Sentiments and Economic Activity:
Evidence from U.S. States

Jess Benhabib

New York University

and

Mark M. Spiegel*

Federal Reserve Bank of San Francisco

September 12, 2017

ABSTRACT

We examine whether sentiment influences aggregate demand by studying the relationship between the Michigan Survey expectations concerning national output growth and future economic activity at the state level. We instrument for local sentiments with political outcomes, positing that agents in states with a higher share of congressmen from the political party of the sitting President will be more optimistic. This instrument is strong in the first stage, and our results confirm a positive relationship between sentiments and future state economic activity that is robust to a battery of sensitivity tests.

J.E.L. Classification Number: E20, E32

Keywords: Sentiments, U.S. states, panel estimation, political economy

*Helpful comments were received from Simon Gilchrist, Christian Gillitzer, Oscar Jorda, Atif Mian, Jianjun Miao, Morten Ravn, Amir Sufi, Dan Wilson, Tao Zha, the SWUFE International Macro-Finance Conference in Chengdu, China, seminar participants at the Federal Reserve Bank of San Francisco, and two anonymous referees. Rebecca Regan and Ben Shapiro provided excellent research assistance. Our views are our own and not necessarily those of the Federal Reserve Board or the Federal Reserve Bank of San Francisco.

1. Introduction

There is some evidence of a contemporaneous correlation between measures of consumer sentiment and economic activity. The headline University of Michigan Index of Consumer Sentiment (ICS) has been shown to be closely correlated with growth in personal consumption expenditures over the postwar period [Carroll, et al (1994)]. This observed correlation in the data can be interpreted in different ways. It is possible that sentiments only reflect knowledge about current or future economic fundamentals. Some of the empirical literature has therefore concentrated on testing the restrictions that predict a causal link between sentiment changes and economic activity. Along these lines, there is also some evidence that sentiment measures unexplained by economic fundamentals are associated with spending shocks [e.g. Oh and Waldman (1990), Carroll, et al (1994), Starr (2012)]. However, the contribution of sentiment shocks "unrelated" to other measures of fundamentals has been found to be only temporary [e.g. Starr (2012)] and small [e.g. Ludvigson (2004)].[1]

Nonetheless, some theories suggest that shifts in consumer sentiments, like positive shocks to expectations concerning future output or future output growth, can indeed be "self-fulfilling," and therefore constitute multiple rational expectations equilibria. Indeed these sentiment-driven, self-fulfilling rational expectations equilibria can arise in various models of endogenous growth, in models of real business cycles with external effects, in models with collateral or borrowing constraints, or in search models with aggregate demand externalities, as well as in OLG models. Stochastic sentiments or sunspots can randomize locally over a continuum of rational expectations equilibria converging to an indeterminate steady state, or

[1] Most of these studies concentrate on the implications of sentiment changes on consumption. Permanent income theories however suggest that agents should spread the impact of improved economic prospects on their consumption over the course of their lifetime. Observed responses over longer time horizons may also be difficult to identify as new shocks emerge and fundamentals respond to changes caused by sentiment shocks.

alternatively across distinct multiple steady states.[2][3] Alternatively, even if the fundamentals-based equilibrium is unique, information frictions and incomplete markets can give rise to distinct sentiment driven stochastic rational expectations equilibria.[4]

Barsky and Sims (2012) distinguish between "animal spirits" shocks and news or information shocks using a VAR framework[5]. They only consider "animal spirits" shocks that are "erroneous" or irrational, and explicitly exclude sunspot shocks that are self-fulfilling under rational expectations. They argue that such animal spirits shocks unrelated to fundamentals are likely to have an immediate but transitory impact on economic activity. Therefore positive shocks to animal spirits are likely to look like positive aggregate demand shocks in the short run, but eventually will peter out if they are not followed by real increases in productivity. Using this assumption as an identification strategy in their VARs, they find that unexplained innovations to consumer confidence are the result of slowly building news about "apparently permanent" current and future economic fundamentals, rather than emanating from changes in "animal spirits" or self-fulfilling expectational shifts that can drive the economy across multiple equilibria or across multiple steady states.

---

[2] See for example Benhabib and Farmer (1999), Benhabib, Schmitt-Grohe and Uribe (2000, 2001), Benhabib and Wang (2013), Howitt and McAfee (1988), and Kaplan and Menzio (2013)].

[3] Self-fulfilling changes in consumer sentiment have also been identified in the literature in a number of alternative ways, as self-fulfilling prophecies [e.g. Azariadis (1981), Farmer (1999), herding [e.g. Blanchard (2016)], and animal spirits [e.g. Keynes (1936) and Akerlof and Shiller (2010)]. Here, unless indicated otherwise, we use changes in sentiments to describe changes in beliefs unrelated to fundamentals.[4] See for example Shell (1977) and Cass and Shell (1983), Maskin and Tirole (1987), Aumann, Peck and Shell (1988), and more recently Angeletos and La'O (2013)], and Benhabib, et al (2015).

[4] See for example Shell (1977) and Cass and Shell (1983), Maskin and Tirole (1987), Aumann, Peck and Shell (1988), and more recently Angeletos and La'O (2013)], and Benhabib, et al (2015).

This interpretation is based on the assumption that sentiment shocks or expectation shocks that do not reflect news about fundamentals only have temporary effects in equilibrium. However many economic models generate persistent stochastic fluctuations under rational expectations, driven by sunspot randomizations over multiple equilibria, or by correlated stochastic equilibria driven by sentiment or expectation shocks. Furthermore, the stochastic properties of such self-fulfilling equilibria depend, at least in part, on the stochastic processes driving the sunspots or sentiment shocks, in addition to model-specific internal propagation mechanisms for such shocks. The sentiment shocks themselves can be driven by general Markov processes with arbitrary degrees of persistence. Indeed, the fluctuations induced by sentiments can take the form of randomizations over multiple equilibrium paths around a locally indeterminate steady state, or alternatively of stochastic fluctuations across multiple steady states, all under full rational expectations. Depending on the model, these fluctuations can be in either levels or in growth rates.

In this paper, however, we do not focus on stochastic properties of sentiments themselves, which could be highly persistent, or alternatively could be temporary or i.i.d. Instead we focus on the empirical question as to whether consumer sentiments directly influence output levels over certain time horizons. It is of course quite possible that sentiment shocks that affect output growth rates are indeed temporary, and our analysis does not exclude this possibility. What we demonstrate instead is that sentiment or consumer confidence shocks at the state level do have a measurable impact on state outputs and consumptions, certainly at one year horizons, and possibly over longer horizons.

We concentrate on U.S. state economic activity as the focus of our analysis. We examine the responses of overall state economic activity to changes in sentiment about national economic conditions. Our identification strategy relies on the notion that changes in local sentiment about national economic prospects are likely to induce local changes in consumption and investment expenditures. As both consumption and investment expenditures are in play, a focus on aggregate demand and overall economic activity seems useful. Moreover, the output response is useful as a guide to the implications of sentiment shocks for optimal stabilization policy.

The use of state data also allows us to condition for aggregate shocks, facilitating the identification of the direct impact of sentiment changes on economic activity at the state level.

We use Michigan Surveys (2016) questions concerning national economic conditions. Our base specification uses the question, "Looking ahead, which would you say is more likely -- that in the country as a whole we'll have continuous good times during the next 5 years or so, or that we will have periods of widespread unemployment or depression, or what?" Our maintained hypothesis is that states are sufficiently small that attitudes about the local economy will not distort the response about national economic conditions. Given this assumption, our cross-sectional treatment should isolate the impact of differences in sentiment across states on future differences in state economic activity.

While sentiment data is available at the county level, it seems plausible that substantive leakage is likely to occur across county lines. This is of course a compromise, as leakages across state lines are also likely to take place, but they are much less likely to be prevalent than those across county lines.

A potential problem with our specification is that household expectations about future national economic activity may be positively related to local experiences, raising the prospect of reverse causality in our empirics. We test for this possibility by examining the impact of past state growth on current sentiment. We do find evidence of such a relationship in the data, as our coefficient of interest enters at statistically significant levels. We respond to this reverse causality challenge using instrumental variables estimation. Instrumental variables estimation will also serve to address the also-likely issue that answers to the Michigan Survey are noisy measures of consumer sentiment.

We turn to political data as an instrument for local sentiment levels that vary systematically across states. There is a large literature that demonstrates a positive relationship between partisanship and economic assessments. A survey respondent that self-identifies as a member of one of the major political party is more optimistic about the national economic picture when the sitting national leader is from that same party. In an early paper, Gerber and Huber (2009) demonstrate that consumption changes following a political election are correlated with whether or not the election was won by the preferred political party of the respondent. They interpret this correlation as working through the sentiment channel.

Political partisanship has also been used as an instrument for identifying a connection between sentiment and consumption. Mian, et al (2015) demonstrate that presidential elections are associated with changes in sentiment about the effectiveness of government policy in line with political partisanship. However, they find no statistically significant relationship between changes in the presidential party at the county level in the United States and changes in actual consumption.[6] In contrast, Gillitzer and Prasad (2015) show in Australian survey data that higher sentiment is associated with having a member from your political party in office at the federal level. These changes in sentiment associated with elections are also associated with increased future vehicle purchase rates.

[6] Mian, et al (2015) examine the cases of the 2000 and 2008 elections. While in neither case do they find evidence of significant changes in consumption at the county level, they do find a significant correlation between the 2008 election outcome and planned consumption measures consistent with the predictions of a partisanship model of politically-driven sentiment changes.

One distinction that favors the Australian study is that Australian survey data has a direct question about political affiliation. U.S. consumer data, such as that used by Mian, et al (2015) - and also used in this study – require proxies for political partisanship. The Mian, et al study uses county-level data on voting in presidential elections. Below, we use the share of state congressional representatives from the same political party as the sitting president. The latter proxy has the advantage of changing every two years, yielding more variability in our sample and allowing us to use our full panel sample. This is desirable because our use of state-level data to mitigate consumption leakages across counties results in a smaller cross-section than the Mian, et al (2015) county-level study.

We demonstrate below that our proxy is a strong instrument in the first stage of our IV specification, and that the instrumented measure of consumer sentiment is shown to be positively and statistically significantly associated with State output growth over the following four quarters.

These results are shown to be robust to the inclusion or exclusion of state and time fixed effects, as well as a variety of sensitivity tests. These include weighting observations by either state size or the number of respondents, or changes in the sample population, dropping specific time periods, states with exceptionally high or low incomes, investment levels or populations, or dropping states identified as outliers based on residual values.[7] We also examine robustness to the use of conventional, rather than heteroskedasticity-corrected standard errors, random instead of fixed effects, and regional instead of state dummies. The results are also robust to a change to

[7] We do find that our results disappear when we drop the pre-crisis portion of our sample. However, this appears to be the result that we have a relatively small number of non-crisis years available in our sample. As we discuss below, our results are also weaker when we drop the post-crisis years (and strengthen when the crisis years are dropped). It appears that the relationship we identify between sentiment and economic activity broke down during the crisis period.

annual frequencies, and to the inclusion of lagged state and national output growth in our specification.

We also examine other sentiment questions. Here our results are more mixed, but they remain relatively robust for most alternative sentiment measures. We also consider cnthe impact of sentiment shocks on state consumption, in the form of personal consumption expenditures at the state level. Sentiment is found to play a positive role in the determination of consumption a year out of sample as well.

Finally, we consider longer-horizon sentiment impacts. We repeat our base specification to investigate the impact of sentiment on state activity and personal consumption expenditures over 2 and 3 year horizons. Our results for these longer horizons generally remain strong for both activity and consumption effects, but we find that our longer horizon activity results lose their significance when we include time dummies and cluster our standard errors by state.

The remainder of this paper is organized into 7 sections. The following section introduces our data summary statistics. Section 3 discusses our IV methodology. Section 4 describes our base specification results. Section 4 subjects these results to a battery of robustness tests. Section 5 discusses our IV specification and results. Section 6 reports our results for longer horizons. Section 7 concludes.

2. Data

Quarterly sentiment data is obtained from 2004 through 2015 from the University of Michigan Surveys of Consumers (2017). Our base gauge of consumer sentiment is the answer to question BUS5 in the survey, “Looking ahead, which would you say is more likely -- that in the country as a whole we'll have continuous good times during the next 5 years or so, or that we will have periods of widespread unemployment or depression, or what?" Respondents' answers are scored 1 through 5, with 1 representing the answer "Good times," 2 representing "Good with qualifications," 3 representing "Pro-Con," 4 representing "Bad with Qualifications," 5 representing "Bad Times." There are also a modest number of responses characterized as "depends," scored 6 through 99 based on the response. The distribution of responses for the entire sample is shown in Figure 1. It can be seen that extreme responses of 1 or 5 are most common.

Figure 2 displays the relationship between national sentiment and national economic activity. As in much of the literature, sentiment appears to track current economic activity closely. For example, it is clear that sentiment declines sharply in tandem with the onset of the Great Recession. Still, sentiment does not track activity perfectly. Sentiment reaches its lowest level in 2011, reflecting volatility in financial markets associated with the euro area debt crisis. Needless to say, while there is a decline in output at this time, it does not match that experienced during the Great Recession. During that period, sentiment appears to have held up on average while the US economy fell into recession, and then continued to fall after the recession had ended. The great recession periods is notable, as sentiment appears to track activity much more closely both before and after the event.

It is also difficult to draw any causal inferences from this national picture. Changes in sentiment may be following national economic conditions, rather than leading them. For that reason, we turn to state activity data for identification. Reassuringly, there appears to be no apparent cross-sectional pattern to state responses, either by income, education, geography, or for the purposes of our IV specification below, political partisanship. In the latter case, note that our sample period spans years with sitting Presidents from both political parties.[8]

As our base measure of lagged sentiment, *GOOD*, we consider the share of a state *i* at time *t-4* whose respondents' answers were scored 1 or 2.

We include other variables obtained from the Michigan Surveys to condition on the characteristics of individual respondents. As our observation is at the state level, these are measured as state respondent averages, also at time *t-4*. Our conditioning variables include income levels by state, *INCOME,* which is calculated as the average of reported levels of respondent incomes within a state, *EDUC*, which is the average of the highest year of education reported by respondents within a state, and *INVEST*, which is the share of state respondents who said that they hold investments. Growth at the state level from period *t-4* to *t*, *GGDP*, is obtained from Haver analytics, as is our measure of the national output gap, *YGAP*.

Summary statistics are shown in Table 1. It can be seen that there is a lot of variability in the data in both growth and sentiment measures, unsurprising since our sample includes the Great Recession period as well as the boom that preceded it. The final three columns show average growth rates in our pooled sample for states exhibiting high (more than one standard deviation above the mean) sentiment levels, *HGGDP,* neutral (within one standard deviation of the mean) sentiment levels, *MGGDP,* and low (more than one standard deviation below the mean), *LGGDP,* sentiment levels. As expected, it can be seen that subsequent growth on average is higher following reports of high sentiment levels, and lower for states where low

[8] Appendix Table A1 displays average consumer sentiment over our sample by state. For our sample period, respondents in the state of Virginia are most optimistic about future national economic activity on average, while those from New Mexico proved the most pessimistic on average.

levels of sentiment on average are reported. However, t-tests for differences in these populations are not statistically significant.

3. Methodology

3.1 Base specification

Our base specification is a conventional panel estimator:

$$\Delta y_{it} = \alpha + \beta GOOD_{it-4} + \gamma X_{it-4} + \{\delta_i\} + \{\varepsilon_t\} + \eta_{it} \quad (1)$$

where Δy_{it} represents income growth in state i from period $t-4$ to the present, $GOOD_{it-4}$ represents the share of respondents with positive sentiment responses in state i in period $t-4$, X_{it-4} is a vector of controls linked to state growth via a set of nuisance parameters γ, $\{\delta_i\}$ and $\{\varepsilon_t\}$ are respectively state and time-specific fixed effects, and η_{it} is a is a residual, assumed to be well behaved. Our coefficient of interest is β, the partial-correlation between sentiment and subsequent state income growth. We use three covariates in X_{it-4} to control for other determinants of state growth available from the respondent survey, including, *INCOME*, *EDUC*, and *INVEST*, all described above.

We consider two alternative methods for conditioning for prevailing economic conditions. First, we include the start-of-period output gap, $YGAP_{t-4}$. Alternatively, we include yearly time dummies, $\{\varepsilon_t\}$. Quarterly time dummies are included below in our robustness checks.[9]

[9] We use yearly time dummies because estimating the full set of quarterly dummies results in a substantial loss of degrees of freedom. Still, as we show in Table 5, our results are robust to the inclusion of quarterly dummies, with our coefficient of interest retaining significance at least at a 10% confidence level for all estimation methods.

It is also quite possible that our data may exhibit heteroscedasticity and correlations within and across state groups. We therefore use heteroscedasticity-corrected standard errors throughout, and also allow for cross-sectional or state-specific dependence. For cross-sectional dependence, which is likely across states in our panel, we use Driscoll-Kraay (1998) estimators. Hoechle (2007) demonstrates that Driscoll-Kraay standard errors are well-calibrated when cross-sectional dependence is present. For state-specific dependence, we include state dummies and cluster by state.

For those specifications that include a comprehensive set of both time and state-specific fixed effects, our specification can be interpreted as a difference-in-differences estimator of the impact of changes in the share of agents within a state holding positive sentiment about future national economic prospects or not.

3.2 Endogeneity issues

One potential challenge in identifying a causal relationship between sentiment and activity is that opinions about future national economic performances may be based on individual experiences, and hence tied to the fortunes of the local economy. [10] This would render a least-squares identification strategy invalid. To test if this is the case, we first examine the relationship between current state economic growth and current sentiment levels. Our reverse specification satisfies

---

[10] We do find positive, albeit modest, correlations in the survey between responses to our national outlook question and individual household experiences and expectations. The estimated correlation coefficient between positive responses to the national economic outlook and positive responses to the question about whether a household's financial condition is better or worse off than it was a year ago (question PAGO in the survey) is 0.32, while the estimated correlation coefficient between expected future national economic conditions and expected household financial conditions five years in the future (question PEXP5 in the survey) is 0.23.

$$GOOD_{it} = \alpha + \beta \Delta y_{it} + \gamma X_{it-4} + \{\delta_i\} + \{\varepsilon_t\} + \eta_{it} \tag{1.1}$$

where variable definitions are the same as in our base specification and we include the same nuisance covariates in X_{it-4} to control for other determinants of state growth available from the respondent survey, *INCOME, EDUC*, and *INVEST*. We run this specification using ordinary least squares, with the same variety of error specifications as those in our base specification.

Our results are shown in the appendix of this paper (Table A2). The data do appear to indicate that sentiment is associated with local activity, although statistical significance is lost when time dummies are substituted for the national output gap. Overall, however, there does indeed seem to be a risk of reverse causality.

We address this potential issue through instrumental variables (IV). We follow the literature in turning to political data as an instrument for differences in sentiment levels that vary systematically across regions. Our posited relationship is that survey respondents will be more optimistic about national economic prospects if the sitting president is from his or her political party.

Sentiment has been shown in the literature to correspond to economic activity, as in Gerber and Huber (2009), who identify a positive relationship between partisanship and economic activity, as consumption changes following a political election are correlated with whether or not the election was won by the preferred political party of the respondent.

We expect that the primary channel through which partisanship can affect economic activity is through changes in sentiment. A positive relationship between political partisanship and economic sentiment has been demonstrated in the literature. Mian, et al (2015) find that presidential elections are associated with changes in sentiment about the effectiveness of government policy in line with political partisanship. However, they find no statistically significant relationship between changes in the presidential party at the county level in the United States and changes in actual consumption. In contrast, Gillitzer and Prasad (2015) show in Australian survey data that higher sentiment is associated with having a member from your political party in office at the federal level. Changes in sentiment associated with elections are shown to be associated with increased future vehicle purchase rates.

One distinction that favors the Australian study is that Australian survey data has a direct question about political affiliation. U.S. consumer data, such as that used by Mian, et al (2015) have to use proxies for political partisanship. Their study uses county-level data on voting in presidential elections. Our study using U.S. data faces the same challenge. To proxy for political partisanship at the state level, we use the share of state congressional representatives from the same political party as the sitting president, which we term *CONGPRES.* Our proxy has the advantage of changing every two years, with each congressional election, and therefore yields more variability in our sample.

We follow this literature in using changes in partisan fortunes as an instrument for sentiment changes. However, such an instrument would be invalid if the political situation directly affected underlying economic fundamentals. In particular, it is possible that states with a higher number of congressional representatives from the same political party as the sitting president will be favored in political outcomes in a manner that directly supports local economic conditions. For example, decision about military base closures may be made in geographically partisan manners.

Mian, et al (2015) provide two pieces of evidence against this possibility. First, they look at income growth in U.S. counties before and after Presidential elections. They find no evidence that Presidential elections are systematically related to changes in county growth in manner associated with local political leanings. Second, they also find no relationship between election outcomes and changes in government transfers to localities.

We provide additional evidence against this potential problem below, by adding taxes and transfers from the federal government to our base specification. As in these previous studies, our demonstrated relationship between sentiment and economic activity is robust to conditioning for federal government economic transfers.

4. Base specification results

We first examine the first stage of our IV specification to demonstrate that we have a strong instrument. We include our base specification conditioning variables as well. Our panel results are shown in Table 2. We consider six variations: Models 1 through 3 include the lagged output gap, while models 4 through 6 include annual time dummies. Models 1 and 4 use robust standard errors, models 2 and 5 use the Driscoll-Kraay estimators, and models 4 and 6 allow for clustered standard errors by state. We include state fixed effects throughout.

Our variable of interest, *CONGPRES,* consistently enters significantly with its expected positive sign at a 1% confidence level, indicating significant correlation with our sentiment variable. Our point estimate for this variable is also largely invariant to perturbations in our specification. We therefore conclude that we have a strong instrument and proceed with our IV estimation using a two-stage least squares approach.

The second stage results are shown in Table 3. It can be seen that our variable of interest, *GOOD,* enters significantly positively. The instrumented coefficient point estimate in our IV specification is large. With the income gap included, it comes in at 0.19 and with state fixed effects included it comes in at 0.13.

Overall, the data support a non-trivial sentiment channel for differences in economic growth across US states. Indeed, if anything, these point estimates appear to be too high. Under our Model 1 sample, a one standard deviation increase in sentiment would be associated with an 3.6 percentage point increase in state output growth in Models 1 through 3, and a 2.5 percentage point increase in Models 4 through 6. [11] However, the 95% confidence intervals for these coefficients allow a one standard deviation increase in sentiment to be associated with an increase in state output of as low as 1.1 percentage points, a much more plausible figure, for our base specification with robust standard errors and the income gap included (Models 1-3), and as low as 75 basis points with robust standard errors and state fixed effects included (Models 4-6).

Our point estimate for the coefficient of interest, changes in the sentiment share, with the output gap included is equal to 0.19 with the income gap included, implying that a one standard deviation increase in our sentiment measure would be associated with a predicted increase in annualized state GDP growth of approximately 3.6 percent. This seems to be a non-trivial effect, but one that is not too large to be plausible. Our point estimate with time dummies instead of the output gap is qualitatively similar, albeit modestly smaller at 0.13.

All of the specifications are statistically significant, at a 1% confidence level using White heteroscedasticity-robust standard errors, at a 1% confidence level, using Driscoll-Kraay standard errors, and at a 5% confidence level with standard errors clustered by states using the output gap and at a 10% confidence level with clustered standard errors with time dummies instead of the output gap.

[11] The estimated standard deviation of the variable *GOOD* in our base sample is equal to 0.19.

The performances of the conditioning variables are mixed. Only the *INCOME* variable enters with statistical significance, and then only in our output gap specifications (Models 1 through 3).

5. Sensitivity analysis

In this section, we demonstrate that our base results, that state sentiments about national economic prospects have a direct impact on future state output, are quite robust. We first demonstrate that our base specifications results are largely robust to a wide variety of sample perturbations (Table 4).

For each sample perturbation, we report the point estimate and standard error for the coefficient of interest, *GOOD*, for the six Models in our base specification in Table 3. It can be seen that the variable of interest is robust to a large majority, but not all permutations. First, we drop the financial crisis period, which we interpret as spanning from 2007Q4 to 2009Q2. It can be seen that our results get stronger with the removal of the crisis period, as we experience increases in point estimates and significance for all specifications.[12] The increased significance when the crisis period is dropped -- which indicates that our results are not driven by the crisis -- suggests that our overall results are likely to strengthen as more data becomes available.

We next drop observations from "high" (more than one standard deviation above the mean) and "low" (more than one standard deviation below the mean) average state income levels. We do the same for high and low reported share of households with investments, and

[12] We also tried dropping the first and last 8 quarters of our sample period, but experienced more sensitivity. In particular, the GOOD variable of interest drops out with the early sub-sample omitted for all specifications. The results are more robust with final 8 quarters dropped, entering at a 5% confidence level for all output gap specifications, but only at a 10% confidence level for the White and Driscoll-Kraay specifications with time dummies, and narrowly misses significance at a 10% level with standard errors clustered by state. These results are available from the authors on request.

states with large and small GDPs. Finally, we drop outlier observations, measured as those with residuals more than two standard deviations above or below zero in our base specification.

Our base specification results are robust to dropping high or low incomes, high and low investments, and large GDP states for all specifications. However, we do observe some sensitivity to either dropping small GDP or outlier observations. Our results for both sample perturbations remain significant at least at a 10% level with White heteroscedasticity standard errors for our base specification with the output gap included, but both alternative samples drop out with Driscoll-Kraay standard errors and only our drop outliers alternative sample is significant with standard errors clustered by state.

Table 5 considers the robustness of our results to changes in estimation methodology and specification. First, given our sample sensitivity to dropping small GDP states, one might be concerned that averages of sentiment responses from larger states might be more informative, as these are taken from a larger sample of individual responses. We respond with two types of weighted least squares estimators, weighting by both state GDP and the number of Michigan Surveys respondents in the state for that time period. The results for the six specifications with these weighting schemes for the sample are in the first two rows of Table 5.

Weighting in either manner leaves our coefficient estimate for the variable of interest significant for all model specifications, with the majority entering at least a 5% confidence level. Moreover, the change does not qualitatively change our point estimate for output gap specifications. It falls modestly when weighted by state GDP and marginally rises when weighted by the n umber of survey responses. We observe larger point estimate declines using time dummies, where weighting by GDP and the number of responses yield 0.11 and 0.10

point estimates respectively. These lower point estimates seem more plausible, as a one standard deviation increase in the *GOOD* variable is predicted here to result in 2.1 and 1.9 percentage point increases in state GDP growth respectively.

We also consider conventional, rather than robust standard errors, random, rather than fixed effects, quarterly, instead of annual, time dummies, and regional dummies instead of state fixed effects. All specifications continue to enter at standard confidence levels.

We next consider annual, rather than quarterly data. Our annual data is available a year earlier, and so our sample spans from 2005-2016 and has 588 observations. Our coefficient estimates for the annual frequency are almost identical to those we obtain with quarterly data, and our variable of interest remains statistically significant for all specifications.

The use of annual data allows for the consideration of lagged GDP under the same time period as our base specification. We consider the inclusion of lagged state GDP, lagged US GDP, and both variables. All specifications enter with qualitatively similar values on our *GOOD* variable of interest, and remain statistically significant. The coefficient point estimates are modestly smaller, but still indicative of a substantive impact of positive sentiment on output. Given a one standard deviation increase in our *GOOD* variable, our point estimate for our base specification indicates a 2.3 percentage point increase in state output growth.[13]

Finally, as we discuss above, our instrument may be rendered invalid if the political situation directly affects underlying economic fundamentals. In particular, if states with a higher number of congressional representatives from the same political party as the sitting president are

[13] We lose a year of data when adding lagged state GDP at a quarterly frequency, leaving our time series only spanning 2007-2016. Our variable of interest, *GOOD*, is positive but insignificant with this truncated sample. However, we run the specification under our original sample by linearly interpolating our 2005 annual state output data. Using this sample, the statistical significance for the variable of interest is restored. This alternative sample was provided to the referees and is available on request from the authors.

favored in political outcomes that support local economic conditions. We therefore also condition for federal taxes and transfers in Table 5.

It can be seen that our results are qualitatively the same with this variable included, as our point estimates are roughly the same and enter significantly for all specifications. Given this evidence, we proceed under the assumption that the sentiment channel is the only channel through which the political characteristics of a state influence its economic activity, rendering our IV specification valid.

Table 6 considers several alternative sentiment measures to the Michigan Surveys as well as the impact of sentiment on consumption. Among the alternative sentiment measures, we first we consider negative responses to the question about future national economic conditions, i.e. those that answered response "5" to the question above. We term this variable *"BAD"*. Second, we consider the share of respondents in state *i* at time *t* that answered the question above that national economic conditions over the next five years would be "good," without qualifications, i.e. with responses that were coded "1" to the question above. We term this variable *"GOOD1."* We also consider responses to question BAGO (109) which asks whether business conditions are better or worse than the previous year, which we term *"BETTER"*. Finally, we use the sentiment measure studied by Mian, et al (2015) on the quality of government performance, which we term *"GOVT."* We measure this variable as the share of respondents who answered "1," indicating that they thought that the government was doing a "good job" in its economic policy.

Our results for these alternative sentiment measures continue to enter with their expected signs, but their performances in terms of statistical significance is uneven. Our results for negative sentiment responses, *"BAD5,"* universally enters negatively at statistically significant levels. Our point estimates for this variable indicate a comparable response to what we saw for our positive sentiment measure. Given the standard error for this variable in our sample of 0.19, our point estimate of, for example, our base specification with the income gap included and robust standard errors (Model 1) implies that a one standard deviation increase in the share of negative sentiment is associated with a 2.7% percentage point decrease in state output growth.

Our coefficient estimates for the *GOOD1* variable are larger than those in our base specification above (which is unsurprising as it is on average a more positive reponse), and are universally statistically significant at standard confidence level for all of our specifications.

Our results are not as strong over a shorter horizon, as the impact of those with higher expectations concerning business conditions a year from now, *BETTER*, consistently obtains a positive coefficient estimate, but is statistically insignificant for our output gap specifications, and is also insignificant for the clustered regression with time dummies included.

We obtain stronger results for the specifications where sentiment is measured in terms of attitudes about the government's performance, *GOVT*. These enter significantly for all of the specifications except Model 6, which clusters and includes time dummies. In that specification, the variable of interest just barely misses a 10% confidence level of statistical significance.

All of the alternative measures considered enter with their expected signs in all of our specifications. Not all of these are statistically significant for our sentiment variable of interest, but then, neither were the specifications with our base sentiment measure in Table 3. Overall, then, while we acknowledge some sensitivity to the sentiment measure used, our results continue to indicate a positive (negative) impact on state output with higher (lower) measures of sentiment about future economic activity used.

Finally, a number of papers in the literature [e.g. Gillitzer and Prasad (2015)], examine the implications of sentiment shocks for consumption, rather than income. These need not go together, as positive responses in income to sentiment shocks may more reflect movements in local investment, rather than consumption. We therefore look at the impact of sentiment shocks on consumption, as measured by movements in personal consumption expenditures. Data for this variable is available by state only at an annual frequency.

Our results are also shown in Table 6. It can be seen that we find a statistically significant and positive response by state in consumption to sentiment differences as well. Moreover, the point estimates for our specifications appear to indicate responses of a plausible and economically significant magnitude, as a one standard deviation increase in *GOOD* is predicted to result in a 2.3 percentage point increase in PCE consumption.

6. Impact on activity over longer horizons

Finally, we consider the impact of sentiment over longer horizons. We redo our base specifications with a two-year lag for sentiment. As the start of our quarterly sample is constrained by data availability for quarterly growth by state, we estimate over these longer lags at an annual frequency over our base sample.[14] Our dependent variable is now $GGDP_{it,t+2}$, average annual state growth from period *t* through period *t+2*.

Our results for the variable of interest, *GOOD*, over a two-year horizon are shown in Table 7. Our base specifications estimated using OLS are shown in the first row. Our specifications are significant at standard confidence levels for all methods of estimating standard errors with the output gap included (models 1 through 3). The coefficient point estimates are a

[14] Recall that coefficient point estimates for our variable of interest are very close for quarterly and annual frequencies.

little smaller, as would be expected over a longer horizon, but still indicate a substantive impact on state growth from changes in sentiment.

The results are more mixed for growth over a two year horizon with state fixed effects included. The coefficient point estimates are smaller, but still indicate a notable impact on growth. However, while we obtain significant coefficient estimates with robust standard errors, our results over a two-year horizon are weaker using either Driscoll-Kraay or clustered standard errors. As above, we have some reservations about our standard error estimates for these clustered specifications because our panel is small in the time dimension. Our point estimates suggest that under our sample a one standard deviation increase in sentiment is associated with an 1.6 percentage point increase in average state growth over the next two years for the specifications with the income gap included (Models 1 through 3), and even smaller effects with state fixed effects. However, a 95% confidence interval would include values similar to our point estimates for the one year horizon.

The point estimates for average annual growth are even smaller and less significant over a three year horizon, as might be expected. Our point estimates suggest that under our sample a one standard deviation increase in sentiment is associated with a still-substantial 0.9 percentage point increase in average state growth over the next two years for the specifications with the income gap included (Models 1 through 3), and even smaller effects with state fixed effects included. Again, a 95% confidence interval includes values similar to our point estimates for the one year horizon. Moreover, it is important to remember that as the impact period increases in length, the probability that further unmeasured sentiment shocks took place increases, which adds noise to our analysis.

Our estimates over the 3-year horizon for state growth are in the second row. These results appear to be weaker, as our point estimates fall, for example to 0.05 in the case of output gap specifications (Models 1 through 3), and we only obtain statistically significant coefficient estimates at standard confidence levels under robust standard error estimation. Our coefficient estimates for models 1 through 3 suggest that a one standard deviation increase in sentiment results in a 1.0 percentage point increase in average annual output growth over a three year horizon.

Our results suggest a more modest, but still positive impact of sentiment shocks over these longer term horizons. Still, we acknowledge that our results over these longer –term horizons are weaker than those we found over a 1 year horizon, indicating that the positive impact of state sentiment shocks on output growth that we find in the data appears to be temporary.

We next turn to longer-term impacts on PCE consumption by state, which is only available at an annual frequency. Our results for two and three year horizons are shown in rows three and four of Table 7. It can be seen that our point estimates are modestly lower, but all of our specifications over both longer horizons are statistically significant at standard confidence levels.

Overall, our results confirm a longer-term relationship between sentiment and future consumption at the state level is persistent. We also obtained positive, but weaker, estimates of the implications of sentiment shocks for state economic activity.

7. Conclusion

We revisit the relationship between consumer sentiment and economic activity. Our identification strategy is based on using the cross-sectional information in state data. We examine individual responses to survey questions about long-term prospects for the national economy. If sentiment drives activity, states in which agents hold more optimistic outlooks about national economic prospects should undertake higher levels of economic activity: Sentiments about economic prospects can thereby affect output at the state level.

A potential problem with our strategy is that it is possible that agents' responses to questions about future national economic conditions may reflect local conditions. Our reverse regression results suggest that reverse causality along these lines may indeed be an issue, although formal Hausman tests do not indicate endogeneity at statistically significant levels. To address this potential problem, we turn to IV estimation, based on a predicted relationship between political partisanship and economic sentiment concerning national economic prospects. Our results demonstrate a strong first stage relationship between our instrument and our sentiment measure. Our instrumented sentiment measure confirms a statistically significant relationship between state sentiment about national economic prospects and next year's state economic activity under a variety of specifications. Our point estimate under our base specification indicates that a one standard deviation increase in sentiment is associated with additional 49 basis points of growth in the following year. This result is robust to a wide variety of robustness tests, including sample perturbations, changes in estimation methods, and the use of alternative sentiment measures. Moreover, we find a significant and robust relationship at annual frequency between state sentiment and next year's PCE consumption at the state level.

We also consider the impact of sentiment over two and three year horizons. Our results over these longer horizons are weaker for state output, always entering positively but except for the specifications with robust standard errors, only statistically significantly at a two-year horizon for our specifications with Driscoll-Kraay or clustered standard errors with the national output gap included instead of state fixed effects. At the three year horizon, our point estimates are smaller and only enter at statistically significant levels for our robust standard error estimates.

Our results for longer horizons with state PCE consumption are stronger. All of specifications enter positive at statistically significant levels, regardless of model or estimation method. Still, our point estimates diminish with horizon, indicating that our data indicate a persistent positive one, but not necessarily a permanent one.

Our overall results therefore support the notion of a positive empirical relationship between sentiments and future economic activity, as well as future consumption expenditures. We find weaker and sometimes insignificant impacts of sentiment shocks in economic growth at the state levels at longer than two and three year horizons, but robust evidence of measurable impacts on consumption at those horizons.

Still, under our specifications it is possible that new sentiment shocks that occur late under the longer horizons add noise to our sample and preclude finding statistically significant results. Therefore, we remain skeptical of rejecting the possibility of measurable and significant sentiment effects on output at longer horizons because of the weaker results we obtain under clustered standard errors. Our quarterly sample is by necessity small in the time dimension, and with longer samples we may obtain stronger results over time.

References

Akerlof, George A. and Robert J. Shiller, (2010), *Animal Spirits: How Human Psychology Drives the Economy, and Why It Matters ...*, Princeton University Press.

Angeletos, George-Marios and Jennifer La'O, (2013). Sentiments, *Econometrica,* 81, 739-779.

Azariadis, Costas, (1981), "Self-Fulfilling Prophecies," *Journal of Economic Theory,* 25, 380-396.

Barsky, Robert B. and Eric R. Sims, (2012), "Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence," American Economic Review, 102(4), 1343-77.

Benhabib, Jess and Roberto Perli (1994), "Uniqueness and Indeterminacy: On the Dynamics of Endogenous Growth," Journal of Economic Theory, 63, 113-142.

Benhabib, Jess, and Roger Farmer, (1999), "Indeterminacy and Sunspots in Macroeconomics," in *Handbook of Macroeconomics,* Vol. 1, J. Taylor and M. Woodford, eds., Amsterdam: North Holland/Elsevier, 387-448.

Benhabib, Jess, Stephanie Schmitt-Grohe and Martin Uribe (2001), "The Perils of Taylor Rules," Journal of Economic Theory, Vol. 96, Jan 2001, 40-69.

Benhabib, Jess, Pengfei Wang, and Yi Wen, (2015), "Sentiments and Aggregate Demand Fluctuations," *Econometrica,* 83(2), 549-585.

Blanchard, Olivier, (2016), "The Price of Oil, China, and Stock Market Herding," mimeo, January 17, http://blogs.piie.com/realtime/?p=5341.

Cameron, Colin A., and Pravin K. Trivedi, (2005), *Microeconometrics*, Cambridge University Press, New York.

Carroll, Christopher D., Jeffrey Fuhrer, and David W. Wilcox, (1994), "Does Consumer Sentiment Forecast Household Spending? If So, Why?," *American Economic Review,* 84(5), 1397-1408.

Cass, David and Karl Shell, (1983), "Do Sunspots Matter?," *Journal of Political Economy,* 91(2), 193-227.

Farmer, Roger, (1999), The Macroeconomics of Self-Fulfilling Prophecies," MIT Press, Boston.

Gerber, Alan S., and Gregory A. Huber, (2009), "Partisanship and Economic Behavior: Do Partisan Differences in Economic Forecasts Predict Real Economic Behavior?," *American Political Science Review,* 103(3), 407-426.

Gillitzer, Christian, and Nalini Prasad, (2015), "The Effect of Consumer Sentiment on Consumption: Cross-Sectional Evidence from Elections," mimeo, September.

Hoechle, Daniel, (2007), "Robust Standard Errors for Panel Regression with Cross-Sectional Dependence," Stata Journal, 7(3), 281-312.

Kaplan, Greg, and Guido Menzio, (2013), “Shopping Externalities and Self-Fulfilling Unemployment Fluctuations,” NBER Working Paper no. 18777, Ferbruary.

Keynes, John Maynard, (1936), *The General Theory of Employment, Interest, and Money*, Palgrave Macmillan, United Kingdom.

Ludvigson, Sydney C., (2004), “Consumer Confidence and Consumer Spending,” *Journal of Economic Perspectives*, 18(2), 29-50.

Mian, Atif, Amir Sufi, and Nasim Khoshkhou, (2015), “Government Economic Policy, Sentiments, and Consumption,” NBER Working Paper 21316, July.

Oh, S. and Michael Waldman, (1990), “The Macroeconomic Effects of False Announcements,” *Quarterly Journal of Economics*, 105(4), 1017-1034.

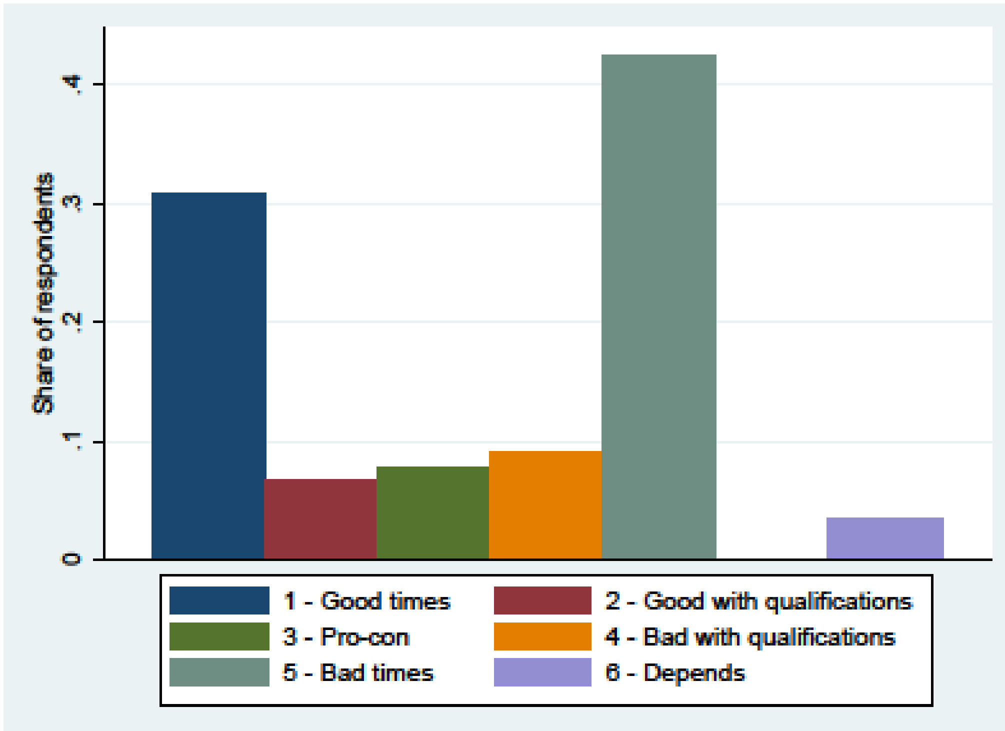
Shell, Karl , (1977), “Monnaie et allocation intertemporelle,” mimeo, Seminaire d'Econometrie Roy Malinvaud, Centre National de la Recherche Scientifique, Paris, November, 1977. [title and abstract in French, text in English]

Starr, Martha A., (2012), “Consumption, Sentiment, and Economic News,” *Economic Inquiry*, 50(4), 1097-1111.

University of Michigan, (2016), “Surveys of Consumers,” http://www.sca.isr.umich.edu/.

Figure 1

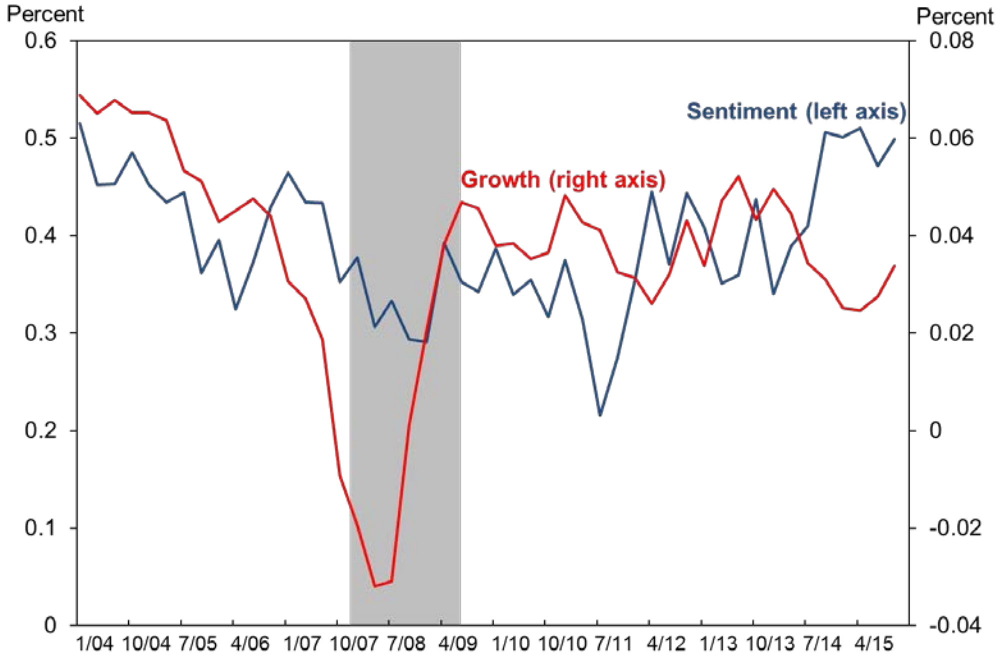
Distribution of responses to base sentiment question



Source: Michigan Survey of Consumers, 2004-2013. Histogram of percentages of each answer to survey question BUS5. See text for question.

Figure 2

Average sentiment and national output growth
(2004-2015)



Note: Source: Michigan Survey of Consumers, 2004-2015. Sentiment measured by average share of "GOOD" responses (1 or 2) to question BUS5. See text for question. Output growth from current quarter to four quarters in future.

Table 1

Summary Statistics

	GOOD	GGDP	INCOME	EDUC	INVEST	HGGDP	MGGDP	LGGDP
Mean	0.39	0.03	83512.27	14.79	0.68	0.04	0.03	0.03
SD	0.19	0.04	32752.83	2.08	0.18	0.05	0.03	0.03
Min	0	-0.19	4416	8	0	-0.12	-0.19	-0.07
Max	1	0.28	330080.78	44.83	1	0.28	0.23	0.12
N	2148	2148	2148	2148	2148	263	1579	306

Notes: See text for variable definitions. Test for HGGDP > MGGDP has p-value 0.016; test for HGGDP > LGGDP has p-value 0.17, while test for MGGDP > LGGDP has p-value 0.16. Data is taken for sample years 2004-2015.

Table 2

1st Stage IV Regression

	(1)	(2)	(3)	(4)	(5)	(6)
Congpres	0.09***	0.09***	0.09***	0.09***	0.09***	0.09***
	(5.05)	(5.15)	(4.26)	(5.29)	(5.31)	(4.59)
YGAP	0.02***	0.02***	0.02***			
	(9.87)	(3.53)	(7.18)			
Income	0.00**	0.00*	0.00**	0.00*	0.00**	0.00**
	(2.17)	(1.94)	(2.41)	(1.77)	(2.34)	(2.10)
Education	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
	(-0.47)	(-0.58)	(-0.42)	(-0.98)	(-1.25)	(-0.89)
Investment	0.07*	0.07*	0.07*	0.07**	0.07**	0.07**
	(1.80)	(1.94)	(1.97)	(2.01)	(2.35)	(2.31)
R^2	0.09	0.09	0.09	0.14	0.14	0.14
N	2343.00	2343.00	2343.00	2343.00	2343.00	2343.00
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes
Standard Error	Robust	DK	Cluster	Robust	DK	Clustered

Notes: * p < 0.1, **p < 0.05, ***p < 0.01. Note: Dependent variable is GOOD. OLS estimation with T-statistics in parentheses. Models 1 and 4 use robust standard errors, models 2 and 5 use the Driscoll-Kraay estimators, and models 4 and 6 allow for clustered standard errors by state. Data is taken for sample years 2004-2015.

Table 3
IV Regression

	(1)	(2)	(3)	(4)	(5)	(6)
GOOD	0.19***	0.19**	0.19**	0.13***	0.13**	0.13*
	(3.29)	(2.13)	(2.16)	(3.22)	(2.33)	(1.90)
YGAP	-0.00***	-0.00*	-0.00**			
	(-3.76)	(-1.87)	(-2.30)			
Income	-0.00***	-0.00**	-0.00***	-0.00	-0.00*	-0.00
	(-2.93)	(-2.11)	(-2.84)	(-1.26)	(-1.84)	(-1.46)
Education	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
	(-0.98)	(-0.77)	(-0.81)	(-0.69)	(-0.82)	(-0.64)
Investment	0.00	0.00	0.00	-0.00	-0.00	-0.00
	(0.29)	(0.42)	(0.38)	(-0.07)	(-0.09)	(-0.09)
R^2	.	-0.69	.	.	-0.02	.
N	2148.00	2148.00	2148.00	2148.00	2148.00	2148.00
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes
Standard Error	Robust	DK	Cluster	Robust	DK	Clustered

Notes: * p < 0.1, **p < 0.05, ***p < 0.01. IV estimation with CONGPRES as instrument for GOOD. T-statistics in parentheses. Models 1 and 4 use robust standard errors, models 2 and 5 use the Driscoll-Kraay estimators, and models 4 and 6 allow for clustered standard errors by state. Data is taken from sample years 2006-2016.

Table 4
Sample Sensitivity

	(1)	(2)	(3)	(4)	(5)	(6)
Base	0.19***	0.19**	0.19**	0.13***	0.13**	0.13*
	(3.29)	(2.13)	(2.16)	(3.22)	(2.33)	(1.90)
Drop financial crisis	0.23***	0.23**	0.23**	0.15***	0.15***	0.15***
	(3.38)	(2.50)	(2.28)	(3.57)	(2.94)	(2.62)
Drop high income	0.27***	0.27*	0.27**	0.16***	0.16***	0.16**
	(2.87)	(1.92)	(2.23)	(2.96)	(2.62)	(2.01)
Drop low income	0.20***	0.20***	0.20**	0.14***	0.14**	0.14*
	(3.15)	(2.60)	(1.97)	(2.96)	(2.48)	(1.67)
Drop high investment	0.21***	0.21*	0.21**	0.14***	0.14**	0.14*
	(3.08)	(1.87)	(2.05)	(3.02)	(2.03)	(1.81)
Drop low investment	0.19***	0.19**	0.19**	0.14***	0.14***	0.14*
	(3.15)	(2.57)	(2.03)	(3.09)	(2.87)	(1.74)
Drop large GDP	0.19***	0.19**	0.19**	0.13***	0.13**	0.13*
	(3.16)	(2.14)	(2.04)	(3.06)	(2.35)	(1.78)
Drop small GDP	0.08*	0.08	0.08	0.04	0.04	0.04
	(1.76)	(0.79)	(1.54)	(1.27)	(0.54)	(0.90)
Drop outliers	0.16**	0.16	0.16**	0.10*	0.10	0.10
	(2.33)	(1.24)	(2.01)	(1.68)	(0.95)	(1.29)

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. IV estimation with CONGPRES as instrument for GOOD. Coefficient estimates are for GOOD sentiment values. T-statistics in parentheses. Models 1 and 4 use robust standard errors, models 2 and 5 use the Driscoll-Kraay estimators, and models 4 and 6 allow for clustered standard errors by state. Data is taken from years 2006-2016.

Table 5
Estimator Sensitivity

	(1)	(2)	(3)	(4)	(5)	(6)
StateGDP weighted	0.16**	0.16	0.16*	0.12**	0.12	0.12
	(2.21)	(1.44)	(1.91)	(2.06)	(1.34)	(1.41)
Number of responses weighted	0.18**	0.18	0.18*	0.10**	0.10	0.10
	(2.42)	(1.29)	(1.90)	(2.31)	(1.39)	(1.51)
Random effects	0.19**		0.19**	0.13*		0.13*
	(2.12)		(2.12)	(1.85)		(1.85)
Quarterly dummies				0.12***	0.12*	0.12*
				(3.15)	(1.95)	(1.68)
Regional dummies	0.16***	0.19**	0.16*	0.13***	0.13**	0.13*
	(2.86)	(2.13)	(1.84)	(3.21)	(2.32)	(1.89)
Annual frequency	0.19***	0.19***	0.19***	0.13***	0.13***	0.13**
	(3.19)	(3.86)	(2.61)	(2.84)	(3.56)	(2.05)
1 Year State GDP Lagged	0.12**	0.12***	0.12**	0.1**	0.1***	0.1*
	(2.38)	(3.34)	(2.46)	(2.16)	(2.81)	(1.87)
1 Year US GDP Lagged	0.19***	0.19***	0.19**	0.13***	0.13***	0.13**
	(3.31)	(5.16)	(2.50)	(2.84)	(3.56)	(2.05)
1 Year State and US GDP Lagged	0.15***	0.15***	0.15**	0.1**	0.1***	0.1*
	(2.86)	(4.63)	(2.48)	(2.16)	(2.81)	(1.87)
Taxes	0.17***	0.17***	0.17***	0.13***	0.13***	0.13**
	(3.45)	(5.12)	(2.90)	(2.80)	(3.59)	(2.02)
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes
Standard Error	Robust	DK	Cluster	Robust	DK	Clustered

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: IV estimation with CONGPRES as instrument for GOOD. Coefficient estimates are for GOOD sentiment values. T-statistics in parentheses. Quarterly data except where indicated and for lagged specifications which have annual frequency. The number of observations for annual is 588; Taxes is 539; and the rest is 2148. Robust standard errors ((1) and (4)), Driscoll-Kraay standard errors ((2) and (5)) and standard errors clustered by state ((3) and (6)) in parentheses. Quarterly data estimated for 2006-2016; annual samples for 2005-2016.

Table 6

Sentiment

	(1)	(2)	(3)	(4)	(5)	(6)
Good1	0.26***	0.26**	0.26*	0.19***	0.19**	0.19*
	(2.75)	(2.19)	(1.93)	(2.74)	(2.29)	(1.71)
BAD5	-0.14***	-0.14*	-0.14**	-0.10***	-0.10*	-0.10*
	(-3.59)	(-1.67)	(-2.20)	(-3.41)	(-1.81)	(-1.84)
Better1	0.51	0.51	0.51	0.15***	0.15**	0.15
	(1.37)	(1.16)	(0.61)	(2.82)	(2.01)	(1.64)
GOVT	0.17***	0.17*	0.17*	0.14***	0.14*	0.14
	(3.58)	(1.80)	(1.95)	(3.12)	(1.91)	(1.57)
PCE	0.11***	0.11***	0.11***	0.07***	0.07***	0.07***
	(3.45)	(6.54)	(3.79)	(3.84)	(5.77)	(3.11)
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes
Standard Error	Robust	DK	Cluster	Robust	DK	Clustered

* $p < 0.1$, **$p < 0.05$, ***$p < 0.01$

Notes: IV estimation with CONGPRES as instrument for indicated sentiment variable. Coefficient estimates for alternative sentiment variables. See text for definitions. Last row reports IV results for consumption, measured by PCE, with GOOD variable as sentiment. GOOD1, BAD5, Better1, GOVT had 2148 observations; BUS5 had 2147 observations; PCE is annual data and had 539 Observations. Robust standard errors ((1) and (4)), Driscoll-Kraay standard errors ((2) and (5)) and standard errors clustered by state ((3) and (6)) in parentheses. Data is taken from years 2006-2016 for everything but PCE which is taken from 2005-2015.

Table 7

Long Horizon Results

	(1)	(2)	(3)	(4)	(5)	(6)
2 year IV	0.11***	0.11*	0.11*	0.08***	0.08	0.08
	(3.08)	(1.75)	(1.74)	(2.84)	(1.58)	(1.42)
3 year IV	0.05**	0.05	0.05	0.04*	0.04	0.04
	(2.03)	(1.13)	(0.79)	(1.93)	(1.09)	(0.71)
No. obs 2yr IV	1954	1954	1954	1954	1954	1954
2 year PCE	0.11***	0.11***	0.11***	0.06***	0.06***	0.06**
	(3.25)	(4.77)	(3.10)	(3.19)	(4.54)	(2.18)
3 year PCE	0.08***	0.08***	0.08**	0.05***	0.05***	0.05*
	(2.82)	(3.26)	(2.45)	(2.65)	(4.02)	(1.74)
No. obs 2yr PCE	490	490	490	490	490	490
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes
Standard Error	Robust	DK	Cluster	Robust	DK	Clustered

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: IV estimation for state GDP and state PCE growth over 2 and 3 years, with CONGPRES ans instrument for GOOD variable. Coefficients shown for GOOD variable. T-statistics in parentheses. Models 1 and 4 use robust standard errors, models 2 and 5 use the Driscoll-Kraay estimators, and models 4 and 6 allow for clustered standard errors by state. For 2 year PCE we are sampling from 2006-2015 data. For 1 year IV we are sampling from 2007-2016.

APPENDIX

A.1 Variable definitions and sources

Better1= Sentiment – percent of people who answered 1 to the "bago" survey question (think business conditions are "better now" than they were one year ago)

Good1= Sentiment **–** Percent of people who answered 1 to the "bus5" survey question (think the country will be doing "Good" in the next 5 years)

Good12= Sentiment- Percent of people who answered 1 or 2 to the "bus5" survey question (think the country will be doing "Good" or "good with qualifications" in the next 5 years)

Congpres= Percent of Congress representatives in each state that share the same party as the sitting president

Educ= highest level of education attained

GGDP = GDP growth by state over the past 4 quarters

GHSENT= gdpgrowth in states that have 4-quarter lagged sentiment greater than 1 SD above the mean

GNSENT= gdpgrowth in states that have 4-quarter lagged sentiment within 1 SD of the mean

GLSENT= gdpgrowth in states that have 4-quarter lagged sentiment greater than 1 SD below the mean

GOVT= Sentiment – percent of people who answered 1 to the "govt" survey question (think the government is doing a "good job" on economic policy)

INVEST= Percent of people who said they invest, by state and quarter

i.region = regional dummies (4 total regions)

i.time = quarter dummies

i.state= state dummies (49 total states – no HI or AK)

LGOOD12 = bus5_1and2 lagged 4 quarters

NRESP= Number of individual respondents to the Michigan Survey in that state in that quarter

STATEGDP= GDP by state

Table A1

Average Consumer Sentiment

State	GOOD	Congpres	GSP Per Capita	Education
Virginia	0.47	0.56	52.90	15.79
Washington	0.44	0.53	55.50	14.90
Nevada	0.44	0.49	47.65	13.95
Georgia	0.43	0.42	45.05	15.28
Delaware	0.43	0.71	66.64	14.71
North Dakota	0.43	0.31	59.20	14.64
Arizona	0.42	0.53	40.73	14.68
Nebraska	0.42	0.39	51.94	14.69
Utah	0.42	0.38	44.69	15.08
Wyoming	0.42	0.37	67.88	15.04
Texas	0.42	0.42	52.43	14.66
North Carolina	0.42	0.47	45.28	14.85
Rhode Island	0.41	0.69	47.80	15.02
Iowa	0.41	0.50	49.01	14.25
Maryland	0.41	0.63	54.81	15.25
Minnesota	0.41	0.60	53.13	14.89
Wisconsin	0.41	0.43	46.37	14.18
California	0.41	0.58	56.09	14.66
Colorado	0.41	0.57	52.60	14.85
Dist. of Columbia	0.41	0.63	169.16	15.64
New Hampshire	0.40	0.64	49.91	14.43
Illinois	0.40	0.52	53.50	15.08
Michigan	0.40	0.51	41.71	14.58
New Jersey	0.40	0.52	57.48	15.02
South Carolina	0.40	0.36	37.08	14.50
South Dakota	0.39	0.32	48.72	14.03
Massachusetts	0.39	0.62	63.15	15.16
Idaho	0.39	0.40	36.54	14.58
Tennessee	0.39	0.36	42.24	14.56
New York	0.39	0.58	63.65	15.30
Kentucky	0.38	0.41	39.09	15.36
Alabama	0.38	0.38	37.81	14.57
Missouri	0.38	0.44	43.58	14.97
Ohio	0.38	0.46	46.04	14.76
Florida	0.38	0.47	40.80	14.65
Oklahoma	0.38	0.36	43.12	14.69
Maine	0.38	0.50	39.22	14.53
Indiana	0.38	0.43	44.80	14.16
Pennsylvania	0.38	0.47	48.41	14.54
Montana	0.37	0.51	39.68	14.74
Louisiana	0.37	0.37	49.60	14.74
Connecticut	0.36	0.75	66.70	15.55
Oregon	0.35	0.65	49.10	15.18
Kansas	0.35	0.30	46.63	14.65
Arkansas	0.35	0.27	36.35	14.88
West Virginia	0.34	0.42	36.06	14.20
Vermont	0.32	0.63	43.41	14.49
Mississippi	0.31	0.39	32.77	14.83
New Mexico	0.29	0.76	42.42	14.75

Note: Data sampled from 2005-2015

Table A2

Reverse Specifications

	(1)	(2)	(3)	(4)	(5)	(6)
GGDP	0.33**	0.33*	0.33**	0.22	0.22	0.22
	(2.05)	(2.02)	(2.02)	(1.16)	(1.35)	(1.40)
YGAP	0.01***	0.01	0.01**			
	(2.80)	(1.00)	(2.19)			
INCOME	0.00	0.00	0.00	-0.00	-0.00	-0.00
	(0.59)	(0.57)	(0.57)	(-0.36)	(-0.45)	(-0.37)
EDUC	0.00*	0.00	0.00**	0.00	0.00	0.00
	(1.76)	(1.73)	(2.17)	(0.98)	(1.12)	(1.25)
INVEST	0.01	0.01	0.01	0.03	0.03	0.03
	(0.15)	(0.15)	(0.11)	(0.89)	(0.83)	(0.67)
Observations	1949.00	1949.00	1949.00	1949.00	1949.00	1949.00
R^2	0.05	0.05	0.05	0.12	0.12	0.12
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	No	Yes	Yes	Yes
Standard Error	Robust	DK	Cluster	Robust	DK	Clustered

* $p < 0.1$, **$p < 0.05$, ***$p < 0.01$

Notes: Dependent variable GOOD. Data sampled from 2006-2015. Robust standard errors ((1) and (4)), Driscoll-Kraay standard errors ((2) and (5)) and standard errors clustered by state ((3) and (6)) in parentheses.