in less than two years is a very sizable impact, although as the discussion above shows, the effect of a news shock on the level of output is much greater and actually leads to permanent changes. Table 2.3 reports the variance decomposition for each of the endogenous variables, with each cell of the table displaying the proportion of the overall variance in each of the two endogenous variables attributable to each of the structural innovations under consideration at the given time horizon.

Table 2.3 demonstrates that at short horizons, nearly all of the variation in annualized growth in the state coincident economic index is accounted for by shocks to news, with only a tiny portion due to animal spirits. While the share of variation made up by animal spirits grows over time, even at fairly long horizons (up to five years), animal spirits can explain less than 15% of total variation in real economic activity.[17] On the other hand, the identified innovation in animal spirits can explain a great deal of the variation in the consumer confidence index, as we might expect. These shocks explain nearly 90% of the variance in the consumer confidence index at one month and up to 94% at the one-year horizon. The innovation in news, by contrast, explains a much smaller share of the changes in consumer confidence.

It is also worthwhile to compare our results for output to those of Blanchard and Quah (1989), who examine the dynamic responses of aggregate output and the unemployment rate to demand and supply disturbances. In particular, we note that our estimated responses of the state coincident economic indexes to news, or fundamentals, shocks is very similar to the response of output to an aggregate supply shock in Blanchard and Quah (1989), while the response of the coincident index to

[17]Note that the variance decompositions reported in Table 2.3 are for the growth rate, not the level, of the state coincident index. While the effect of the innovations in animal spirits on the level of output must go to zero, the effect on the growth rate need not go to zero within the same time period. This is because all of the positive effects that animal spirits have on output must be undone in accordance with the identifying assumption and this manifests in a negative effect on the growth rate of the state coincident index at longer horizons.

an animal spirits shock is has the same hump shape and persistence as the response of output to a demand disturbance in that paper. In that sense, our identified news shocks have similar effects as supply shocks and our animal spirits shocks act like demand shocks.

We next revisit the question of whether or not the shock that we identify can in fact be characterized as a shock to animal spirits. In particular, we discussed in Section 2.3 that there was a risk that the true animal spirits shock we aimed to identify might in practice be conflated with an idiosyncratic transitory shock to productivity. We argue that our results are more supportive of the idea that the shock we identify is in fact an animal sprits shock. The fact that the response to output is hump shaped is crucial to this argument. In Blanchard, L'Huillier, and Lorenzoni (2013), the theoretical response of output to a transitory productivity shock is an immediate jump up on impact, followed by a slow decline back to the pre-shock level. That is not the response that we observe. Rather, our estimated response more closely resembles the theoretical response of output to a sentiments shock in Angeletos and La'O (2013). Here, it is important to recall that Blanchard, L'Huillier, and Lorenzoni (2013) do not allow their noise shocks (which might be treated as being more analogous to our animal spirits shocks or the sentiments shocks of Angeletos and La'O (2013)) to have any effect on output at all, an assumption that may be too restrictive.

In Figures 2.6 through 2.10, we report the impulse responses for the various state-level labor market variables that make up the state coincident indexes. We first consider the response of nonfarm payrolls to shocks to news and animal spirits.

The shapes of the impulse responses with respect to both structural shocks roughly mirror the responses of the coincident indexes. After an innovation in news, the number of employees on nonfarm payrolls slowly rises before plateauing at an increase of six percent relative to the pre-shock level. The peak in the response of this variable occurs after about three years. In contrast, after a shock to animal spirits, the number of employees on nonfarm payrolls rises initially, but peaks (after about eighteen months), and then returns to its pre-shock level. At the peak of the impulse response to an animal spirits innovation, nonfarm payrolls are about two percent higher than they were before the shock.

Figure 2.7 displays the response of the unemployment rate at the state level to each of the innovations. We see similar patterns in the behavior of the unemployment rate, although the signs are reversed, of course. After a news shock, the unemployment rate quickly falls, dropping to a level about 0.1% below its pre-shock level at a horizon of around two years. Then, the unemployment rate begins to slowly trend back toward its initial level. This is intuitive. Considering that we are looking at U.S. states and that there is generally greater labor mobility across states than across countries, we would expect that when a state experiences a fundamental news shock that leads to an increase in employment (as shown in Figure 2.6), agents may be encouraged to enter the labor market in that state, whether they were already state residents but not in the labor force before the news shock or they moved to that state from another state. The relatively non-persistent response of the unemployment rate to a news shock is consistent with such a story. After an animal spirits shock, the unemployment rate falls a little more than 0.05 percentage points after about two years, but subsequently quickly reverts to its pre-shock level.

We next consider the response of average weekly hours worked by production workers in manufacturing, which responses can be found in Figure 2.8. We can observe that in response to neither shock do average weekly hours significantly respond, either statistically or economically. After a news shock, hours are generally higher, but, especially by the end of the estimation window, the difference relative to the pre-shock amount of hours worked is not statistically significant. Similarly, after an animal spirits shock, hours actually turn negative, but not significantly so. They turn positive after about two years, but, again, the response is not significant. After both shocks, the responses are quite volatile.

In Figures 2.9 and 2.10, we report the responses of aggregate wages and salaries and average wages and salaries per nonfarm worker, respectively. As with nonfarm payrolls, aggregate wages and salaries respond significantly positively, and permanently, to a news shock. The response reaches its peak level after about two and a half years, and the total amount of wages and salaries paid in the state after a positive news shock is five percent higher at the peak than it was before the state was hit by the news shock. After a positive animal spirits innovation, aggregate wages and salaries rise initially to about 1.5% higher than the pre-shock level, but they quickly revert back. When we consider the responses of wages and salaries per nonfarm worker, we see slightly different patterns. Wages and salaries per worker rise more slowly after a positive news shock than aggregate wages and salaries, and they peak at only about two percent higher than the pre-shock level. Again, this is not surprising when we recall that the response of the number of workers to a news innovation in a given state was strongly positive. Thus, some of the rise in aggregate wages and salaries is due to an influx in the number of workers. In response to an animal spirits shock, the response of wages per worker is roughly flat.

We next examine the response of state-level retail sales to a structural innovation in news and in animal spirits, and these estimates can be found in Figure 2.11.[18] Although our estimate of the response of retail sales to a news shock is not very precise, the point estimate is well above zero up to about eight years and is significant at the 95% confidence level. At the peak of the response, retail sales are close to five percent higher than before the news innovation. We would certainly expect that, if output fluctuations can indeed be generated by innovations in animal spirits, that one channel through which they would work would be the consumption channel. This is what we observe. The estimates are much noisier than those for the state coincident economic index and its components, but we do see a positive and significant (both statistically and economically) consumption response to a shock to animal spirits.

In the case of the responses to news, both retail sales and the labor market variables that make up the state coincident economic indexes see similar time patterns. The responses rise slowly at first before plateauing a couple of years after the innovation has hit the state economy. In general, it is also the case that, in response to a shock to animal spirits, the time paths are also pretty similar, exhibiting a hump-shaped response at intermediate horizons. One curious aspect of the findings that we have reported so far is that, in general, the effects of animal spirits shocks are fairly persistent. For many of the macroeconomic time series that we study, it takes

[18] We set the number of lags of the structural innovations in the regression to be 36 quarters.

up to two years for the response to animal spirits innovations to peak. This might seem surprising, even though (by construction) the effects erode to zero eventually, because it is not clear that such spontaneous urges to action should last as long as they appear to. This is a question that we aim to return to in the future.

2.5 Conclusion

In this paper, we address whether "animal spirits," as described by Keynes (1936), can have real effects on business cycle fluctuations. This question is motivated by recent theoretical literature that looks to formalize a role for such sentiments shocks within a modern rational expectations framework. We leverage data on consumer confidence provided by the Conference Board and indexes of coincident economic activity, both at monthly frequency and at the level of U.S. states to provide an answer to this question. To identify shocks to animal spirits, separate from shocks to news about future economic fundamentals, we employ long run restrictions in a structural vector autoregression framework, following the seminal paper by Blanchard and Quah (1989). In particular, we assume that shocks to animal spirits can not have an effect on the level of output in the long run and that, on impact, they must have a positive effect on both output and on measures of consumer confidence.

Our findings suggest that animal spirits shocks can indeed generate business cycle fluctuations. Specifically, two years after a positive shock to animal spirits, the level of output is about three percent higher, but this positive effect fades after about five years. In contrast, the bulk of variation (more than 85% at the five-year horizon)in output growth is caused by shocks to economic fundamentals (or "news"). Thus, although animal spirits can give important fluctuations at business cycle frequencies, the main driver of output in the short run and in the long run is our estimated news shock.

Retail sales, a proxy for consumption, respond positively to both news and animal spirits shocks, though again, the response is stronger for news shocks. Among the components that make up the state-level coincident economic indexes, the most precisely estimated responses are in nonfarm payrolls, the unemployment rate, and aggregate wages and salaries.

Certainly, these results would benefit from more formal structural modeling. In future, we hope to reconsider them in the context of a Dynamic Stochastic General Equilibrium (DSGE) model to better understand certain dynamics that we observe. In particular, it would be worthwhile to understand why innovations in animal spirits should have such persistent effects on output. Questions such as this will be left to future research.

Table 2.1: Conference Board Categorization of States into Regions

Region				
New England	Middle Atlantic	South Atlantic	East North Central	East South Central
Connecticut	New Jersey	Delaware	Illinois	Alabama
Maine	New York	Florida	Indiana	Kentucky
Massachusetts	Pennsylvania	Georgia	Michigan	Mississippi
New Hampshire		Maryland	Ohio	Tennessee
Rhode Island		North Carolina	Wisconsin	
Vermont		South Carolina		
		Virginia		
		West Virginia		

Region				
West North Central	West South Central	Mountain	Pacific	
Iowa	Arkansas	Arizona	Alaska	
Kansas	Louisiana	Colorado	California	
Minnesota	Oklahoma	Idaho	Hawaii	
Missouri	Texas	Montana	Oregon	
Nebraska		Nevada	Washington	
North Dakota		New Mexico		
South Dakota		Utah		
		Wyoming		

Notes: This table provides a list of states belonging to each of the regions for which the Conference Board compiles a separate consumer confidence index.

Table 2.2: Summary Statistics of Key Variables

Variable	Mean	Std Dev 1	Std Dev 2	Observations
State-Level Consumer Confidence	92.46	14.88	28.57	20400
Coincident Economic Index Growth (Annualized)	2.26	2.60	3.85	20350
State-Level Retail Sales Growth (Annualized)	1.05	7.22	7.34	3486

Notes: Consumer Confidence units are index points. Growth in the coincident economic index and retail sales are expressed in percentage points. "Std Dev 1" is defined as the time average of $[(1/n)\sum_i (X_{it} - \bar{X}_t)^2]^{1/2}$. "Std Dev 2" is defined as the cross sectional average of $[(1/T)\sum_t (X_{it} - \bar{X}_i)^2]^{1/2}$.

Table 2.3: Variance Decomposition of Effects on State Coincident Economic Index and Consumer Confidence

Horizon	*Proportion of Variance in Coincident Economic Index due to:*	
	News	Animal Spirits
1 month	0.997	0.003
6 months	0.965	0.035
12 months	0.910	0.090
24 months	0.903	0.097
36 months	0.902	0.098
48 months	0.890	0.110
60 months	0.860	0.139
Horizon	*Proportion of Variance in Consumer Confidence due to:*	
	News	Animal Spirits
1 month	0.105	0.895
6 months	0.067	0.933
12 months	0.060	0.940
24 months	0.082	0.918
36 months	0.100	0.900
48 months	0.109	0.891
60 months	0.119	0.881

Notes: In each panel, each cell of the table reports the proportion of the overall variance in the given endogenous variable attributable to the shock given by the column heading at the time horizon given by the row heading.

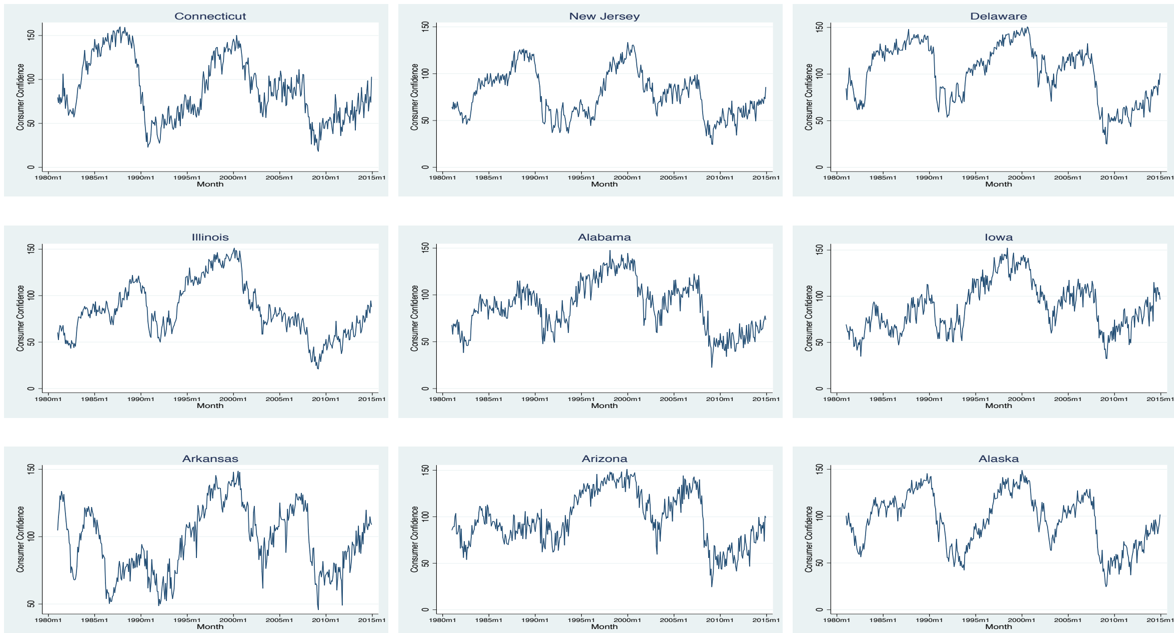

Figure 2.1: Consumer Confidence

Notes: Each figure plots the state-level consumer confidence measure over time. The states are chosen as the first state alphabetically in the region for which the Conference Board tabulates its consumer confidence measure. The derivation of the state level consumer confidence measure from the region level index is explained in the text.

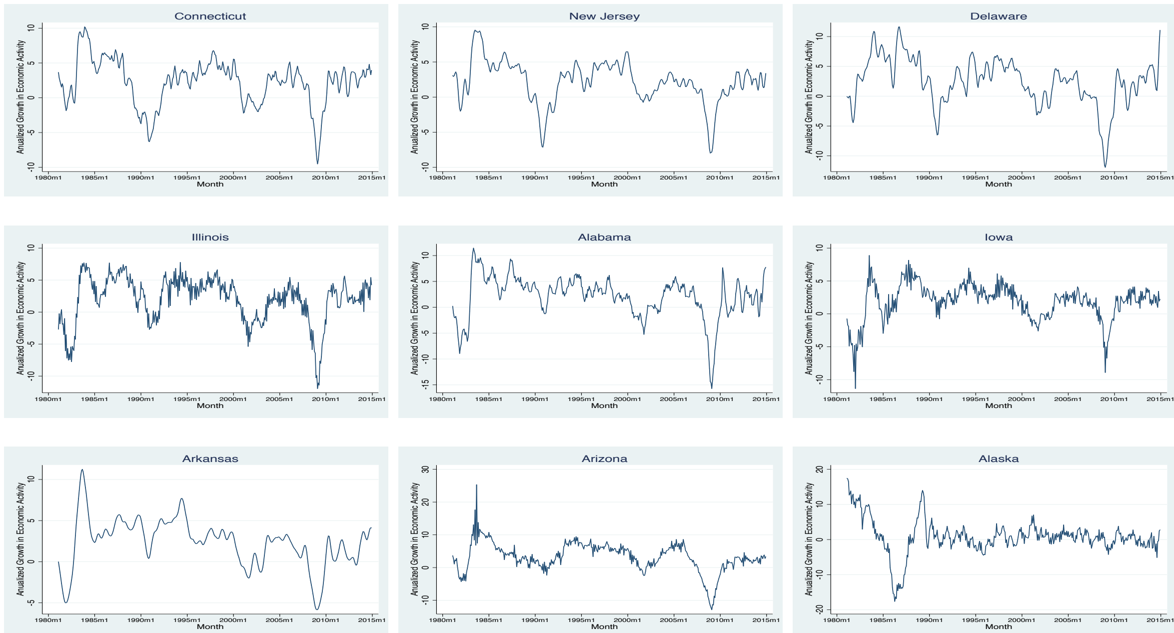

Figure 2.2: State Economic Indexes

Notes: Each figure plots the annualized growth rate in the state-level coincident economic index computed by the Federal Reserve Bank of Philadelphia. The states are chosen as the first state alphabetically in the region for which the Conference Board tabulates its consumer confidence measure.

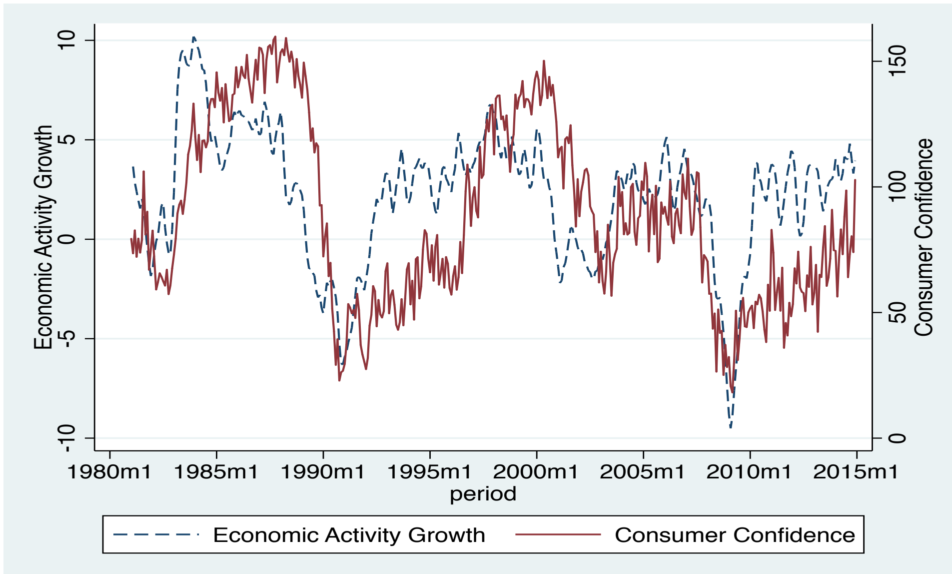

Figure 2.3: Consumer Confidence and Economic Activity Growth in Connecticut

Notes: The figure plots the state-level consumer confidence measure computed by the Conference Board and annualized growth in the state-level coincident economic index computed by the Federal Reserve Bank of Philadelphia over time in Connecticut.

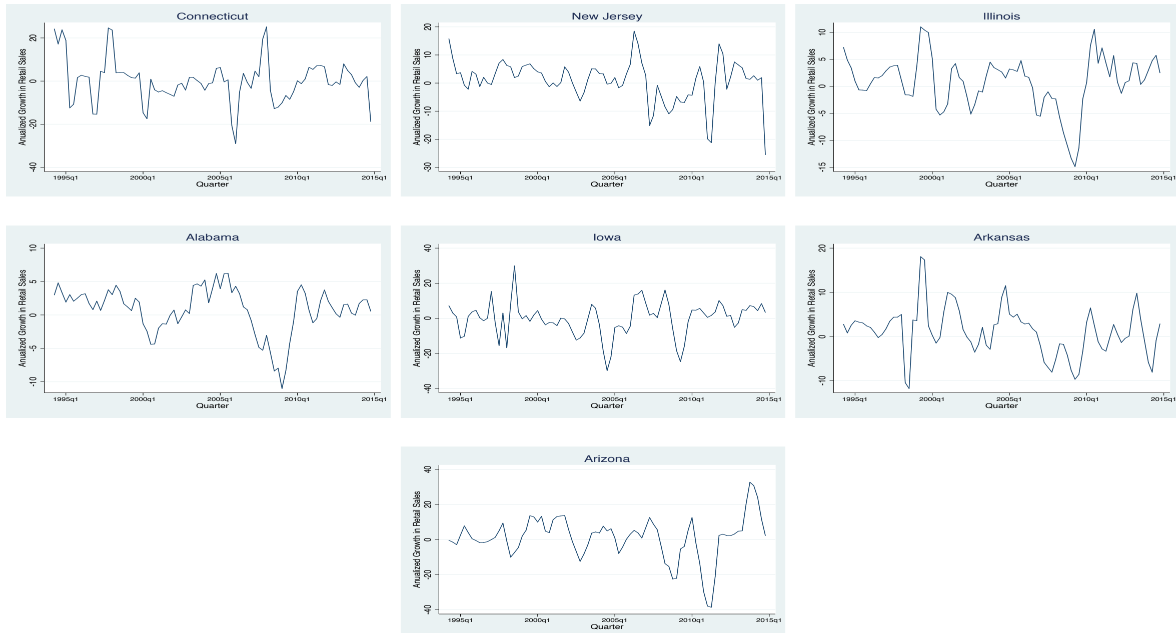

Figure 2.4: State Retail Sales

Notes: Each figure plots the annualized growth rate in state-level retail sales. The states are chosen as the first state alphabetically in the region for which the Conference Board tabulates its consumer confidence measure.

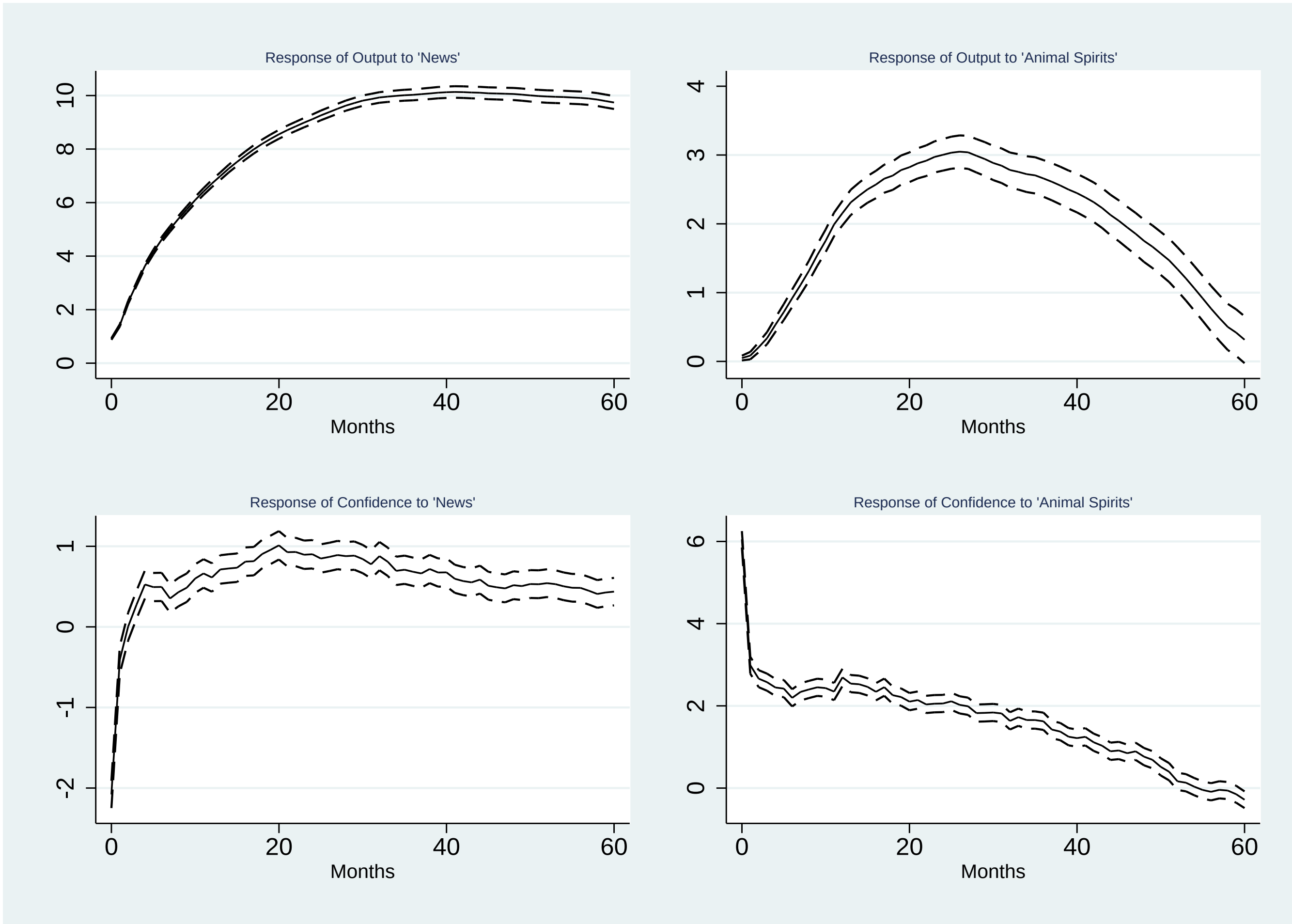

Figure 2.5: Impulse Responses of Output and Consumer Confidence

Notes: The figures in the top row plot the accumulated sums of $\{\beta_{1,j}\}$ (left) and $\{\beta_{2,j}\}$ (right) from the regression $\Delta y_{i,t} = \alpha_{1,i} + \delta_{1,t} + \sum_{j=0}^{T} \beta_{1,j}\epsilon_{NEWS,i,t-j} + \sum_{j=0}^{T} \beta_{2,j}\epsilon_{SPIRITS,i,t-j} + v_{1,i,t}$, where $\Delta y_{i,t}$ denotes annualized growth in the state economic activity index for state i in time t. The figures in the bottom row plot the sequences of $\{\beta_{3,j}\}$ (left) and $\{\beta_{4,j}\}$ (right) from the regression $Conf_{i,t} = \alpha_{2,i} + \delta_{2,t} + \sum_{j=0}^{T} \beta_{3,j}\epsilon_{NEWS,i,t-j} + \sum_{j=0}^{T} \beta_{4,j}\epsilon_{SPIRITS,i,t-j} + v_{2,i,t}$, where $Conf_{i,t}$ denotes the consumer confidence index in state i at time t. The frequency of the regression is monthly. The dashed lines represent 95% confidence bands.

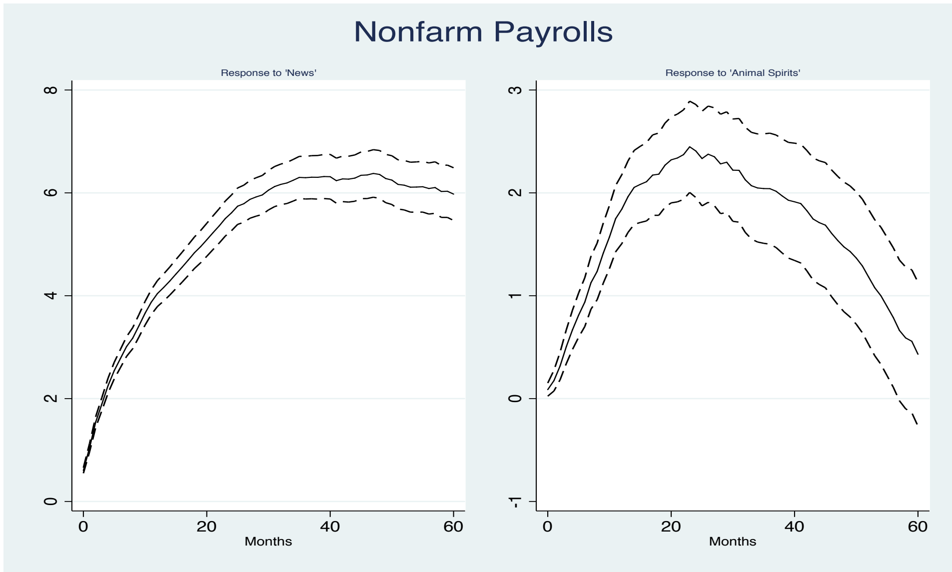

Figure 2.6: Impulse Responses of Nonfarm Payrolls

Notes: The figures plot the accumulated sums of $\{\beta_{1,j}\}$ (left) and $\{\beta_{2,j}\}$ (right) from the regression $\Delta nonfarm_{i,t} = \alpha_{1,i} + \delta_{1,t} + \sum_{j=0}^{T} \beta_{1,j}\epsilon_{NEWS,i,t-j} + \sum_{j=0}^{T} \beta_{2,j}\epsilon_{SPIRITS,i,t-j} + v_{1,i,t}$, where $\Delta nonfarm_{i,t}$ denotes annualized growth in the state nonfarm payroll employees for state i in time t. The frequency of the regression is monthly. The dashed lines represent 95% confidence bands.

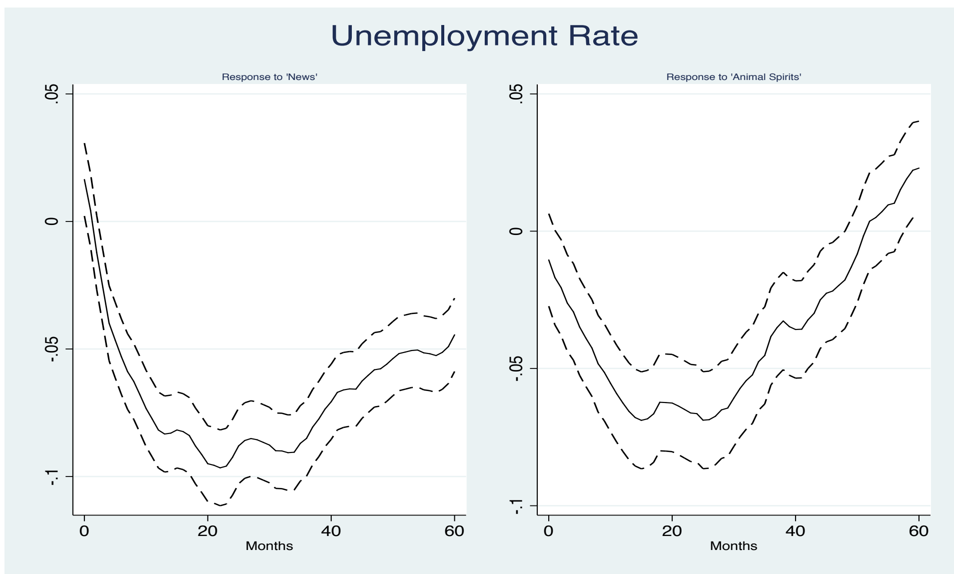

Figure 2.7: Impulse Responses of Unemployment Rate

Notes: The figures plot the accumulated sums of $\{\beta_{1,j}\}$ (left) and $\{\beta_{2,j}\}$ (right) from the regression $U_{i,t} = \alpha_{1,i} + \delta_{1,t} + \sum_{j=0}^{T} \beta_{1,j}\epsilon_{NEWS,i,t-j} + \sum_{j=0}^{T} \beta_{2,j}\epsilon_{SPIRITS,i,t-j} + v_{1,i,t}$, where $U_{i,t}$ denotes the unemployment rate in state i in time t. The frequency of the regression is monthly. The dashed lines represent 95% confidence bands.

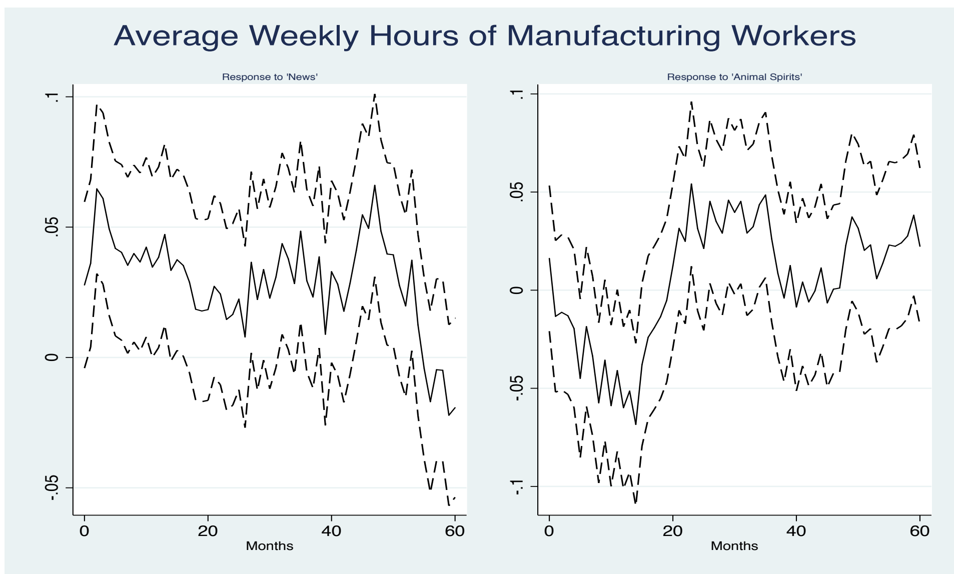

Figure 2.8: Impulse Responses of Average Weekly Hours of Manufacturing Workers

Notes: The figures plot the accumulated sums of $\{\beta_{1,j}\}$ (left) and $\{\beta_{2,j}\}$ (right) from the regression $N_{i,t} = \alpha_{1,i} + \delta_{1,t} + \sum_{j=0}^{T} \beta_{1,j} \epsilon_{NEWS,i,t-j} + \sum_{j=0}^{T} \beta_{2,j} \epsilon_{SPIRITS,i,t-j} + v_{1,i,t}$, where $N_{i,t}$ denotes average weekly hours for manufacturing production workers in state i in time t. The frequency of the regression is monthly. The dashed lines represent 95% confidence bands.

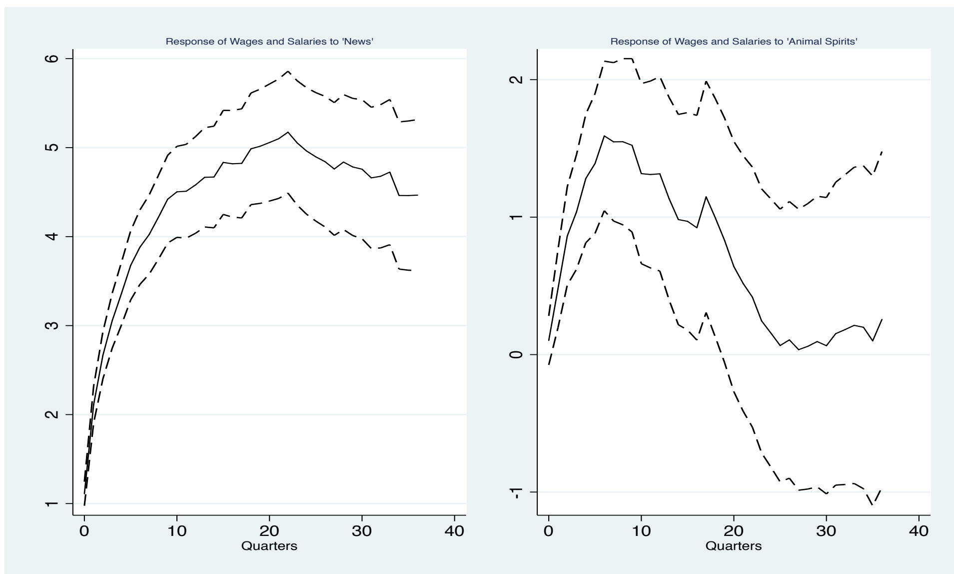

Figure 2.9: Impulse Responses of Real Wages and Salaries

Notes: The figures plot the accumulated sums of $\{\beta_{1,j}\}$ (left) and $\{\beta_{2,j}\}$ (right) from the regression $\Delta w_{i,t} = \alpha_{1,i} + \delta_{1,t} + \sum_{j=0}^{T} \beta_{1,j} \epsilon_{NEWS,i,t-j} + \sum_{j=0}^{T} \beta_{2,j} \epsilon_{SPIRITS,i,t-j} + v_{1,i,t}$, where $\Delta w_{i,t}$ denotes annualized growth in the wages and salaries deflated by the national consumer price index in state i in time t. The frequency of the regression is quarterly. The dashed lines represent 95% confidence bands.

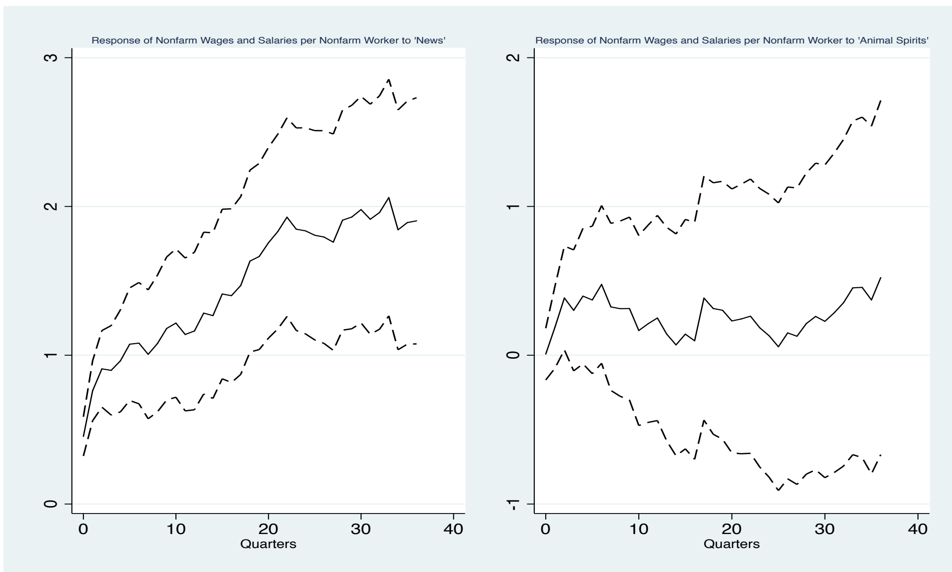

Figure 2.10: Impulse Responses of Real Wages and Salaries per Worker

Notes: The figures plot the accumulated sums of $\{\beta_{1,j}\}$ (left) and $\{\beta_{2,j}\}$ (right) from the regression $\Delta w_{i,t} = \alpha_{1,i} + \delta_{1,t} + \sum_{j=0}^{T} \beta_{1,j}\epsilon_{NEWS,i,t-j} + \sum_{j=0}^{T} \beta_{2,j}\epsilon_{SPIRITS,i,t-j} + v_{1,i,t}$, where $\Delta w_{i,t}$ denotes annualized growth in nonfarm wages and salaries per nonfarm worker deflated by the national consumer price index in state i in time t. The frequency of the regression is quarterly. The dashed lines represent 95% confidence bands.

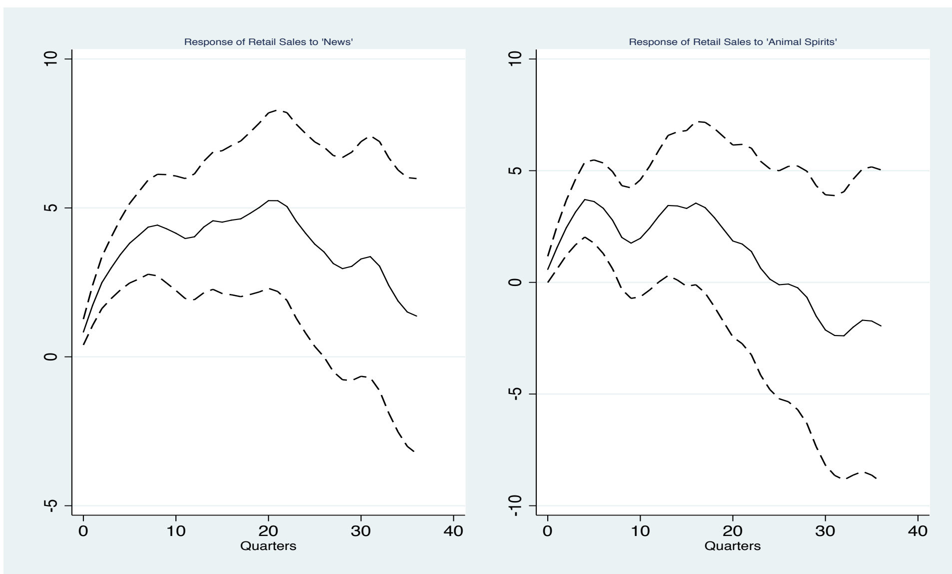

Figure 2.11: Impulse Responses of Retail Sales

Notes: The figures plot the accumulated sums of $\{\beta_{1,j}\}$ (left) and $\{\beta_{2,j}\}$ (right) from the regression $\Delta c_{i,t} = \alpha_{1,i} + \delta_{1,t} + \sum_{j=0}^{T} \beta_{1,j}\epsilon_{NEWS,i,t-j} + \sum_{j=0}^{T} \beta_{2,j}\epsilon_{SPIRITS,i,t-j} + v_{1,i,t}$, where $\Delta c_{i,t}$ denotes annualized growth in the state retail sales for state i in time t. The frequency of the regression is quarterly. The dashed lines represent 95% confidence bands.

Reference

Angelidis, Timotheos, Alexandros Benos and Stavros Degiannakis. 2004, “The Use of GARCH Models in VaR Estimation”. *Statistical Methodology*, 1(2), 105-128.

Angeletos, George-Marios and Jennifer LaO. 2013. “Sentiments”. *Econometrica* 81 (2):739779.

Arias, Maria A., Charles S. Gascon, and David E. Rapach. 2015. “Metro Business Cycles”. Working Paper, Federal Reserve Bank of St. Louis and Saint Louis University.

Bachmann, Rudiger and Eric R. Sims. 2012. “Confidence and the Transmission of Government Spending Shocks”. *Journal of Monetary Economics* 59 (3):235249.

Barsky, Robert B. and Eric R. Sims. 2012. “Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence”. *The American Economic Review* 102 (4):13431377.

Basu, Susanto, John G. Fernald, and Miles S. Kimball. 2006. “Are Technology Improvements Contractionary?”. *The American Economic Review* 96 (5) 14181448.

Benhabib, Jess, Pengfei Wang, and Yi Wen. 2013. “Uncertainty and Sentiment-Driven Equilibria”. NBER Working Paper 18878.

Benhabib, Jess, Pengfei Wang, and Yi Wen. 2015. “Sentiments and Aggregate Demand Fluctuations”. *Econometrica* 83 (2):549585.

Bera, Anil K. and Matthew L. Higgins. 1993. “ARCH Models: Properties, Estimation, and Testing”. *Journal of Economic Surveys* 7: 305366. doi:10.1111/j.1467-6419.1993.tb00170.x.

Blanchard, Olivier, Jean-Paul LHuillier, and Guido Lorenzoni. 2013. “News, Noise, and and Fluctuations: An Empirical Exploration”. *The American Economic Review* 103 (7):30453070.

Blanchard, Olivier and Danny Quah. 1989. “The Dynamic Effects of Aggregate Demand and Supply Disturbances”. *The American Economic Review* 79 (4):655673.

Bollerslev, Tim, 1986,“Generalized autoregressive conditional heteroskedasticity”. *Journal of Econometrics* Elsevier, vol. 31(3), pages 307-327, April.

Bollerslev, Tim, Ray Y. Chou and Kenneth F. Kroner. 1992. “ARCH modeling in finance: A review of the theory and empirical evidence”, *Journal of Econometrics*, Elsevier, vol. 52(1-2), pages 5-59.

Bollerslev, Tim. and Robert F. Engle, 1993, "Common Persistence in Conditional Variance" *Econometrica*, Econometric Society, vol. 61(1), pages 167-86, January.

Christiano, Lawrence J., Martin Eichenbaum, and Robert Vigfusson. 2004. "The Response of Hours to a Technology Shock: Evidence Based on Direct Measures of Technology. *Journal of the European Economic Association* 2 (2-3):381395.

Crone, Theodore M. and Alan Clayton-Matthews. 2005. "Consistent Economic Indexes for the 50 States." *The Review of Economics and Statistics* 87 (4):593603.

Engle, Robert F., 1982. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica*, Econometric Society, vol. 50(4), pages 987-1007, July.

Engle, Robert F. and Tim Bollerslev, 1986, "Modeling the persistence of conditional variances". *Econometric Reviews*, 5:1, 1-50, DOI: 10.1080/07474938608800095.

Enders, Walter. 2010. "Applied Econometric Time Series". New York, NY: John Wiley & Sons, Inc

Fernandez-Villaverde, Jes'us, Juan F. Rubio-Ram'rez, Thomas J. Sargent, and Mark W. Watson. 2007. "ABCs (and Ds) of Understanding VARs. *The American Economic Review* 97 (3):10211026.

Garrett, Thomas A., Rub'en Hern'andez-Murillo, and Michael T. Owyang. 2004. "Does Consumer Sentiment Predict Regional Consumption? *Federal Reserve Bank of St. Louis Review* 87 (2, Part 1):123135.

Huo, Zhen and Naoki Takayama. 2015a. “Higher Order Beliefs, Confidence, and Business Cycles. Working Paper, University of Minnesota and Cabinet Office of the Government of Japan.

Huo, Zhen and Naoki Takayama.. 2015b. “Rational Expectations Models with Higher Order Beliefs. Working Paper, University of Minnesota and Cabinet Office of the Government of Japan.

Keynes, John Maynard. 1936. “The General Theory of Employment, Interest, and Money. London, UK: The Macmillan Press.

Lorenzoni, Guido. 2009. A Theory of Demand Shocks. The American Economic Review 99 (5):20502084.

Lucas, Robert E. 1972. “Expectations and the Neutrality of Money. *Journal of Economic Theory* 4 (2):103124.

Lumsdaine, Robin L., 1996, “Consistency and Asymptotic Normality of the Quasi-Maximum Likelihood Estimator in IGARCH(1,1) and Covariance Stationary GARCH(1,1) Models, *Econometrica*, 64, 575-596.

Ma, Jun, 2008, “A Closed-Form Asymptotic Variance-Covariance Matrix for the Maximum Likelihood Estimator of the GARCH(1,1) Model, working paper, Department of Economics, Finance and Legal Studies, University of Alabama.

Ma, J., Charles R. Nelson, and Richard Startz, 2007, “Spurious Inference in the GARCH(1,1) Model When It Is Weakly Identified, Working Papers UWEC-2006-14-P, University of Washington, Department of Economics, revised Mar 2007.

Nelson, Charles R. and Richard Startz, 2007, “The Zero-Information-Limit Condition and Spurious Inference in Weakly Identified Models, *Journal of Econometrics*, Elsevier, vol. 138(1), pages 47-62, May.

Owyang, Michael T., Jeremy Piger, and Howard J. Wall. 2005. “Business Cycle Phases in U.S. States. *The Review of Economics and Statistics* 87 (4):604616.

Owyang, Michael T., David E. Rapach, and Howard J. Wall. 2009. “States and the Business Cycle. *Journal of Urban Economics* 65 (2):181194.

Pagan, Adrian. 1984. “Econometric Issues in the Analysis of Regressions with Generated Regressors. *International Economic Review* 25 (1):221247.24

Schmitt-Grohe, Stephanie and Martin Uribe. 2012. “Whats News in Business Cycles. *Econometrica* 80 (6):27332764.

Stock, James H. and Mark W. Watson. 1989. “New Indexes of Coincident and Leading Economic Indicators. NBER Macroeconomics Annual 4:351394.

The Conference Board. 2011. “Consumer Confidence Survey Technical Note.

Zivot, Eric, Richard Startz and Charles R. Nelson, 1998, "Valid Confidence Intervals and Inference in the Presence of Weak Instruments. *International Economic Review*, vol. 39(4), pages 1119-46, November.