NBER WORKING PAPER SERIES

ON MEASURING THE EFFECTS OF FISCAL POLICY IN RECESSIONS

Jonathan A. Parker

Working Paper 17240
http://www.nber.org/papers/w17240

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2011

This paper was prepared for the Journal of Economic Literature forum, The Multiplier. The author thanks Martin Eichenbaum, Lawrence Christiano, Pierre-Olivier Gourinchas, Ricardo Reis, and Michael Woodford for very helpful discussions and comments. This research was funded by the Kellogg School. In addition to being faculty at the Kellogg School, the author serves as an academic consultant and adviser for the Federal Reserve Bank of Chicago, and served as a Special Adviser on Financial Stability for the US Department of the Treasury during 2009. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

On Measuring the Effects of Fiscal Policy in Recessions
Jonathan A. Parker
NBER Working Paper No. 17240
July 2011
JEL No. E17,E5,E62

ABSTRACT

We do not have a good measure of the effects of fiscal policy in a recession because the methods that we use to estimate the effects of fiscal policy — both those using the observed outcomes following different policies in aggregate data and those studying counterfactuals in fitted model economies -- almost entirely ignore the state of the economy and estimate 'the' government multiplier, which is presumably a weighted average of the one we care about — the multiplier in a recession — and one we care less about — the multiplier in an expansion. Notable exceptions to this general claim suggest this difference is potentially large. Our lack of knowledge stems significantly from the focus on linear dynamics: VARs and linearized (or close-to-linear) DSGEs. Our lack of knowledge also reflects a lack of data: deep recessions are few and nonlinearities hard to measure. The lack of statistical power in the estimation of nonlinear models using aggregate data can be addressed by exploiting estimates of partial-equilibrium responses in dissaggregated data. Microeconomic estimates of the partial-equilibrium causal effects of a policy can discipline the causal channels inherent in any DSGE model of the general equilibrium effects of policy. Microeconomic studies can also provide measures of the dependence of the effects of a policy on the states of different agents which is a key component of the dependence of the general-equilibrium effects of fiscal policy on the state of the economy.

Jonathan A. Parker
Finance Department
Kellogg School of Management
Northwestern University
2001 Sheridan Road
Evanston, IL 60208-2001
and NBER
Jonathan-Parker@Kellogg.Northwestern.edu

When the Obama Administration came into office in January 2009, the US economy had been in a recession for more than a year and real GDP was falling at a 6 percent annual rate. The Federal Reserve's target for the interest rate on Federal funds was at the zero lower bound, so that conventional monetary policy had reached its limits. The Obama Administration pursued continuity of policy with the Bush Administration on stabilization of financial markets (continuing the Troubled Asset Relief Program), but considered cuts in taxes and increases in government spending and transfers to state governments to maintain aggregate demand and to increase GDP and employment. The Administration turned to economists – significantly academic economists – to help craft the size and details of the stimulus package. How much fiscal stimulus if any should be enacted? Which taxes should be cut or what goods should be purchased? How persistent should spending or tax cuts be?

But, almost as useless as no answer, academic economics provided a wide range of answers. As examples, Krugman (2009) called for much more stimulus spending than actually enacted, arguing that each additional dollar of government purchases would raise output by 1.5 dollars, while Barro (2009) argued for no additional spending, writing that from 1943 to 1944 "World War II raised U.S. defense expenditures by $540 billion (1996 dollars) per year . . . [and] the war raised real GDP by $430 billion per year . . .the multiplier was 0.8 . . ." The academic disagreement left policymakers using a multi-equation macroeconomic forecasting model, one inconsistent with the best practices of modern macroeconomics which identify the effects of fiscal policy either by imposing fewer and more explicit restrictions on time series dynamics or imposing many more restrictions but based on equilibrium conditions from optimizing behavior.

In this article, I address why we had so little evidence on the question of interest in early 2009 and how research can proceed to rectify this problem. I consider the central question to be: *how much, if any, fiscal stimulus is appropriate when facing a deep recession?* That is, I maintain as an important possibility that countercyclical fiscal policy can be effective in the spirit of Keynes (1936), meaning that increases in government spending or tax cuts may have significant positive effects on output and consumption if and only if there are significant slack resources in the economy. This view of the world corresponds to that of many practitioners and to that contained in many non-quantitative discussions and textbook descriptions of the

effects of fiscal policy: fiscal stimulus, if effective, is only effective when unemployment is high and capacity utilization low. In this (old) Keynesian view, in recessions, markets are somehow failing, and these failures can be (probably imperfectly) rectified through fiscal policy which will return idle resources to work. In an expansion however, market are somehow working and there are no benefits to expansionary fiscal policy as it would primarily raise interest rates and crowd out private consumption and investment, as in the neoclassical model of economic fluctuations.

Returning to Barro (2009)'s examples of 1943-44, the unemployment rate during 1942 averaged 4.7%, and was steadily falling reaching 0.7% by 1944. In contrast, in January of 2009 the unemployment rate was 7.8 percent and has steadily risen since, and still stands at 9.1% more than two years later. If idle resources matter for the multiplier, then the larger share of idle workers in 2009 may imply a larger multiplier in 2009.[1] Unfortunately, we have very little evidence on whether the government multiplier differs with the state of the economy. That is, we do not know the efficacy of fiscal policy *in recessions*, which is dispiriting given the central importance of the question.

This article begins by discussing the reasons for our lack of knowledge and then lays out a methodology for improving our estimate of the efficacy of fiscal policy in recessions. This article is not a review of the literature on the multiplier; several papers contain good discussions of the extant methods and issues (see Hall (2009), Auerbach and Gale (2010), Woodford (2011), and Ramey (forthcoming)). Instead, this article is an idiosyncratic, selective discussion of why we could not answer the question I am posing and what to do about it.

One major reason for our inability to answer the central question is methodological – our main methodologies answer a different question. A Keynesian-style multiplier – one that is large in recessions and small at other times – is ruled out in the dynamics of any DSGE model linearized around a single point and in any standard VAR because it does not allow state dependence. That is, in almost every DSGE model and VAR analysis, fiscal policy is as effective in a roaring boom as a deep recession.

This point highlights a deeper issue: in many cases where the nonlinearities of DSGE

[1] Romer (1992) and Gordon and Krenn (2010) also discuss the relevance of the unemployment rate for the effects of fiscal policy in the Great Depression.

models – including new Keynesian models – have been studied, the dynamics are in fact close to linear in structure. Thus, either most models mismeasure the multiplier in that their components imply a lack of state-dependence in the government multiplier or the models are correct and 'idle' resources do not matter for the effects of fiscal policy. Which conclusion you reach depends on your commitment to extant models.

Turning back to measurement, a second distinct point is that even research that seeks to estimate the multiplier in recessions actually estimates the marginal multiplier not the total multiplier. If the marginal multiplier declines with the size of the stimulus, then government actions - countercyclical fiscal policy – will tend to make the marginal multiplier significantly lower than the average multiplier in recessions.

To be clear about the issue with DSGE models, we can, do, and have learned about the effects of cyclical fiscal policy from extant models because the welfare effects of fiscal policy are state dependent. And there are notable extant models that do have an important role for state dependence, some of which are discussed subsequently.

To be clear about the issue with VARs, we do learn a lot about 'the' marginal multiplier from VARs. And there are notable extant empirical studies that are nonlinear and investigate the role of idle resources for the marginal effects of government spending, some of which are discussed subsequently. But while research sometimes finds evidence of larger effects of government spending in recessions, the evidence is statistically weak, highlighting the real reason for our lack of knowledge: lack of data.

The final section of this articles considers how economic research can measure the multiplier as a function of the state of the economy. From a theoretical perspective, it is feasible to extend existing DSGE models to include frictions that induce a high degree of curvature in the cost of supplying output and to extend analysis of these models to study nonlinearities that cause the elasticity of output with respect to demand to be high in recessions and low in expansions. Researchers are beginning to investigate which frictions can generate interesting nonlinearities.

But in terms of quantification, the difficulties seem more significant. First, methodologically, large changes are necessary to model state dependence in VAR-type empirical models; much of VAR inference is conditioned on its linear structure. Second, there are simply few

recessions, and fewer deep recessions, to learn from. The problems of weak inference that already plague the linear VAR literature on the multiplier are more severe for the more data-intensive estimation of state-dependent dynamics. Thus, while we can build DSGE models with state dependence that is potentially quantitatively important, inference based on aggregate data alone is likely to yield imprecise measurement of the multiplier.

But there is another source of evidence: causal evidence from microeconomic studies of the effect of policies on economic outcomes taking prices as given. Microeconomic studies can estimate the direct, partial-equilibrium response of agents to a policy – the effect of policy on microeconomic spending or investment excluding the multiplier for example – which is an important component of the general-equilibrium response of the economy to the policy. Further, such studies can measure the dependence of these responses on the states of different agents which is an important component of the dependence of the general-equilibrium response on the state of the economy. These studies represent a largely unexploited source of information for structural macroeconomic models designed to give quantitative evaluation of the multiplier. In any quantitative DSGE model, the more links of the implicit causal chain running from policy to general equilibrium multiplier that are disciplined by explicit microeconomic estimates of causation, the more confidence we can have in our model-based interpretations of impulse response functions and model-based counterfactual policy analyses. This method stands in contrast to most of the current literature which derives quantitative discipline from best-fitting (linear) aggregate dynamics and/or from external estimates of structural parameters.

While most of this paper focusses on the quantification of the efficacy of fiscal policy, many of the main points are as relevant to the quantification of the efficacy of monetary policy.[2]

1 What do we know about the multiplier?

To infer the effect of countercyclical fiscal policy on the economy, one must estimate a counterfactual history in which the policy is different from the observed policy. There are two

[2] Any study of the effects of government spending is already in part the study of monetary policy since fiscal and monetary policy are necessarily linked.

main ways this is done: DSGE models that analyze counterfactual policies in fitted model economies and structural VAR's that infer the effects of policy from differences in the evolution of the economy following different policies controlling for the state of the aggregate economy.

The former, DSGE models, are designed to approximate the dynamics of the U.S. economy and provide structural interpretations of the correlations observed in macroeconomic time series. The calibrated or fitted models are used to conduct counterfactual policy experiments and generate estimates of the effects of alternative levels of government transfers or purchases. The size of the multiplier depends on the model chosen. In the neoclassical model, increases in government spending raise output through reducing household wealth and increasing their labor supply. Plausible estimates of the parameters governing this response imply small effects of government spending on output, typically less than 0.5 and sometimes even negative. More recently, fiscal policy has been studied in New Keynesian models where the effects can be significantly larger, but still typically less than unity.[3]

The latter, VAR models, use a small number of orthogonality assumptions to first identify changes in policy that are unexpected given the state of the economy and then trace out the effects of these policy innovations on the evolution of the economy and policy. There are a large number of interesting and challenging issues in implementation and interpretation of these findings, many discussed by Ramey (forthcoming). In terms of magnitude, Hall (2009) concludes that VARs ". . . find fairly consistently that the output multiplier is in the range from 0.5 to 1.0 and that the consumption multiplier is somewhat positive."

But almost all DSGE and VAR analyses do not answer the questions asked in the introduction. I now turn to several reasons why this is so. The list I cover is incomplete; most notably, I do not discuss the limited evidence on what goods to purchase and which taxes to cut and how these decisions might lead to different multipliers from that for government purchases of good that do not enter the utility function.[4]

[3]Examples include Barro and King (1984), Aiyagari, Christiano, and Eichenbaum (1992), Ramey and Shapiro (1998), Burnside, Eichenbaum, and Fisher (2004), Galí, López-Salido, and Vallés (2007), and Woodford (2011).

[4]Taylor (forthcoming) discusses the issue of the effects of different types of goods or transfers.

2 The problem of 'the' multiplier

Almost all the research using either VARs to measure the response of macroeconomic variables to different policies or DSGE models fitted to macroeconomic data employs linear models or solution methods which has two problems. First, there is a time-invariant government multiplier that is independent of the state of the economy: fiscal policy is *assumed* as effective in a recession, when there may be idle resources, as in a boom, when there are not. Second, a linear model forces a multiplier independent of the size of the stimulus. Econometric studies estimate the effects of a marginal change. Reasonable arguments suggest that, if the marginal and total multiplier differ, the marginal is smaller.

2.1 DSGE-based inference

Consider first estimates of the government multiplier derived from model economies.[5] In principle, research using this methodology could be useful for measuring a Keynesian multiplier as I define it. In practice, little of it is. Why? Because most research solves the dynamics of the model by employing a single first-order linearization of the non-linear system of equations. This leads to two related problems for understanding the multiplier, although, for most models in both the Neoclassical and New Keynesian tradition, the issue is deeper than the solution method.

A linearized solution to a DSGE model implies that the partial derivative of output or consumption (or their present discounted values) with respect to government spending or taxes is the same following a large positive shock to the economy as following a large negative shock to the economy. The effects of government spending on output or consumption are the same in a roaring boom, when the unemployment rate is low and capacity utilization high, as in the depths of a recession, when the unemployment rate is high and capacity utilization low.

This is not to say that optimal fiscal policy cannot stabilize output in a linearized DSGE model. A choice of fiscal policy is the specification of a fiscal policy rule that depends on the state variables in the model economy. The amount of government spending can thus be

[5] I am considering the standard definition of the government multiplier as the effect of a change in government spending on output or consumption, which are directly measurable, rather than model-based welfare.

conditioned on variables such as the level of the capital stock or the output gap, and so can go down when good shocks hit the economy and up when bad shocks hit the economy. And the choice of this rule is a critical determinant of welfare and business cycle dynamics, and countercyclical fiscal policy may or may not be optimal in any given model.

The problem arises with quantitative measurement. Government spending or tax policy does not have the same effect on discounted utility in a recession as in a boom, but it does have the same effect on output or consumption.[6] The fiscal policy rule in a linearized DSGE is linear and has linear effects on measurable economic quantities. In the linearized model, the study of optimal fiscal policy is based on the answer to the question 'can the government raise model-based utility by conditioning government spending linearly on the state of the economy given that its effects are always the same?' and not 'can the government raise output or consumption more by increasing government spending in a recession than a boom and so should it?'[7]

This is not just (or really) a problem in the application of a methodology, it is also a problem or at least an issue for model specification. There are methods for studying economic dynamics allowing for state dependence, such as linearizing around multiple points, employing higher-order expansions of the system, and numerical characterization of the full dynamics of the nonlinear system.[8] These methods have been used to evaluate the accuracy of the approximation provided by the linearized system of equations. And most models – certainly the textbook neoclassical model and textbook New Keynesian model – are close to linear in structure so that the linear approximation provides an extremely good characterization of the actual dynamics of the non-linear system. What does one make of this? On the one hand, if one has a high degree of confidence in these models, then one believes that the average multiplier is always the multiplier and that the quantitative recommendations about fiscal policy from these models are correct. In particular this implies that the optimal response to a

[6] So the optimal policy in response to a shock to the economy in a boom is the same as that to a shock to the economy in a recession.

[7] Another way to understand this point is to note that in most modern business cycle theory, the choice of a policy rule is made to minimize the variance of output of consumption around a given mean. In the Keynesian proposition considered in the paper, fiscal policy is not just about the variance of output or consumption, but also its level and about its skewness. It admits the possibility that a rule to pursue expansionary policy only in deep recessions for example could raise average output.

[8] See for example Eggertsson and Woodford (2003), Kim, Kim, Schaumburg, and Sims (2005), and Fernandez-Villaverde and Rubio-Ramirez (2007).

small boom should be the optimal response to a large recession, scaled-down and in reverse.[9] But if one entertains the possibility that the multiplier might be countercyclical, then one would like models to test that admit this possibility.

And in fact there are notable examples of nonlinear models used to estimate quantitative effects of fiscal policy. The most important of these is the literature on fiscal policy at the zero lower bound (Krugman (1998), Eggertsson and Woodford (2003), Eggertsson (2008), Christiano, Eichenbaum, and Relebo (2011), and Woodford (2011)). Christiano, Eichenbaum, and Relebo (2011) in particular show that the government spending multiplier can be very large – 3 or more – when the interest rate is held constant, such as when nominal interest rates are zero in a model with sticky prices. The same model delivers multipliers significantly less than one when the interest rate is allowed to vary. But this nonlinearity is subtle: it depends on the fixity of the interest rate and not on the degree of slack in the economy. A commitment to maintain a constant real interest rate would lead to the same multiplier in a boom as in a bust, although the monetary authority may find it more costly to pursue such a policy in a boom and the type of economic shocks that would tend to push the economy into the zero lower bound are also the type of shocks that tend to cause output to fall and would cause utilization of capital and labor to decline.

2.2 VAR-based inference

Vector autoregressions provide evidence on the government multiplier by characterizing the evolution of the economy following changes in fiscal policy that are not responses to the current state of the economy, such as statistical innovations in fiscal variables (as in Blanchard and Perotti (2002)) assumed not caused by the economy or changes in military spending (as in Ramey and Shapiro (1998)) assumed driven by external factors that do not themselves impact the economy. These estimates provide estimates of the effect of government spending – as measured by the innovation in and impulse response of government spending – on the economy – as measured by the impulse response of output, prices, etc.

But VARs are linear systems. Writing the evolution of the economy as an (infinite-order)

[9] In any close-to-linear model in which the American Reinvestment and Recovery Act is an optimal response to a financial collapse, an American Disinvestment and Slowdown Act would be an optimal response to a financial boom.

moving average of structural shocks, the derivative of any variable with respect to any lagged innovation to fiscal policy is independent of the other shocks in that period or any shocks in any nearby periods. Thus, *by assumption*, the impulse response to an unexpected or exogenous increase in government spending is constrained to be the same independent of the state of the business cycle. Thus, the majority of the VAR literature on the efficacy of government spending in recessions suffers a similar problem to that of its DSGE sibling: the multiplier is constrained to be independent of the economic situation.

Also, the policy responses to other innovations in the economy are constrained to be linear in the size of the innovation. Again, writing the VAR in structural moving average form, the fiscal response to any shock to the economy is a linear impulse response, which therefore responds symmetrically to positive and negative shocks and increases in size linearly with the size of any given shock to the economy.

Finally, and distinct from estimating the average multiplier over time, VAR impulse responses also typically estimate the marginal multiplier from an incremental change in government spending, not the total multiplier from all countercyclical government spending. That is, the impulse responses to innovations in policy summarize the effect of a small unpredicted change in government spending – the effect of a deviation from the average countercyclical policy – not the change that occurs in responses to usual countercyclical fiscal policy – the total multiplier for countercyclical policy.

Might this be quantitatively important? If the marginal multiplier declines with the size of the stimulus, then government actions will tend to make the marginal multipliers significantly lower than average multipliers in recessions.

Specifically, suppose that the government were optimally responding to all economic disturbances, increasing spending or cutting taxes to the point where the present discounted benefits of a small increase in spending equalled its present discounted costs. Then measurement of the government multiplier derived from shocks to the VAR would be identified from errors in implementing the optimal policy given the state of the economy or changes in policy due to improved understanding of how best to conduct it, or even possibly to policy experimentation.[10] If we were to ask in this world whether the present discounted utility

[10] This assumes that these things exist and that we are not identifying errors using a misspecified model of

of consumption was increased by any small change in fiscal policy from the typical optimal path, the answer would be no. That is, small deviations from the optimal path of government spending would lead to no change in expected present discounted utility.[11]

Thus, if the effects of government spending are not linear in the size of the fiscal action, the marginal multiplier for consumption or output in a recession could be significantly smaller than the total multiplier for all countercyclical policy in a recession. If governments are more shortsighted than agents, more risk averse, or pursue robust policies, in a recession, the measured, marginal multiplier could be below one while the total multiplier is well above one.

Instead of the issues just discussed, the existing literature has focussed on a different difficulty in measuring the efficacy of macroeconomic policy in recessions: measurement of the effects of policies that are responses to the economic situation. These are policies that we could consider pursuing again. But the performance of the economy following a policy that is a response to the economic situation is a mixture of the natural response of the economy to the economic situation and the response of the economy to the policy. There are two ways to make the step from estimated structural VAR to quantitative analysis of countercyclical macroeconomic policy.

One approach is to treat the estimated impulse responses of the economy as a measure of the differential effects of different policies considered by a policymaker.[12] Considering different policy options is considering deviating from an average policy rule. If the policy shocks identified by the VAR measure previous deviations from the average policy rule, then the impulse response of these previous deviations provide a good guide to the effects of the current deviation. But if the actual effects of policy differ in booms and recession, the impulse responses estimated in most existing VAR analyses will significantly misestimate the multiplier in a recession because they measures the average of the multiplier in a boom and the multiplier in a recession.[13] Further, because these VAR estimates represent marginal multipliers, one learns only about the effects of small deviations from rules. That is, supposing

agents' expectations, such as if we omitted lags, variables, or nonlinearities from our VAR.

[11] A standard result of first-order conditions from government optimization.

[12] For example, this is the view with respect to monetary policy of Sims (1980) and Leeper, Sims, and Zha (1996), less explicitly in Romer and Romer (2010), and in much of the debate over the size of the multiplier in early 2009.

[13] Again, this is not the only issue here – there are many other assumptions in this approach that are debated in the VAR literature.

that estimates could be conditioned on being in a recession, even small estimated marginal multipliers would at most imply that government spending responses to recessions be scaled back not that they be abandoned.

The other use of VARs in the study of stabilization policy is as part of identification and estimation of a DSGE model. The parameters of a DSGE model are chosen to make the causal effects of shocks in the DSGE as consistent as possible with those of the estimated impulse responses to one or more structural shocks in the structural VAR.[14] This approach imposes more theoretical assumptions on the impulse responses, but gives them structural interpretations that permit counterfactual analysis and analysis of optimal policy. This methodology is tractable and elegant, but as currently implemented relies heavily on the linearity of both VAR and DSGE dynamics – the irrelevance of the unidentified shocks and of the state of the economy more generally is central to the identification of the effects of policy. To the extent that the real world has different economic responses to policy in different economic situations, this aspect of the data is never bought to the DSGE model and is ignored in estimation and inference.

There are of course exceptions to the general rule that VAR analysis has not been concerned with the nonlinear effects of policy and, as was the case for DSGE models, the extant evidence suggests that nonlinearities may be quite important. While some earlier work estimates VARs over sub-periods to investigate differences in impulse responses in different periods, the most significant work to date is Auerbach and Gorodnichenko (2010), which employs a smooth transition structural VAR in which the dynamics of the economy are characterized by a time-varying linear combination of two different linear autoregressive structures.[15] In any period, three variables – output, government purchases, and taxes – are determined by a weighted average of two different linear autoregressive models with weights depending on a seven-quarter moving average of output growth. This gives the model a non-linear structure, allowing estimation of multipliers that differ in recessions and in booms but also losing many of the standard benefits of VARs.

Auerbach and Gorodnichenko (2010) finds that multipliers are similar on impact in reces-

[14] This methodology and issues originate as discussions in the literature on monetary policy, see Rotemberg and Woodford (1997) and Christiano, Eichenbaum, and Evans (1998).

[15] Again, this article is not a survey, but see also Perotti (1999), Tagkalakis (2008), and Taylor (forthcoming).

sions and expansions, but that the impulse responses are quite different, implying cumulative multipliers over 5 years of 0 to 0.5 in a boom and 1 to 1.5 in a recession. While this evidence is consistent with the importance of state-dependence in the efficacy of fiscal policy, the paper raises many issues about how best to statistically model nonlinear dynamics. In particular, the Auerbach and Gorodnichenko (2010) approach smooths the impulse responses so that the data in normal times helps to inform the dynamics in both boom times and recessions.

This final point directs us to the central difficulty confronting VAR-based inference: lack of data. We have few severe recessions from which to learn (although we could also learn from avoided severe recessions), and each (potential) recession has a different set of causal factors, a different set of contemporaneous confounding factors, and different policies themselves – different spending on different goods and different types of tax changes, both coupled with different monetary policy responses. Thus, there is little information in macroeconomic data alone to inform policy.[16]

One response to the lack of data is to study more carefully the data that we do have. A vector autoregression constructs policy innovations by estimating expectations from small sets of aggregate variables and bases counterfactuals (the world without a shock) on outcomes following different levels of this small set of aggregate variables. Historical analysis can examine more information than that contained in limited aggregate data and decide with more precision – although more judgement – what a given counterfactual might look like. As a concrete example, the statistics from the financial sector that portended poor performance in this sector and possibly the economy, such as issuance of asset backed securities and various credit spreads, are typically excluded from VARs yet were taken into consideration by policymakers and private agents in 2007 and 2008. Notable examples of this approach include Ramey and Shapiro (1998) and Romer and Romer (2010), as well as a vast amount of research on the Great Depression. While historical analysis can potentially better determine the baseline expected dynamics of output and government spending and how they both change over time, the use of historical evidence is to some extent inherently subjective.[17]

[16] This is why many studies consider additional data – most commonly data from policy during expansions – to improve power, but, as I am arguing, this may come at the cost of bias.

[17] Of course in practice so too is the choice of tastes and technology in a DSGE model and the choice of variables in a VAR.

Another response to the lack of data is to use asset market data to better measure market expectations, or surveys of expectations to better measure counterfactuals. As examples, Ramey (2011) and Auerbach and Gorodnichenko (2010) both use surveys of professional forecasters and Fisher and Peters (2009) uses information from asset markets on the returns on the stocks of defense contractors to better measure the arrival of information about changes in government spending.[18] Finally, data from more countries can give more information, as in Perotti (1999) and Ilzetzki, Mendoza, and Vegh (2010). But few countries have aggregate data of the quality and length of that in the U.S., and different countries likely have different multipliers because they differ in many ways such as monetary policy regimes, tax systems, and labor market flexibility to name just a few.[19]

While all of these approaches are useful in that they bring more information to bear, the next section describes a different approach to increasing the information used in inference: the use of estimates from microeconomic studies to discipline steps in the causal chains from policy to outcomes that are embedded in any macroeconomic model.

3 Measuring the multiplier in a recession

To understand whether the multiplier for fiscal policy differs in booms and recession, we need to identify and measure the extent to which the effects of policy are state dependent. From a theoretical perspective, it seems feasible to extend existing DSGE models to include frictions that induce strong curvature or a kink in the cost of supplying output and that cause the elasticity of output with respect to demand to be high in recessions and low in expansions. While solution methods for such models exist, for practical use they may require refinement.[20] And while it is not obvious which frictions are the right ones, these issues seem ripe for research. As illustrated by the literature on the zero lower bound, such curvature can be induced by constraints which occasionally bind. Are there relevant constraints that generate curvature and slack resources when binding? Informational frictions seem like plausible candidates;

[18] See also Cochrane and Piazzesi (2002) which uses asset price data to study monetary policy.

[19] Favero, Giavazzi, and Perego (2011) studies multipliers across countries and relates the different estimates to different country characteristics, some of which are time-varying states.

[20] The solution methods already exist, "but oh man is that hard in practice!" (Christiano (personal communication)).

both the Great Depression and the 2008-2009 recession suggest that informational frictions in financial markets might be a promising starting point for investigation.[21]

But while theoretical models seem poised to investigate nonlinearities, the difficulties in extending inference to nonlinear environments are more serious. Asymptotically, given a correct model, this statement is of course incorrect. Nonlinear (statistical or DSGE) models can be estimated structurally from informative moments from the data. But data are limited. Research using VARs has made enormous strides in extracting the economic evolution caused by an innovation to a well-identified structural shock, such as monetary or fiscal policy, but many of these steps rely heavily on the linear structure of the VAR.[22] For example, linearity is required by current methods to parse the impulse response to an identified shock when one is unable to identify the remaining shocks in the system and to study this response or use it to estimate a DSGE model. Nonlinear methods are far more demanding of the data, or, worse, far more reliant on the specific functional forms of the model being estimated. In a model in which the effects of fiscal policy are allowed to differ across states of the economy, more parameters are estimated from the data than if the effects are constrained to be the same across states. Thus, while we can build DSGE models in which state dependence is potentially quantitatively important, inference based on aggregate dynamics, while still informative, is alone unlikely to be sufficient to give precise measurement of the multiplier.

In considering the use of additional information for measurement of the multiplier, two properties are desirable: independence from macroeconomic data and relevance to the multiplier.[23] Independence is desirable as it provides some guard against overfitting; the model can be estimated on one set of data and tested on the macroeconomic dynamics. Current practice in macroeconomics, following Prescott (1986) among others, often exploits independent sources of information by setting model parameters on the basis of trend relationships in aggregate data or cross-sectional estimates. But since every model is misspecified, the model

[21] While the finance literature has lots of nonlinearity, macroeconomic models with financial frictions do not. In the canonical macroeconomic model with financial frictions, Bernanke and Gertler (1989), the constraints always bind so that the model is close to linear – the constraint moves amplifying fluctuations but dynamics are linear. Older work contains nonlinearities: Mankiw (1986) for example contains an important nonlinearity but is not readily estimable. And much current work is incorporating occassionally-binding constraints into macroeconomic models and studying non-linear effects. See Brunnermeier and Sannikov (2011) for an example with financial constraints.

[22] Not that all issues have been resolved, just that many issues have been nicely resolved.

[23] I use the term independence in the sense used by Prescott (1986).

parameters that maximize fit for one aspect of the real world may not maximize fit or applicability of the model for another. Thus, relevance is also desirable. For example, cross-sectional estimates of parameters based on variation in the data that is closely related to the variation in fiscal policy in question are likely to be useful for fitting and understanding the macroeconomic dynamics following this fiscal policy. But even relevant variation has the problem that, when models are misspecified, best-fitting "structural" parameters are generally not invariant across environments.

An alternative approach when available is to base the parameters of the DSGE model on evidence from cleanly-identified estimates of the partial-equilibrium causal effects of policies. Evidence (or moments) from well-identified microeconomic studies of causation can satisfy both independence and relevance, and, because they can avoid imposing model-based restrictions, they are invariant across environments.[24] When aggregate state dependence is imprecisely measured in aggregate data alone, model-based inference about the multiplier can be made more precise by estimating the parameters of the model so that experiments in the model that replicate those in microeconomic studies of partial-equilibrium causal relationships match the causal patterns found in the microeconomic studies. Most helpful are studies of causal relationships that are central to the impact of the policy in general equilibrium. Further, to the extent that aggregate nonlinearity is due to the fact that individual behavior is dependent on the aggregate state (rather than dependent on individual-level states, the distribution of which are part of the aggregate state), then even more informative are cross-sectional analyses at times of deep recession that can be contrasted to analyses at more normal times.[25]

Relatedly, to the extent that the state-dependence in the macroeconomic model follows from a state-dependence at the microeconomic level, microeconomic studies can measure the importance of this nonlinearity at the microeconomic level. That is, microeconomic

[24] That is, estimates of causation can be (largely) model independent so that one can (largely) avoid model misspecification in the estimation in the cross-sectional data. Instead, more relevant estimates are desirable because they provide statistical power more useful for identifying the dynamics of interest.

[25] Like linear times series analysis, causal effects that do not depend on aggregate state are still informative and add information. Chetty, Guren, Manoli, and Weber (2011) for example survey microeconomic estimates of labor supply elasticities which is useful information for DSGE's. But the paper ignores possible state dependence and averages across all studies. It would be useful to know whether the elasticity varies with the state of the business cycle.

studies can measure the dependence of agents' responses on their states which is an important component of the dependence of the general-equilibrium response on the state of the economy. For example, if the nonlinearity in the macroeconomic model follows from variations in the extent to which a given constraint binds for agents, then the macroeconomic model can be made more quantitatively relevant if the behavior of constrained and unconstrained agents in the DSGE model matches the microeconomic evidence on the behavior of constrained and unconstrained agents.

Consider the example of the countercyclical tax policy. Johnson, Parker, and Souleles (2006) and Parker, Souleles, Johnson, and McClelland (2011) estimate the effect of the receipt of a Federal economic stimulus tax rebate on household spending. The estimation identifies the causal effect from the effective randomization in the timing of the disbursement of the rebates distributed in the recessions of 2001 and 2008. The papers find that the distribution of the rebates caused an economically-significant partial-equilibrium increase in the aggregate demand for consumption of nondurable goods (and some services) of a quarter to a third of the amount of the rebates, an amount which is much larger than that consistent with the models of household behavior under rational expectations that are embedded in both the neoclassical and New Keynesian models.[26] It seems unlikely that standard DSGE models that are inconsistent with this measured household response provide accurate quantitative evaluations of the economic dynamics following this policy or accurate measurements of the multiplier in particular.

Further, many papers including Johnson, Parker, and Souleles (2006) find evidence that the spending responses of households are stronger among low-income or low-asset households. This finding reveals state dependence in behavior at the microeconomic level, suggesting state dependence at the macroeconomic level. Specifically, variations in the share of households with low levels of liquid wealth or income may change the aggregate effects of fiscal policy, which suggests that liquidity constraints or incomplete markets across households are a potentially-important source of nonlinearity for a DSGE.[27] This exemplifies the standard way

[26] Hinting at some form of state dependence, Johnson, Parker, and Souleles (2006) finds statistically weak evidence of no response in spending on durable goods while Parker, Souleles, Johnson, and McClelland (2011) finds statistically stronger evidence of economically large increases in spending on durable goods.

[27] Foreshadowing issues discussed subsequently, Krusell and Smith (1998) show that, in a DSGE model with incomplete markets and borrowing constraints, the share of households that are constrained does affect

in which microeconomic findings often inform macroeconomic modelling.

Several (preliminary at writing) papers more directly build models for the study of fiscal policy that are designed and parameterized to be consistent with the microeconomic evidence on the effects of tax rebates on consumption demand.[28] Kaplan and Violante (2011) is a partial equilibrium model that focuses on how the state of the economy influences the quantitative response of aggregate demand to the policy. In the model, heterogeneous life-cycle households have access to a liquid asset and an illiquid asset with a higher rate of return that can be accessed only by paying a transaction cost. Consumers are impatient between adjustments and so consume at high rates from their liquid assets, and thus from liquid asset infusions like tax rebates. Kaplan and Violante (2011) finds significant state dependence: the model response of consumption demand to tax rebate policies varies by a factor of two across different economic scenarios considered.

In contrast, Huntley and Michelangeli (2010) and Oh and Reis (2011) study the linear dynamics of DSGE models – so that the effects of the fiscal program are constrained to be the same in recessions and in booms – but discipline the models to match the microeconomic evidence on the additional spending caused by tax rebates – so that one of the main channels through which the rebates effect the economy is disciplined by the causal microeconomic evidence. Huntley and Michelangeli (2010) studies a heterogeneous-agent economy with incomplete markets and borrowing constraints (as in Aiyagari (1994)). Oh and Reis (2011) also include nominal rigidities. The papers illustrate nicely how moments that represent model-free estimates of causation can be used to fit DSGE models.

Huntley and Michelangeli (2010) estimates that the multiplier for a tax rebate like the ones implemented is very small (less than 0.2) and the effect of the rebate transitory.[29] On the one hand, one might have expected the fitted model to produce a larger-than-average multiplier because the model is fit to data from a recession, the state of the world in which the policy might have larger impact. On the other hand Huntley and Michelangeli (2010) chooses model features that match the microeconomic evidence without other inefficiencies

macroeconomic dynamics but the incomplete markets and changing number of constrained agents does not make the dynamics of the model quantitatively non-linear.

[28] See also Reis (2006).

[29] Huntley and Michelangeli (2010) is a moving target at this point, but Taylor (forthcoming) reaches similar conclusions.

that could generate a role for policy to improve the allocation of inefficiently used (or unused) productive resources. So the nonlinear model is probably close to linear.[30]

These results suggest that if the multiplier is to vary with the state of the business cycle in a DSGE model, nonlinearities need to come from other or additional sources besides household liquidity constraints, such as either an aggregate constraint that, like the zero lower bound, binds occasionally or agent-specific constraints that generate an important role for idle resources (failures to make productive transactions or matches). Finally, this example illustrates the limits of additional evidence from microeconomic studies: it only provide information about microeconomic effects and leaves the rest of the economy for specification by the researcher.

Another example of the use of estimates of relevant causal effects from disaggregated data is Nakamura and Steinsson (2011). This paper estimates the effect of an increase in Federal military spending in a given U.S. state on output in that U.S. state relative to spending and outcomes in other states. Like most studies that use cross-sectional variation, this measured effect is not a measure of the national multiplier. A far larger share of spending in a state goes to goods produced outside of that state than for the country as a whole; Federal spending affects the Federal budget constraint and so the tax ramifications of higher spending in one state are shared across states; finally, common economic reactions to the aggregate policy such as monetary policy responses are not captured by the cross-state difference in outcomes. However, these features can be viewed as advantages. By exploiting the greater amount of variation in the cross-state data, Nakamura and Steinsson (2011) estimates reasonably precisely a statistic that is both independent of aggregate dynamics and relevant for the aggregate government multiplier. In fact, it also turns out that the statistic provides a sharp test among a set of quantitative DSGE models, and one that is largely independent of the monetary policy rule in each model. A New Keynesian DSGE model with nonseparable preferences between consumption and leisure matches the cross-sectional impact of government spending well while the baseline New Keynesian model and various variants of the neoclassical model do not.

While Nakamura and Steinsson (2011) does not analyze possibly nonlinearities or state

[30] This is also the case for Oh and Reis (2011). The near-linearity again confirms a main finding of Krusell and Smith (1998).

dependence in the DSGE models considered, the paper does investigate whether the estimated statistic – the impact of state-specific government spending on state-specific output and employment – varies with the unemployment rate of the U.S. economy in a given period. The effects of military spending on output and employment are roughly twice as large as in times in which the national unemployment rate is above its median as when it is below its median. Thus this evidence also suggests an important empirical role for slack resources, one that structural modelling at the moment does not match and one which again provides both motivation for such a nonlinear DSGE model and some clues as to what model ingredients might be required.

While I discuss these few examples, there any many papers that estimate the effects of fiscal stabilization policies on partial-equilibrium behaviors or outcomes, including Mian and Sufi (2010), Nekarda and Ramey (2011), and House and Shapiro (2008), that could be used to provide more discipline for macroeconomic modelling. There are also many more papers in macroeconomics that employ microeconomic evidence to discipline their theoretical or quantitative investigations. The point of this section is not to survey either set of papers, but to emphasize that cleanly-identified microeconomic estimates of partial-equilibrium causal effects of policies are both independent and relevant and so can provide additional evidence to improve the informativeness of model-based estimates of the size of the multiplier in recessions.

4 Conclusion

To date, much recent work on the effects of fiscal policy imposes that its impact on consumption, output and other economic outcomes is the same in a booming economy as in the depths of a recession. Estimates based on this assumption were used to evaluate the potential effects of fiscal policy in the Great Recession. It seems desirable to relax this assumption. Some theoretical and some empirical work that allows state-dependence in the effects of policy suggests that state dependence may be quite important, and that recessions, or only some recessions, may be times when fiscal policy is particularly potent. But an important difficulty with further investigation is the limited macroeconomic data available on the effects of policy

in recessions (or deep recessions).

Thus, this article argues that a fruitful avenue for ongoing and future research is to build (possibly highly) nonlinear models and match their implications not just to correlations in aggregate or microeconomic data, but also to estimates from microeconomic studies of the causal effects of policies on economic outcomes taking prices as given. Microeconomic studies can estimate for example the direct impact of countercyclical policy on demand, which measures the effect of policy on microeconomic spending and investment, excluding the multiplier. These studies provide largely unexploited quantitative discipline for structural macroeconomic models designed to give quantitative evaluation of the multiplier.

References

Aiyagari, R. (1994): "Uninsured Idiosyncratic Risk and Aggregate Savings," *Quarterly Journal of Economics*, 109, 659–684.

Aiyagari, S. R., L. J. Christiano, and M. Eichenbaum (1992): "The output, employment, and interest rate effects of government consumption," *Journal of Monetary Economics*, 30(1), 73–86.

Auerbach, A. J., and W. Gale (2010): "Activist Fiscal Policy to Stabilize Economic Activity," in *Financial Stability and Macroeconomic Policy: a symposium sponsored by the Federal Reserve Bank of Kansas City*, pp. 327–74. Federal Reserve Bank of Kansas City.

Auerbach, A. J., and Y. Gorodnichenko (2010): "Measuring the Output Responses to Fiscal Policy," NBER Working Papers 16311, National Bureau of Economic Research, Inc.

Barro, R. J. (2009): "Government Spending Is No Free Lunch," *The Wall Street Journal*, p. A17.

Barro, R. J., and R. G. King (1984): "Time-separable Preferences and Intertemporal-Substitution Models of Business Cycles," *The Quarterly Journal of Economics*, 99(4), 817–39.

Bernanke, B., and M. Gertler (1989): "Agency Costs, Net Worth, and Business Fluctuations," *The American Economic Review*, 79(1), 14–31.

Blanchard, O., and R. Perotti (2002): "An Empirical Characterization Of The Dynamic Effects Of Changes In Government Spending And Taxes On Output," *The Quarterly Journal of Economics*, 117(4), 1329–1368.

Brunnermeier, M. K., and Y. Sannikov (2011): "A Macroeconomic Model with a Financial Sector," Princeton University Manuscript.

Burnside, C., M. Eichenbaum, and J. D. M. Fisher (2004): "Fiscal shocks and their consequences," *Journal of Economic Theory*, 115(1), 89–117.

Chetty, R., A. Guren, D. S. Manoli, and A. Weber (2011): "Does Indivisible Labor Explain the Difference Between Micro and Macro Elasticities? A Meta-Analysis of Extensive Margin Elasticities," NBER Working Papers 16729, National Bureau of Economic Research, Inc.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (1998): "Modeling Money," NBER Working Papers 6371, National Bureau of Economic Research, Inc.

Christiano, L. J., M. Eichenbaum, and S. Relebo (2011): "When is the Government Spending Multiplier Large?," *Journal of Political Economy*, 119(1), 78–121.

Cochrane, J., and M. Piazzesi (2002): "The Fed and Interest Rates - A High-Frequency Identification," *American Economic Review*, 92(2), 90–95.

Eggertsson, G. B. (2008): "Great Expectations and the End of the Depression," *American Economic Review*, 98(4), 1476–1516.

Eggertsson, G. B., and M. Woodford (2003): "The Zero Bound on Interest Rates and Optimal Monetary Policy," *Brookings Papers on Economic Activity*, 34(1), 139–235.

Favero, C., F. Giavazzi, and J. Perego (2011): "Country Heterogeneity and the International Evidence on the Effects of Fiscal Policy," Manuscript for conference on fiscal policy, stabilization, and sustainability, International Monetary Fund.

Fernandez-Villaverde, J., and J. F. Rubio-Ramirez (2007): "Estimating Macroeconomic Models: A Likelihood Approach," *Review of Economic Studies*, 74(7), 1059–1087.

Fisher, J. D. M., and R. Peters (2009): "Using stock returns to identify government spending shocks," Working paper series, Federal Reserve Bank of Chicago.

Galí, J., J. D. López-Salido, and J. Vallés (2007): "Understanding the Effects of Government Spending on Consumption," *Journal of the European Economic Association*, 5(1), 227–270.

Gordon, R. J., and R. Krenn (2010): "The End of the Great Depression 1939-41: Policy Contributions and Fiscal Multipliers," NBER Working paper 16380.

Hall, R. E. (2009): "By How Much Does GDP Rise if the Government Buys More Output?," *Brookings Papers on Economic Activity*, 2, 183–231.

House, C., and M. D. Shapiro (2008): "Temporary Investment Tax Incentives: Theory with Evidence from Bonus Depreciation," *American Economic Review*, 98(3), 737–68.

Huntley, J., and V. Michelangeli (2010): "Tax Rebates and Fiscal Multipliers," Manuscript, Congressional Budget Office.

Ilzetzki, E., E. G. Mendoza, and C. Vegh (2010): "How Big (Small?) are Fiscal Multipliers?," NBER Working Papers 16479, National Bureau of Economic Research, Inc.

Johnson, D. S., J. A. Parker, and N. S. Souleles (2006): "Household Expenditure and the Income Tax Rebates of 2001," *American Economic Review*, 96(5), 1589–1610.

Kaplan, G., and G. Violante (2011): "A Model of the Consumption Response to Fiscal Stimulus Payments," Slides.

Kim, J., S. H. Kim, E. Schaumburg, and C. Sims (2005): "Calculating and Using Second Order Accurate Solutions of Discrete Time," Discussion paper, Princeton University, Department of Economics.

Krugman, P. (1998): "It's Baaack: Japan's Slump and the Return of the Liquidity Trap," *Brookings Papers on Economic Activity*, 2, 137–187.

——— (2009): "The Obama Gap," *The New York Times*, p. A27.

Krusell, P., and A. A. Smith (1998): "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106(5), 867–96.

Leeper, E., C. Sims, and T. Zha (1996): "What Does Monetary Policy Do?," *Brookings Papers on Economic Activity*, 27(2), 1–78.

Mankiw, N. G. (1986): "The Allocation of Credit and Financial Collapse," *The Quarterly Journal of Economics*, 101(3), 455–470.

Mian, A., and A. Sufi (2010): "The Effects of Fiscal Stimulus: Evidence from the 2009 ŚCash for ClunkersŠ Program," NBER Working Papers 16351, National Bureau of Economic Research, Inc.

Nakamura, E., and J. Steinsson (2011): "Fiscal Stimulus in a Monetary Union: Evidence from U.S. Regions," Columbia Department of Economics, mimeo.

Nekarda, C. J., and V. A. Ramey (2011): "Industry Evidence on the Effects of Government Spending," *American Economic Journal: Macroeconomics*, 3(1), 36–59.

Oh, H., and R. Reis (2011): "Targeted Transfers and the Fiscal Response to the Great Recession," NBER Working Papers 16775, National Bureau of Economic Research, Inc.

Parker, J. A., N. S. Souleles, D. S. Johnson, and R. McClelland (2011): "Consumer Spending and the Economic Stimulus Payments of 2008," NBER Working Papers 16684, National Bureau of Economic Research, Inc.

Perotti, R. (1999): "Fiscal Policy In Good Times And Bad," *Quarterly Journal of Economics*, 114(4), 1399–1436.

Prescott, E. C. (1986): "Theory Ahead of Business-Cycle Measurement," *Carnegie-Rochester Conference Series on Public Policy*, 25(1), 11–44.

Ramey, V. A. (2011): "Identifying Government Spending Shocks: It's all in the Timing," *Quarterly Journal of Economics*, 126(1), 1–50.

——— (forthcoming): "Can Government Purchases Stimulate the Economy?," *Journal of Economic Literature*.

Ramey, V. A., and M. D. Shapiro (1998): "Costly capital reallocation and the effects of government spending," *Carnegie-Rochester Conference Series on Public Policy*, 48(1), 145–194.

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

Romer, C. (1992): "What Ended the Great Depression?," *Journal of Economic History*, 52(4), 757–84.

Romer, C. D., and D. Romer (2010): "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks," *American Economic Review*, 100(3), 763–801.

Rotemberg, J. J., and M. Woodford (1997): "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy," in *N.B.E.R. Macroeconomics Annual, 1991*, ed. by B. Bernanke, and J. J. Rotemberg, vol. 12, pp. 297–346. Cambridge, MA: MIT Press.

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

Tagkalakis (2008): "The Effects of Fiscal Policy on Consumption in Recessions and Expansions," *Journal of Public Economics*, 92, 1486–1508.

Taylor, J. (forthcoming): "An Empirical Analysis of the Revival of Fiscal Activism in the 2000s," *Journal of Economic Literature*.

WOODFORD, M. (2011): "Simple Analytics of the Government Expenditure Multiplier," *American Economic Journal: Macroeconomics*, 3(1), 1–35.