Is Economics Research Replicable? Sixty Published Papers from Thirteen Journals Say "Usually Not"

Andrew C. Chang* and Phillip Li†

December 10, 2015

Abstract

We attempt to replicate 67 macroeconomic articles published in 13 well-regarded economics journals using author-provided replication files that include both data and code. Some journals in our sample require data and code replication files, and other journals do not require such files. Aside from 6 papers that use confidential data, we obtain data and code replication files for 29 of 35 papers (83%) that are required to provide such files as a condition of publication, compared to 11 of 26 papers (42%) that are not required to provide data and code replication files. Defining replication success as our ability to use the author-provided data and code files to produce the key qualitative conclusions of the original paper, we successfully replicate 22 of 67 papers (33%) without contacting the authors. Excluding the 6 papers that use confidential data and the 2 papers that use software we do not possess, we replicate 29 of 59 papers (49%) with assistance from the authors. Because we are able to replicate less than half of the papers in our sample even with help from the authors, we assert that economics research is usually not replicable. We conclude with recommendations on improving replication of economics research.

*Chang: Board of Governors of the Federal Reserve System. 20th St. NW and Constitution Ave., Washington DC 20551 USA. +1 (657) 464-3286. a.christopher.chang@gmail.com. https://sites.google.com/site/andrewchristopherchang/.

†Li: Office of the Comptroller of the Currency. phillip.li@occ.treas.gov.

‡The views and opinions expressed here are those of the authors and are not necessarily those of the Board of Governors of the Federal Reserve System, the Department of the Treasury, or the Office of the Comptroller of the Currency. We thank Jan H. Höeffler, Lawrence Katz, Evan F. Koenig, B.D. McCullough, Michael Nahas, Min Qi, W. Robert Reed, Bo Sun, Richard H. Thaler, J. Luke Van Cleve, Kurt von Tish, Christian Zimmermann, and seminar participants at FRB Cleveland, OCC, OFR, Penn State, and UC - Irvine for helpful comments. We thank Tyler J. Hanson, Erik Larsson, Kim T. Mai, Anthony Marcozzi, Shawn M. Martin, Tyler Radler, Adam Scherling, and John Stromme for research assistance. We also thank Felix Galbis-Reig and Spencer C. Li for technical assistance. We are responsible for any errors.

JEL Codes: B41; C80; C82; C87; C88; E01
Keywords: Data and Code Archives; Gross Domestic Product; GDP; Journals; Macroeconomics; National Income and Product Accounts; Publication; Research; Replication

1 Introduction

In response to McCullough and Vinod (2003)'s failed replication attempt of several articles in the *American Economic Review* (AER), then-editor of the AER Ben Bernanke strengthened the AER's data and code availability policy to allow for successful replication of published results by requiring authors to submit to the AER data and code replication files (Bernanke, 2004). Since the AER strengthened its policy, many of the other top journals in economics, such as *Econometrica* and the *Journal of Political Economy*, also started requiring data and code replication files.

There are two main goals of these replication files: (1) to bring economics more in line with the natural sciences by embracing the scientific method's power to verify published results, and (2) to help improve and extend existing research, which presumes the original research is replicable. These benefits are illustrated by the policy-relevant debates between Card and Krueger (1994, 2000) and Neumark and Wascher (2000) on minimum wages and employment; Hoxby (2000, 2007) and Rothstein (2007) on school choice; Levitt (1997, 2002) and McCrary (2002) on the causal impact of police on crime; and, more recently, Reinhart and Rogoff (2010) and Herndon, Ash, and Pollin (2014) on fiscal austerity. In extreme cases, replication can also facilitate the discovery of scientific fraud, as in the case of Broockman, Kalla, and Aronow (2015)'s investigation of the retracted article by LaCour and Green (2014).

This article is a cross-journal, broad analysis of the state of replication in economics.[1] We attempt to replicate articles using author-provided data and code files from 67 papers

[1] We follow existing work by, among others, Dewald, Thursby, and Anderson (1986); McCullough and Vinod (2003, 2004); McCullough, McGeary, and Harrison (2006); Hamermesh (2007); Glandon (2010) and Duvendack, Palmer-Jones, and Reed (2015).

published in 13 well-regarded general interest and macroeconomics journals from July 2008 to October 2013. This sampling frame is designed to be more comprehensive across well-regarded economics journals than used by existing research. Previous work has tended to focus on a single journal, such as McCullough, McGeary, and Harrison (2006), who look at the *Journal of Money, Credit and Banking* (JMCB); McCullough and Vinod (2003), who attempt to replicate a single issue of the AER (but end up replicating only Shachar and Nalebuff (1999) with multiple software packages); or Glandon (2010), who replicates a selected sample of nine papers only from the AER.

We define a successful replication as when we can use author-provided files to qualitatively produce the key results of the paper.[2] Under this definition, we are able to replicate 22 of 67 papers (33%) independently of the authors by following the instructions in the author-provided readme files. The most common reason we are unable to replicate the remaining 45 papers is that the authors do not provide data and code replication files. We find that some authors do not provide data and code replication files even when their article is published in a journal with a policy that requires submission of such files as a condition of publication, indicating that editorial offices do not strictly enforce these policies, although provision of replication files is more common at journals that have such a policy than at journals that do not. Excluding 6 papers that rely on confidential data for all of their results and 2 papers that provide code written for software we do not possess, we successfully replicate 29 of 59 papers (49%) with help from the authors. Because we successfully replicate less than half of the papers in our sample even with assistance from the authors, we conclude that economics research is usually not replicable.[3]

Despite our finding that economics research is usually not replicable, our replication success rates are still notably higher than those reported by existing studies of replication in economics. McCullough, McGeary, and Harrison (2006) find a replication success rate

[2]There are widespread differences in the definitions of "replication." See, for example, Clemens (2015).

[3]This result is in line with recent evidence in replicating psychology studies by Open Science Collaboration (2015), where the authors fail to replicate the qualitative result of the majority of their sample of psychology experiments.

for articles published in the JMCB of 14 of 186 papers (8%), conditioned on the replicators' access to appropriate software, the original article's use of non-proprietary data, and without assistance from the original article's authors. Adding a requirement that the JMCB archive contain data and code replication files the paper increases their success rate to 14 of 62 papers (23%). Our comparable success rates are 22 of 59 papers (37%), conditioned on our having appropriate software and non-proprietary data, and 22 of 38 papers (58%) when we impose the additional requirement of having data and code files. Dewald, Thursby, and Anderson (1986) successfully replicate 7 of 54 papers (13%) from the JMCB, conditioned on the replicators having data and code files, the original article's use of non-confidential data, help from the original article's authors, and appropriate software. Our comparable figure is 29 of 38 papers (76%).

2 Methodology and Sampling Frame

Our sampling frame includes papers from 13 well-regarded macroeconomics and general interest economics journals: *American Economic Journal: Economic Policy*, *American Economic Journal: Macroeconomics*, *American Economic Review*, *American Economic Review: Papers and Proceedings* (P&P), *Canadian Journal of Economics*, *Econometrica*, *Economic Journal*, *Journal of Applied Econometrics*, *Journal of Political Economy*, *Review of Economic Dynamics*, *Review of Economic Studies*, *Review of Economics and Statistics*, and *Quarterly Journal of Economics*. We choose papers from these journals because of the relative likelihood that such papers will have a policy effect and also influence future research.[4] We do not select these journals to single out a particular author, methodology, institution, or ideology.

From our sample of journals, we browse for original research articles published in issues

[4] For example, according to the 2014 social science edition of Thompson Reuters Journal Citation Reports, these journals represent half of the top 10 impact factor journals in economics and, aside from the *Canadian Journal of Economics*, all journals are in the top impact factor quartile for economics.

from July 2008 to October 2013.[5,6] Within these issues, we identify all papers with the following three characteristics: (1) an empirical component, (2) model estimation with only US data, and (3) a key empirical result produced by inclusion of US gross domestic product (GDP), published by the Bureau of Economic Analysis (BEA), in an estimated model.[7] We choose to focus on GDP because of its status as a standard macroeconomic statistic and its widespread use in research.[8]

For each paper in this set, we attempt to replicate the key empirical results.[9] We focus on the key empirical results for two reasons: (1) replicating only the key results allows us to expand the sample to more papers, and (2) the key result of the paper is presumably what drove the paper's publication; robustness checks merely serve as confirming evidence.

Defining a key result is subjective and requires judgmental decisions on our part. We attribute a key result of the paper to GDP when the authors themselves refer to GDP as driving a key result, or when a discussion of GDP is featured either in the abstract or prominently in the introduction of their work (or both). We also take key results as those that appear in figures and tables.[10]

We find 67 papers that fit these criteria. Of these papers, 6 papers use proprietary data for all of the key results, so we do not include them in our replication exercise (Fisher and Peters, 2010; Alexopoulos, 2011; Alexopoulos and Cohen, 2011; Hall and Sargent, 2011; Bansak, Graham, and Zebedee, 2012; Gilchrist and Zakrajšek, 2012). If a subset of the key results could be obtained using non-proprietary data, then we attempt to replicate those

[5] This definition excludes lectures, symposiums, and literature reviews. We also exclude articles published prior to July 2008 when a comment on the article is published on or after July 2008.

[6] We choose an arbitrary start date of July 2008. We select the end date of October 2013 to match when we began this project.

[7] The authors may calibrate a subset of their model's parameters and still have their paper fall within our sample. We exclude papers that have only completely calibrated models.

[8] Several of the most cited articles in economics use GDP or its predecessor gross national product (GNP) (Sims, 1980; Kydland and Prescott, 1982; Barro, 1991; Hodrick and Prescott, 1997). Recognizability of GDP extends to the popular press as well. For example, on the HBO late-night talk show "Last Week Tonight," John Oliver cites GDP as a measure of the importance of the District of Columbia (Last Week Tonight, 2015).

[9] A paper's primary contribution may be theoretical, but we focus only on the empirical component.

[10] The web appendix details what we interpret as the key result of every paper in our sample and is available on Chang's website, https://sites.google.com/site/andrewchristopherchang/research.

results.

For the remaining papers that use public data and are published in journals that maintain data and code archives, we download the replication files provided by the authors through the online archives provided by the journals. Unlike prior work by McCullough, McGeary, and Harrison (2006), who found difficulty in accessing the archives of selected journals, we had no trouble doing so through the Board of Governors of the Federal Reserve System or Office of the Comptroller of the Currency subscriptions. However, consistent with McCullough, McGeary, and Harrison (2006) and Vlaeminck and Herrmann (2015), we find that journal data and code archives are incomplete. Of the 35 papers that use public data and are published in journals that require data and code replication files, we obtain files for 28 papers (80%) from journal archives.

For papers where we are unable to obtain replication data and code files from journal archive sites, either because the mandatory files are is missing or because the paper is not subject to a data availability policy, we check the personal websites of each of the authors for replication files. If we are unable to locate replication files online, then we email each of the authors individually requesting the replication files.[11] Of the 7 papers that use public data, are subject to a data and code policy, and do not have replication files on the journal's archive site, this procedure nets us one additional set of replication files. Therefore, we are unable to locate replication files for 6 of 35 papers (17%) that are published in journals that require submission of data and code replication files. For papers published in journals without a data and code availability policy and that use public data, we are unable to obtain data and code

[11] We email each of the authors sequentially using the email addresses obtained in the following three-step manner, moving down a step when our email was undeliverable: (1) the address listed on the author's personal site, (2) the address listed on the author's current institution's site, and (3) the address listed with the published paper. We wait at least one week between contacting each author. If a corresponding author is listed on the paper, then we start with contacting the corresponding author and work our way through the authors in the order they are presented in the paper. If there is no corresponding author, then we initiate contact with the first author. We stop querying different authors to request data and code after receiving a response from any author, unless we are specifically directed to by an author we had already queried or we fail to receive a response after an author-requested delay. We mark the data and code files as incomplete when we both are unable to locate a complete set of files online and also do not receive a complete set of files in response to our emails after waiting a minimum of one month for each author to respond to our queries. We stopped considering responses as of July 15th, 2015.

replication files for 15 of 26 papers (58%). We do not single out any paper or author that fails to comply with a journal's mandatory data and code policy, both because we are interested in the general state of replication in economics and also because of the potential ill-will that singling out a particular paper or author could generate (Zimmermann, 2015). We therefore only report these summary statistics of compliance with data availability policies and only cite papers that we either successfully replicate, that use proprietary data, or where we have what appears to be a complete set of replication files in a software we do not possess.

To determine whether a paper was subject to a data availability policy, we check the implementation dates of the journal data policies and compare them to the publication and submission dates of the published work. If the journal's website does not allow us to extract this information, then we query the editorial office as to when their data availability policy became effective. We do not ask the editorial offices whether a particular paper was subject to a data availability policy. Aside from papers with proprietary data, we find that journal data archives do not provide lists of potentially exempt papers. Therefore, we are unable to determine whether a paper is exempt for a reason other than using proprietary data, although we are not aware of reasons why journals would grant a paper a data and code exemption other than for proprietary data. The authors we query whose papers we believe are subject to a data availability policy yet whose replication files we are unable to locate do not volunteer whether their papers are exempt from the policy, and we do not ask the authors for this information.

For the papers for which we are able to obtain data and code replication files, we attempt to replicate the key results of the paper using only the instructions provided in the author readme files. If the readme files are insufficient or if the replication files are incomplete (or both) and the paper is subject to a replication policy, then we email the corresponding author (if no corresponding author, then the first author) for either clarification or to request the missing files. If we do not receive a response within a week, then we query the second author, and so on, until all authors on the paper had been contacted.[12]

We define a successful replication as when the authors or journal provide data and code files that allow us to qualitatively reproduce the key results of the paper. For example, if the paper estimates a fiscal multiplier for GDP of 2.0, then any multiplier greater than 1.0 would produce the same qualitative result (i.e., there is a positive multiplier effect and that government spending is not merely a transfer or crowding out private investment).[13] We define success using this extremely loose definition to get an upper bound on what the replication success rate could potentially be.[14] We allow for minimal re-working of the provided files, following the procedure of McCullough, McGeary, and Harrison (2006).[15]

One dimension where we are unable to follow the authors exactly is the software version they use. To execute the replications, we make use of the following software version-operating system combinations: Dynare 4.3 and 4.4.2 (Windows), EViews 6, 7, and 8 (Windows), Gauss 9.0.2 (Linux), FAME 10.2 64-bit, Fortran f90 (Linux), Matlab R2008a and R2012a and R2013a (Windows), Matlab R2010a and R2012a (Linux), OX 6.30 (Windows), Oxmetrics 6.30 (Windows), Stata 11.0 and 13.1 (Windows), Stata 13.0 (Windows and Linux), R 2.15.1 and 3.0.1 and 3.0.2 and 3.0.3 and 3.1.0 (Linux), and RATS 7.10 (Linux).[16] When available in the readme, we attempt to run the software version-operating system combination specified by the authors. When the replication files fail to execute on a given software

[12]If we already contacted the authors to request data or code but were having difficulty executing the code, then we only queried the authors whom we did not yet contact. We initiate contact with each author a maximum of one time.

[13]This definition corresponds to replication - verification by Clemens (2015)'s Table 1 with the added condition of the authors providing us data and code files, a rating of three out of five, "minor discrepancies," or better by Glandon (2010), and "partially successful replication" or better by McCullough, McGeary, and Harrison (2006).

[14]This definition is less stringent than the definition for replication success of McCullough, McGeary, and Harrison (2006, 2008)

[15]McCullough, McGeary, and Harrison (2008), in their appendix, suggest that "the author [whose study is being replicated] provides code such that data and code, when placed in the same subdirectory, will execute; and that the output from doing this also will be provided... and produces the results in his paper," which implies that replication files should contain the data and code that requires no re-working. If the code is clearly missing the ability to replicate results, then we do not attempt to re-code the procedures ourselves.

[16]We check the replication results of a small sample of selected papers across different versions of Matlab for Windows and find very minor differences. None of the differences in results across different versions of Matlab are qualitatively significant. The web appendix lists the software version-operating system combination we use for each successfully replicated paper.

version-operating system combination, the author readme did not specify a particular software version-operating system combination, and it appeared that the data and code were complete, we email the authors to find out which combination they use.

3 Results

This section presents summary statistics of our replication attempts.[17]

Table 1 lists the papers we successfully replicate. Table 2 breaks down our replication results by journal type. Panel A of Table 2 shows that our overall replication success rate is 29 of 67 papers (43%).

Table 2, Panel B shows that we successfully replicate 23 of 39 papers (59%) from journals that require data and code replication files. This rate compares to 6 of 28 (21%) of the papers from journals that do not require such files, shown in Table 2, Panel C. These replication rates are similar when we only consider papers with publicly available data: we successfully replicate 23 of 35 (66%) of the papers from journals with mandatory data and code policies and 6 of 26 (23%) of the papers from journals without such policies. The presence of a mandatory data and code policy does not necessarily imply a causal relationship from the policy to successful replication. Authors select which journals to submit papers to, taking into account idiosyncratic journal policies such as mandatory submission of replication data and code. However, we find that it is significantly easier to replicate published research that comes from journals that require authors to submit their data and code.

Table 3, Panel A provides explanations for why we are unable to replicate papers according to four broad classifications: "missing public data or code," "incorrect public data or code," "missing software," or "proprietary data." Panel B provides the breakdown for journals that require data and code. Panel C shows the results for journals that do not require data and code.

[17] Interested readers can find detailed results for each paper we successfully replicate in the web appendix on Chang's website, https://sites.google.com/site/andrewchristopherchang/research.

From Table 3, Panel A we find that we are unable to replicate 21 papers because of "missing data or code," which constitutes the majority of our failed replications (55%). As we outline in our methodology, for each of these unsuccessful replications we attempt to secure data and code from the authors by visiting their personal websites, visiting the journal websites (when the journal requires authors to submit data or code), and sending email requests. We classify an unsuccessful replication as "missing data or code" when at least one of two events occur: (1) the replication code file(s) are clearly missing necessary author-written functions for a subset or all of the key results or (2) the replication data file(s) are missing at least one variable. If the replication data has a shorter data sample than reported in the paper, then we still attempt the estimation and do not necessarily classify the paper as "missing data or code."

We are unable to replicate 9 papers (24% of failed replications) because of "incorrect data or code." We classify an unsuccessful replication as "incorrect data or code" when all variables are present in the dataset and the authors self-identify code for each of the key figures and tables we attempt to replicate. The author-provided code may finish executing and give different results or the code may not finish executing and still fall into this category.

We believe we do not have the needed software to run two papers (Senyuz, 2011; Jermann and Quadrini, 2012) because we are unable to locate a necessary packaged function in our versions of the appropriate software, because of significant syntax changes between software versions, or because the authors declared that they use a particular software version and we are aware that our software would not be compatible. However, it is tricky differentiating between an unsuccessful replication due to "incorrect data or code" or due to "missing software." Because the implementation of packaged functions may differ across software versions even without syntax changes, we believe the number of failed replications we classify as "missing software" is a lower bound. It is possible that a paper we classify as "incorrect data or code" is actually replicable with the appropriate operating system-software combination, so some of the papers that we classify as "incorrect data or code" may belong in the "missing software" category. However, we cannot verify this statement without additional documentation.

Table 4 shows our summary statistics for successful replications independent of the authors versus replications that were successful with the author's help. Overall, we find that contacting the authors marginally improves our success rate for replication. Of the 29 successful replications, we complete 22 without any help from the authors.

4 Conclusion and Recommendations

In this article, we attempt to replicate 67 papers from 13 well-regarded economics journals using author-provided data and code replication files. Improving on existing work evaluating the state of replication in economics, our sampling frame is broader across different journals and covers a large number of original research articles. We replicate 22 of 67 papers (33%) by using only the authors' data and code files, and an additional 7 papers (for a total of 29 papers, 43%) with assistance from the authors. The most common cause of our inability to replicate findings is that authors do not provide files to the journal replication archives, which constitutes approximately half of our failed replication attempts (21 of 38 papers, 55%). Because we are able to replicate less than half of the papers in our sample, we conclude that economics research is generally not replicable.

We now turn to some recommendations that we feel would improve the ability for researchers to replicate and extend published articles, largely echoing the recommendations of McCullough, McGeary, and Harrison (2006).

- *Mandatory data and code files should be a condition of publication.*

Our replication success rate is significantly higher when we attempt to replicate papers from journals that have a mandatory replication data and code submission policy. We believe that replication files need to encompass both data and code. As shown in Table 2, the data-only archives at *Economic Journal* and *Journal of Applied Econometrics* only allow for replication of 4 of 20 papers (20%) that use non-confidential data, compared to the replication success rate of 23 of 35 papers (66%) that use non-confidential data from journals that require both data and code.

- *An entry in the journal's data and code archive should indicate whether a paper without replication files in the journal's archive is exempt from the journal's replication policy.*

Among papers that we believe were subject to a mandatory data and code policy, we are unable to acquire replication files for 6 of 35 papers (15%) even after emailing, and often receiving a response from, the authors. However, we are unsure whether these six papers are exempt from their respective journal's mandatory data and code policies, and the authors did not volunteer whether their papers are exempt in response to our requests for replication files. Therefore, we suggest that journals include an exemption entry in their replication archives. This note in the replication archives would have four virtues: (1) it is low-cost for the journal, (2) it would save authors who are exempt from submitting replication files from needing to respond to queries about replication files, (3) it would save would-be-replicators from searching for replication files for papers that are exempt from the journal's policy, and (4) it would identify those authors who are not compliant with the journal's mandatory data and code policy.

- *Readme files should indicate the operating system-software version combination used in the analysis.*

We attempt to use the operating system-software version combination reported by the authors in their readme files, but we notice that very few readmes include the operating system-software version combination used to conduct their analysis. When we ask authors about the operating system or software version they use to run their models, most authors do not recall this information. Although it is not a focus of our paper, we notice minor discrepancies for a selected subset of papers when running programs on different versions of Matlab (although the discrepancies are not large enough to change the key qualitative results).

- *Readme files should contain an expected model estimation time.*

Many macroeconomic models are estimated with Bayesian (i.e., Markov Chain Monte Carlo) methods, which can take a considerable amount of processing time to execute even under the best of circumstances. We encountered a few instances where we believed an estimation was executing, only to find out weeks later that the programs were stuck in an infinite loop and were supposed to run in much less time. In addition, frequently programs are not written to optimize computation time and also frequently written without a progress bar, so there is no way to track the expected completion time of estimation. A low-cost alternative to a progress bar is simply writing the expected estimation time in the readme file.[18]

- *Code that relies on random number generators should set seeds and specify the random number generator.*

Optimization algorithms often rely on a set of initial conditions, which are commonly specified through a random number generator. For any research that relies on a random number generator, replication requires the same set of numbers that are generated in the published article.[19]

- *Readme files should clearly delineate which files should be executed in what order to produce desired results.*

Occasionally, we are presented with replication data and code that requires files to be executed in a particular order to furnish published results. In cases where the execution order is critical but unspecified, we spend a considerable amount of time attempting to determine the proper order of execution and, in some cases, ultimately fail to do so.

We now turn to two recommendations that will improve the ability of researchers to extend published work, in addition to merely replicating it.

[18] We are happy to report that, in the days after the initial working paper version of this article was posted to the Board of Governors of the Federal Reserve System Finance and Economics Discussion Series on October 5th, 2015, the *Review of Economic Dynamics* changed its data and code policy to include several of our recommendations.

[19] As researchers we hope that the conclusions of a given article would be robust to the random number generator and seed used. For purposes of exactly replicating an author's result, the random number generator and seed would be required.

- *Authors should provide raw data in addition to transformed series.*

While only the transformed data are needed to conduct replication of published results, raw data facilitate potential extensions of research.[20] For example, raw data allow for the investigation of the effect that revisions to macroeconomic data have on previously published research, as in Croushore and Stark (2003) and Chang and Li (2015).

- *Programs that replicate estimation results should carry out the estimation.*

We notice that the replication files for a few papers run smoothly and exactly furnish the results of the tables and figures as published. However, oftentimes the results in tables and figures depend on a model's parameters being estimated. Some of these replication files, instead of estimating the models, take the relevant parameters as given to produce results in tables and figures. For verification of published results, and particularly for purposes of extending research, we assert that code that actually estimates the relevant models would be far more useful.

[20] This point is also mentioned by Duvendack, Palmer-Jones, and Reed (2015). The raw data would also have pedagogical uses, as outlined by Höffler (2014).

Table 1: Successfully Replicated Papers

Auerbach and Gorodnichenko (2012)
Barro and Redlick (2011)
Baumeister and Peersman (2013)
Canova and Gambetti (2010)
Carey and Shore (2013)
Chen, Curdia, and Ferrero (2012)
Clark and McCracken (2010)
Corsetti, Meier, and Müller (2012)
D'Agostino and Surico (2012)
Den Haan and Sterk (2011)
Favero and Giavazzi (2012)
Gabaix (2011)
Hansen, Lunde, and Nason (2011)
Heutel (2012)
Inoue and Rossi (2011)
Ireland (2009)
Kilian (2009)
Kormilitsina (2011)
Krishnamurthy and Vissing-Jorgensen (2012)
Mavroeidis (2010)
Mertens and Ravn (2011)
Mertens and Ravn (2013)
Morley and Piger (2012)
Nakov and Pescatori (2010)
Ramey (2011)
Reis and Watson (2010)
Romer and Romer (2010)
Schmitt-Grohé and Uribe (2011)
Schmitt-Grohé and Uribe (2012)

We replicate the corrected results of Auerbach and Gorodnichenko (2012) found in Auerbach and Gorodnichenko (2013).

Table 2: Replication Sample and Results By Journal

	Papers Replicated Successfully	Papers With Public Data	Total Papers
Panel A: All Journals	29	61	67
Panel B: Journals that Require Data and Code:			
American Economic Journal: Economic Policy	2	4	4
American Economic Journal: Macroeconomics	3	3	4
American Economic Review	5	8	10
Canadian Journal of Economics	0	0	1
Econometrica	3	3	3
Journal of Political Economy	1	1	1
Review of Economic Dynamics	4	7	7
Review of Economic Studies	0	2	2
Review of Economics and Statistics	5	7	7
Total for Journals that Require Data and Code	23	35	39
Panel C: Journals that Do Not Require Data and Code:			
American Economic Review: P&P	0	4	5
Economic Journal	3	10	11
Journal of Applied Econometrics	1	10	10
Quarterly Journal of Economics	2	2	2
Total for Journals that Do Not Require Data and Code	6	26	28

Journal of Applied Econometrics requires data only. Economic Journal currently requires data and code, but the papers in our sample were not subject to a data and code policy according to the Economic Journal's editorial office.

Table 3: Failed Replication Results, Including Causes of Failure, By Journal Type

	Paper Count	Percentage of Sample
Panel A: All Journals		
Replication Unsuccessful	38	100
Unsuccessful Because of:		
Missing Public Data or Code	21	55
Incorrect Public Data or Code	9	24
Missing Software	2	5
Proprietary Data	6	16
Panel B: Journals With Mandatory Data and Code Policies		
Replication Unsuccessful	16	100
Unsuccessful Because of:		
Missing Public Data or Code	6	38
Incorrect Public Data or Code	5	31
Missing Software	1	6
Proprietary Data	4	25
Panel C: Journals Without Mandatory Data and Code Policies		
Replication Unsuccessful	22	100
Unsuccessful Because of:		
Missing Public Data or Code	15	68
Incorrect Public Data or Code	4	18
Missing Software	1	5
Proprietary Data	2	9

Table 4: Successful Replication Results By Journal Type

	Paper Count	Percentage of Sample
Panel A: All Journals		
Replication Successful	29	100
Successful:		
With Contacting Authors	7	24
Without Contacting Authors	22	76
Panel B: Journals With Mandatory Data and Code Policies		
Replication Successful	23	100
Successful:		
With Contacting Authors	3	13
Without Contacting Authors	20	87
Panel C: Journals Without Mandatory Data and Code Policies		
Replication Successful	6	100
Successful:		
With Contacting Authors	4	67
Without Contacting Authors	2	33

References

Alexopoulos, Michelle, "Read All About it!! What Happens Following a Technology Shock?" American Economic Review 101:4 (2011), 1144-1179.

Alexopoulos, Michelle, and Jon Cohen, "Volumes of Evidence: Examining Technical Change in the Last Century Through a New Lens," Canadian Journal of Economics 44:2 (2011), 413-450.

Andrews, Donald W.K., "Tests for Parameter Instability and Structural Change with Unknown Change Point," Econometrica 61:4 (1993), 821-856.

Auerbach, Alan J., and Yuriy Gorodnichenko, "Measuring the Output Responses to Fiscal Policy," American Economic Journal: Economic Policy 4:2 (2012), 1-27.

Auerbach, Alan J., and Yuriy Gorodnichenko, "Corrigendum: Measuring the Output Responses to Fiscal Policy," American Economic Journal: Economic Policy 5:3 (2013), 320-322.

Bansak, Cynthia, Mary E. Graham, and Allan A. Zebedee, "Business Cycles and Gender Diversification: An Analysis of Establishment-Level Gender Dissimilarity," American Economic Review: Papers & Proceedings 102:3 (2012), 561-565.

Barro, Robert J., "Economic Growth in a Cross Section of Countries," Quarterly Journal of Economics 106:2 (1991), 407-443.

Barro, Robert J., and Charles J. Redlick, "Macroeconomic Effects from Government Purchases and Taxes," Quarterly Journal of Economics 126:1 (2011), 51-102.

Baumeister, Christiane, and Gert Peersman, "Time-Varying Effects of Oil Supply Shocks on the US Economy," American Economic Journal: Macroeconomics 5:4 (2013), 1-28.

Bernanke, Ben S., "Editorial Statement," American Economic Review 94:1 (2004), 404.

Broockman, David, Joshua Kalla, and Peter Aronow, "Irregularities in LaCour (2014)," Working Paper (2015).

Card, David, and Alan B. Krueger, "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania," American Economic Review 84:4 (1994), 772-793.

Card, David, and Alan B. Krueger, "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania: Reply," American Economic Review 90:5 (2000), 1397-1420.

Canova, Fabio, and Luca Gambetti, "Do Expectations Matter? The Great Moderation Revisited," American Economic Journal: Macroeconomics 2:3 (2010), 183-205.

Carey, Colleen, and Stephen H. Shore, "From the Peaks to the Valleys: Cross-State Evidence on Income Volatility Over the Business Cycle," Review of Economics and Statistics 95:2 (2013), 549-562.

Chang, Andrew C., and Phillip Li, "Measurement Error in Macroeconomic Data and Economics Research: Data Revisions, Gross Domestic Product, and Gross Domestic Income," Board of Governors of the Federal Reserve System FEDS Working Paper 2015-102 (2015).

Chen, Han, Vasco Curdia, and Andrea Ferrero, "The Macroeconomic Effects of Large-scale Asset Purchase Programmes," Economic Journal 122:564 (2012), F289-F315.

Clark, Todd E., and Michael W. McCracken, "Averaging Forecasts from VARs with Uncertain Instabilities," Journal of Applied Econometrics 25:1 (2010), 5-29.

Clemens, Michael, "The Meaning of Failed Replications: A Review and Proposal," Center for Global Development Working Paper 399 (2015).

Corsetti, Giancarlo, André Meier, and Gernot J. Müller, “Fiscal Stimulus with Spending Reversals,” Review of Economics and Statistics 94:4 (2012), 878-895.

Croushore, Dean, and Tom Stark, “A Real-Time Data Set for Macroeconomists: Does the Data Vintage Matter?” Review of Economics and Statistics 85:3 (2003), 605-617.

D’Agostino, Antonello, and Palo Surico, “A Century of Inflation Forecasts,” Review of Economics and Statistics 94:4 (2012), 1097-1106.

Den Haan, Wouter J., and Vincent Sterk, “The Myth of Financial Innovation and the Great Moderation,” Economic Journal 121:553 (2011), 707-739.

Dewald, William G., Jerry G. Thursby, and Richard G. Anderson, “Replication in Empirical Economics: The Journal of Money, Credit and Banking Project,” American Economic Review 76:4 (1986), 587-603.

Duvendack, Maren, Richard W. Palmer-Jones, and W. Robert Reed, “Replication in Economics: A Progress Report,” Econ Journal Watch 12:2 (2015), 164-191.

Evanschitzky, Heiner, Carsten Bamgarth, Raymond Hubbard, and J. Scott Armstrong, “Replication Research in Marketing Revisited: A Note on a Disturbing Trend,” Journal of Business Research 60:4 (2007), 411-415.

Favero, Carlo, and Giavazzi, Francesco, “Measuring Tax Multipliers: The Narrative Method in Fiscal VARs,” American Economic Journal: Economic Policy 4:2 (2012), 69-94.

Fisher, Jonas D.M., and Ryan Peters, “Using Stock Returns to Identify Government Spending Shocks,” Economic Journal 120:544 (2010), 414-436.

Gabaix, Xavier, “The Granular Origins of Aggregate Fluctuations,” Econometrica 79:3 (2011), 733-772.

Glandon, Philip, “Report on the American Economic Review Data Availability Compliance Project,” Working Paper (2010).

Gilchrist, Simon, and Egon Zakrajšek, "Credit Spreads and Business Cycle Fluctuations," American Economic Review 102:4 (2012), 1692-1720.

Hall, George J., and Thomas J. Sargent, "Interest Rate Risk and Other Determinants of Post-WWII US Government Debt/GDP Dynamics," American Economic Journal: Macroeconomics 3:3 (2011), 192-214.

Hamermesh, Daniel S., "Viewpoint: Replication in Economics," Canadian Journal of Economics 40:3 (2007), 715-733.

Hansen, Peter R., Asger Lunde, and James M. Nason, "The Model Confidence Set," Econometrica 79:2 (2011), 453-497.

Herndon, Thomas, Michael Ash, and Robert Pollin, "Does High Public Debt Consistently Stifle Economic Growth? A Critique of Reinhart and Rogoff," Cambridge Journal of Economics 38:2 (2014), 257-279.

Heutel, Garth, "How Should Environmental Policy Respond to Business Cycles? Optimal Policy Under Persistent Productivity Shocks," Review of Economic Dynamics 15:2 (2012), 244-264.

Hodrick, Robert J., and Edward C. Prescott, "Postwar U.S. Business Cycles: An Empirical Investigation," Journal of Money, Credit and Banking 29:1 (1997), 1-16.

Hoxby, Caroline, "Does Competition Among Public Schools Benefit Students and Taxpayers," American Economic Review 90:5 (2000), 1209-1238.

Hoxby, Caroline, "Does Competition Among Public Schools Benefit Students and Taxpayers? A Reply to Rothstein," American Economic Review 97:5 (2007), 2038-2055.

Höffler, Jan H., "Teaching Replication in Quantitative Empirical Economics," Working Paper for the 68th European Meeting of the Econometric Society (2014).

Inoue, Atsushi, and Barbara Rossi, "Identifying the Sources of Instabilities in Macroeconomic Fluctuations," Review of Economics and Statistics 93:4 (2011), 1186-1204.

Ireland, Peter N., "On the Welfare Cost of Inflation and the Recent Behavior of Money Demand," American Economic Review 99:3 (2009), 1040-1052.

Jermann, Urban, and Vincenzo Quadrini, "Macroeconomic Effects of Financial Shocks," American Economic Review 102:1 (2012), 238-271.

Kilian, Lutz, "Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market," American Economic Review 99:3 (2009), 1053-1069.

Kormilitsina, Anna, "Oil Price Shocks and the Optimality of Monetary Policy," Review of Economic Dynamics 14:1 (2011), 199-223.

Krishnamurthy, Arvind, and Annette Vissing-Jorgensen, "The Aggregate Demand for Treasury Debt," Journal of Political Economy 120:2 (2012), 233-267.

Kydland, Finn E., and Edward C. Prescott, "Time to Build and Aggregate Fluctuations," Econometrica 50:6 (1982), 1345-1370.

LaCour, Michael J., and Donald P. Green, "When Contact Changes Minds: An Experiment on Transmission of Support for Gay Equality," Science 346:6215 (2014), 1366-1369.

Last Week Tonight with John Oliver, Segment on Washington D.C. Voting Rights, airdate August 2, 2015.

Levitt, Steven, "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime," American Economic Review 87:3 (1997), 270-290.

Levitt, Steven, "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime: Reply," American Economic Review 92:4 (2002), 1244-1250.

Mavroeidis, Sophocles, "Monetary Policy Rules and Macroeconomic Stability: Some New Evidence," American Economic Review 100:1 (2010), 491-503.

McCrary, Justin, "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime: Comment," American Economic Review 92:4 (2002), 1236-1243.

McCullough, B.D., Kerry Anne McGeary, and Teresa D. Harrison, "Lessons from the JMCB Archive," Journal of Money, Credit and Banking 38:4 (2006), 1093-1107.

McCullough, B.D., Kerry Anne McGeary, and Teresa D. Harrison, "Do Economics Journal Archives Promote Replicable Research?" Canadian Journal of Economics 41:4 (2008), 1406-1420.

McCullough, B.D., and H.D. Vinod, "Verifying the Solution from a Nonlinear Solver: A Case Study," American Economic Review 93 (2003), 873-892.

McCullough, B.D., and H.D. Vinod, "Verifying the Solution from a Nonlinear Solver: Reply," American Economic Review 94:1 (2004), 391-396.

Mertens, Karel, and Morten O. Ravn, "Understanding the Aggregate Effects of Anticipated and Unanticipated Tax Policy Shocks," Review of Economic Dynamics 14:1 (2011), 27-54.

Mertens, Karel, and Morten O. Ravn, "The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States," American Economic Review 103:4 (2013), 1212-1247.

Morley, James, and Jeremy Piger, "The Asymmetric Business Cycle," Review of Economics and Statistics 94:1 (2012), 208-221.

Nakov, Anton, and Andrea Pescatori, "Oil and the Great Moderation," Economic Journal 120:543 (2010), 131-156.

Neumark, David, and William Wascher, “Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania: Comment,” American Economic Review 90:5 (2000), 1362-1396.

Open Science Collaboration, “Estimating the Reproducibility of Psychological Science,” Science 349:6251 (2015), aac4716.

Ramey, Valerie A., “Identifying Government Spending Shocks: It’s all in the Timing,” Quarterly Journal of Economics 126:1 (2011), 1-50.

Reinhart, Carmen M., and Kenneth S. Rogoff, “Growth in a Time of Debt,” American Economic Review 100:2 (2010), 573-578.

Reis, Ricardo, and Mark W. Watson, “Relative Goods’ Prices, Pure Inflation, and The Phillips Correlation,” American Economic Journal: Macroeconomics 2:3 (2010), 128-157.

Romer, Christina D., and David H. Romer, “The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks,” American Economic Review 100:3 (2010), 763-801.

Rothstein, Jesse, “Does Competition Among Public Schools Benefit Students and Taxpayers? A Comment on Hoxby (2000),” American Economic Review 97:5 (2007), 2026-2037.

Shachar, Ron, and Barry Nalebuff, “Follow the Leader: Theory and Evidence on Political Participation,” American Economic Review 83:3 (1999), 525-547.

Schmitt-Grohé, Stephanie, and Martín Uribe, “Business Cycles with a Common Trend in Neutral and Investment-Specific Productivity,” Review of Economic Dynamics 14:1 (2011), 122-135.

Schmitt-Grohé, Stephanie, and Martín Uribe, “What’s News in Business Cycles,” Econometrica 80:6 (2012), 2733-2764.

Senyuz, Zeynep, "Factor Analysis of Permanent and Transitory Dynamics of the US Economy and the Stock Market," Journal of Applied Econometrics 26:6 (2011), 975-998.

Sims, Christopher A., "Macroeconomics and Reality," Econometrica 48:1 (1980), 1-48.

Taylor, John B., "Discretion Versus Policy Rules in Practice," Carnegie-Rochester Conference Series on Public Policy 39 (1993), 195-214.

Thompson Reuters' Web of Science Journal Citation Reports: Social Science Edition. https://isiknowledge.com/ (2014).

Vlaeminck, Sven, and Lisa-Kristin Herrmann, "Data Policies and Data Archives: A New Paradigm for Academic Publishing in Economic Sciences?" New Avenues for Electronic Publishing in the Age of Infinite Collections and Citizen Science, B. Schmidt and M. Dobreva (Eds.) (2015). doi:10.3233/978-1-61499-562-3-145

Zimmermann, Christian, "On the Need for a Replication Journal," Federal Reserve Bank of St. Louis Working Paper 2015-016A (2015).

A Appendix to Chang and Li (2015): Replication Results and Sources of Replication Files

This appendix details what we interpret to be the key result of the papers in our sample, our individual replication results for papers that we were eventually able to replicate, and a brief summary of our replication procedure for each paper. We include the relevant figures and tables when our results are close to but not an exact replication of the published results. For a given replication, we occasionally supplemented the author-provided files with FAME code (version 10.2).

Auerbach and Gorodnichenko (2012, 2013)

We take the key results of Auerbach and Gorodnichenko (2012) as the top panel of Table 1, and Figures 2, 4, 5, and 7, which correspond to the key results from their abstract: "fiscal policy [is] considerably more effective in recessions than expansions," "military spending [has] the largest multiplier," and "controlling for predictable components of fiscal shocks tends to increase the size of the multipliers in recessions." We attempt to replicate the corrected versions of the top panel of Table 1 and Figure 2 in Auerbach and Gorodnichenko (2013) instead of the original erroneous versions using Matlab code from the American Economic Journal: Economic Policy's website. Our replication results are qualitatively similar to the published paper, but the results do not match exactly with Matlab R2013a (Windows). We were able to replicate the paper without assistance from the authors.

Table 5: Replication of Auerbach and Gorodnichenko (2012) Table 1, Top Panel

	Max Point Estimate	Standard Error	Cumulative Point Estimate	Standard Error
Total Spending				
Linear	0.89	0.29	0.60	0.23
Expansion	0.49	0.13	-0.80	0.16
Recession	2.12	0.18	2.17	0.19
Defense Spending				
Linear	1.53	0.56	0.39	0.22
Expansion	0.76	0.21	-0.94	0.26
Recession	4.27	0.93	2.18	0.78
Nondefense Spending				
Linear	1.69	0.08	2.09	0.15
Expansion	1.20	0.16	1.16	0.15
Recession	1.06	0.30	1.10	0.32
Consumption Spending				
Linear	0.83	0.28	0.90	0.29
Expansion	0.10	0.12	-0.16	0.12
Recession	2.16	0.65	1.33	0.36
Investment Spending				
Linear	2.06	0.60	2.75	0.60
Expansion	2.86	0.27	2.03	0.17
Recession	2.79	0.53	4.18	0.46

Corrected results from Auerbach and Gorodnichenko (2013). Table shows output multipliers for a $1 increase in government spending.

Figure 1: Replication of Auerbach and Gorodnichenko (2012) Figure 2

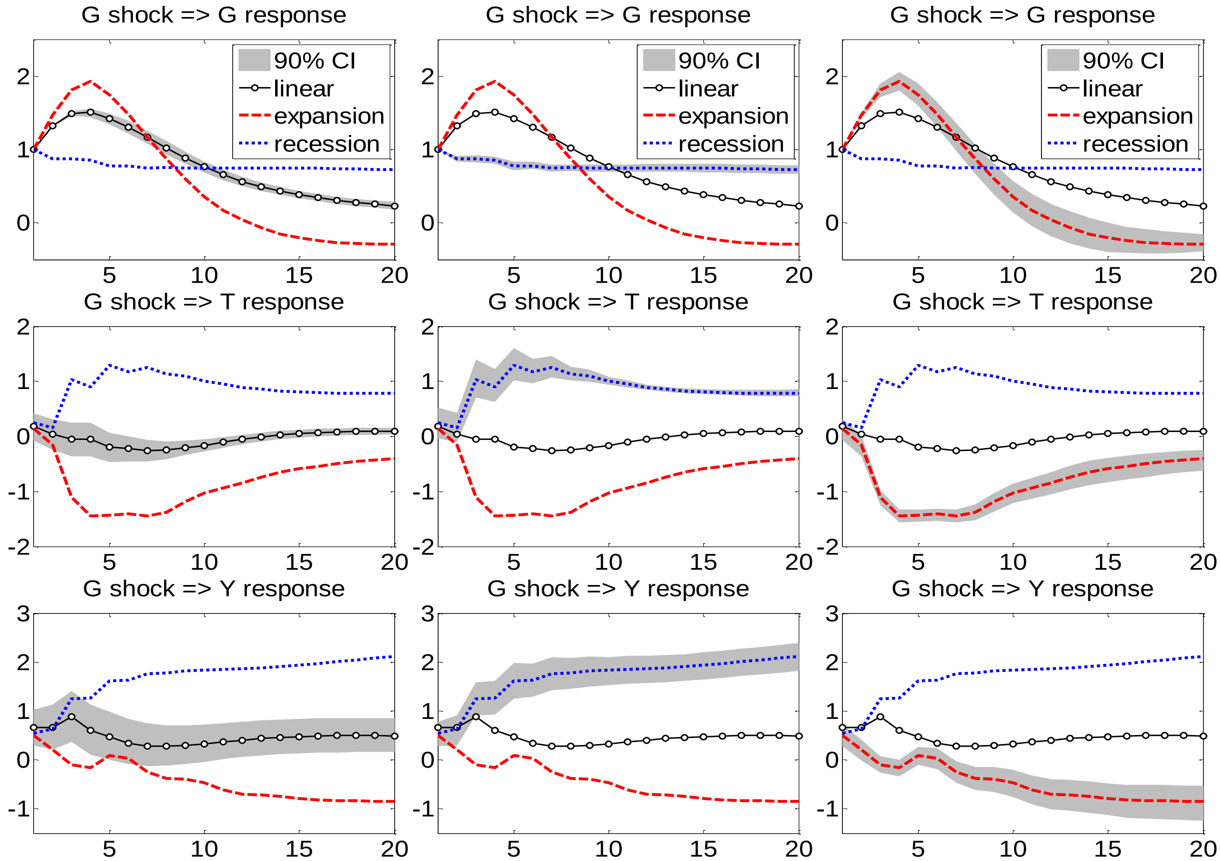


Figure 2: Replication of Auerbach and Gorodnichenko (2012) Figure 4

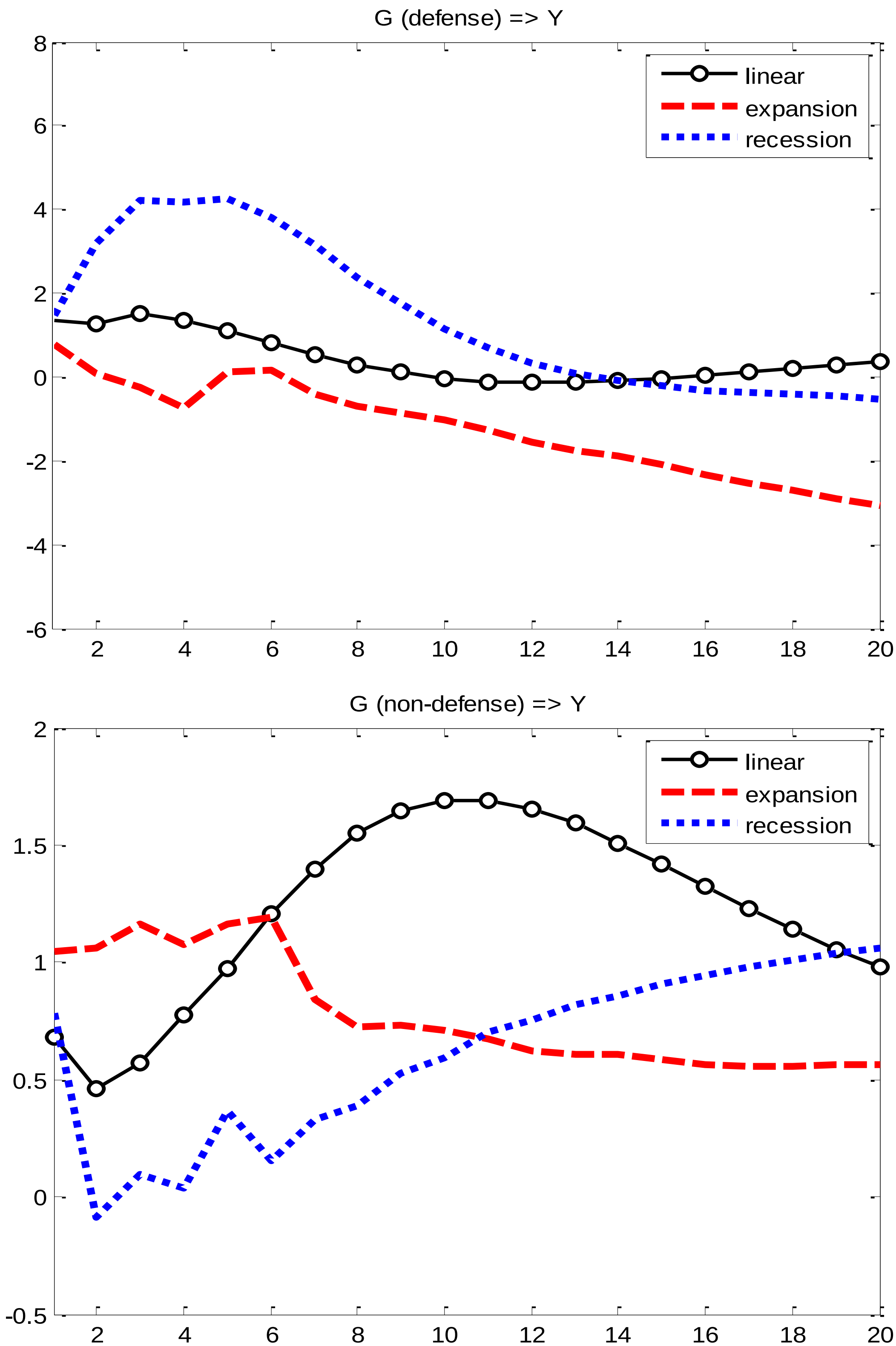


Figure 3: Replication of Auerbach and Gorodnichenko (2012) Figure 5

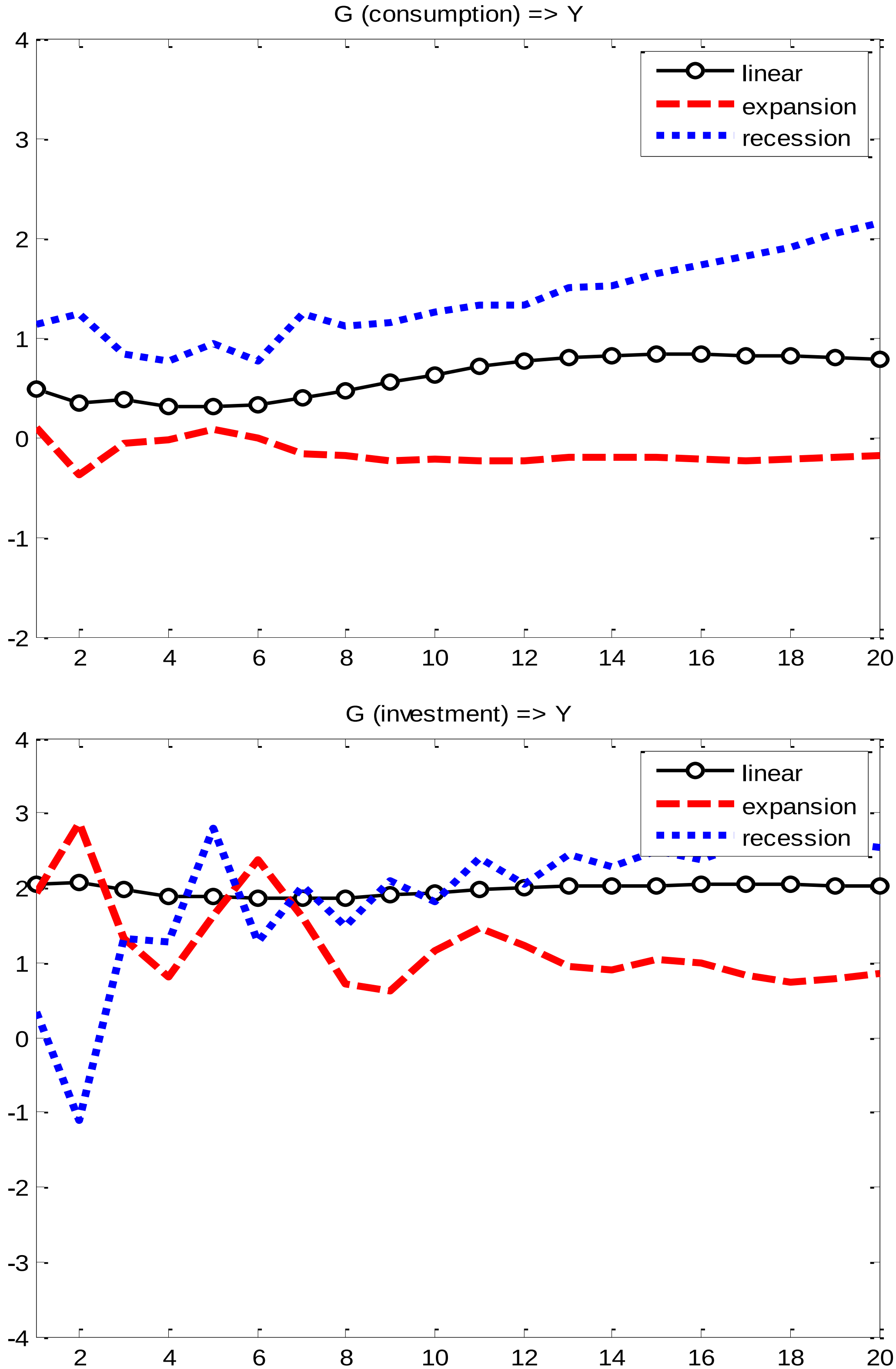
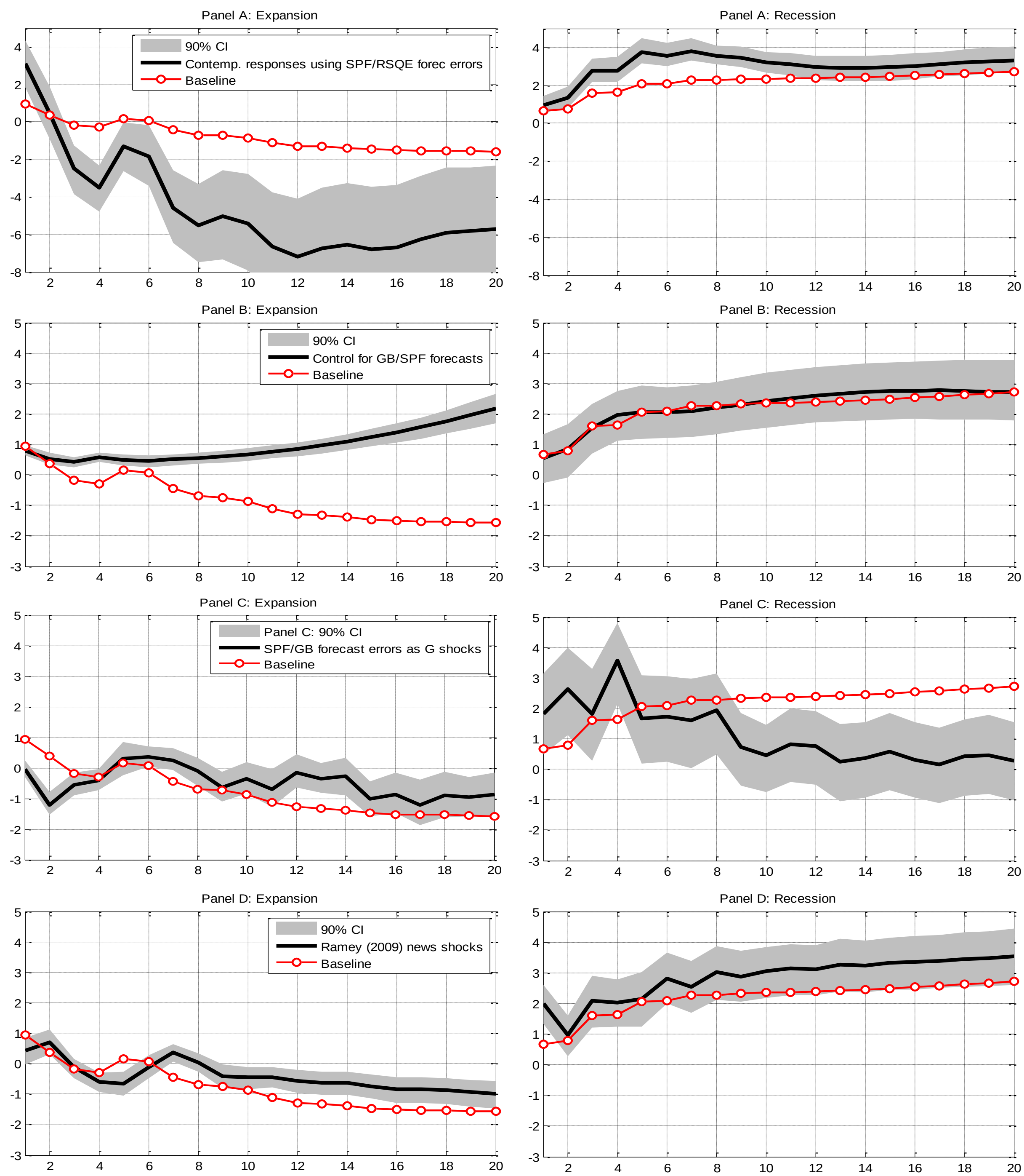


Figure 4: Replication of Auerbach and Gorodnichenko (2012) Figure 7



Barro and Redlick (2011)

We take the key results of Barro and Redlick (2011) as Tables 2, 3, 5, and 7, namely that "all estimated [spending] multipliers are significantly less than one" (abstract). Robert Barro provided us with an EViews file to replicate the results of Barro and Redlick (2011), although we ran the regressions based on the EViews template in Stata 11.0 (Windows). We are able to replicate their point estimates exactly, except for column (8) of their Table 7, where we find slightly different estimates. Our standard errors are a bit smaller than the published estimates across all specifications.

Table 6: Replication of Barro and Redlick (2011) Table 2

	Starting date					
	(1)	(2)	(3)	(4)	(5)	(6)
	1950	1939	1930	1930 (w/o 1949)	1917	1954
$\Delta g : defense$	0.68	0.44	0.46	0.48	0.49	0.98
	(0.26)**	(0.06)**	(0.08)**	(0.08)**	(0.08)**	(0.60)
$\Delta g : defense$	0.01	0.20	0.21	0.25	0.15	-0.54
(-1)	(0.26)	(0.06)**	(0.08)*	(0.08)**	(0.08)	(0.52)
$\Delta g* : defense$	0.026	0.039	0.034	0.034	0.024	-0.12
$news$	(0.015)	(0.010)**	(0.014)*	(0.014)*	(0.015)	(0.10)
$U(-1)$	0.50	0.58	0.61	0.58	0.47	0.51
	(0.16)**	(0.14)**	(0.10)**	(0.09)**	(0.10)	(0.17)**
$\Delta\tau(-1)$	-0.54	-0.16	-0.26	-0.52	-0.18	-0.48
	(0.20)**	(0.15)	(0.21)	(0.22)*	(0.24)	(0.21)*
Yield Spread	-43.9	-37.8	-101.5	-103.4	-73.9	-0.43
Squared	(19.4)*	(20.9)	(12.2)**	(11.8)**	(11.9)**	(20.3)*
p-value, defense variables	0.02	0.00	0.00	0.00	0.00	0.41

Table 7: Replication of Barro and Redlick (2011) Table 3

	Starting date					
	(1)	(2)	(3)	(4)	(5)	(6)
	1950	1930	1950	1930	1950	1950
$\Delta g : defense$	0.89	0.46	0.34	0.51	0.84	0.46
	(0.25)**	(0.08)**	(0.30)	(0.09)**	(0.22)**	(0.24)
$\Delta g : defense$	-0.13	0.21	0.08	0.18	-0.36	0.02
(-1)	(0.25)	(0.08)*	(0.26)	(0.08)*	(0.24)	(0.24)
$\Delta g* : defense$	0.040	0.036	0.028	0.032	0.014	0.016
$news$	(0.015)**	(0.015)*	(0.014)	(0.014)*	(0.012)	(0.013)
$U(-1)$	0.65	0.60	0.43	0.62	0.26	0.55
	(0.16)**	(0.10)**	(0.17)**	(0.10)**	(0.15)	(0.15)**
$\Delta\tau(-1)$	-0.45	-0.25	-0.56	-0.25	-0.26	-0.38
	(0.18)*	(0.22)	(0.19)**	(0.21)	(0.18)	(0.18)*
Yield Spread	-31.2	-100.9	-28.2	-102.4	-38.9	-21.6
Squared	(18.6)	(12.6)**	(23.7)	(12.3)**	(16.8)*	(19.0)
$\Delta g : nondefense$	2.65	0.12				
	(0.86)**	(0.59)				
$\Delta(transfers)$			-1.53	0.64		
			(0.86)	(0.64)		
$\Delta(GMsales)$					3.66	
					(0.80)**	
$\Delta(GEsales)$						17.6
						(4.39)**

Table 8: Replication of Barro and Redlick (2011) Table 5

	Dependent Variable				
	$\Delta(c:nondur)$	$\Delta(c:dur)$	$\Delta(invest)$	$\Delta(g:nondef)$	$\Delta(x-m)$
Sample: 1950-2006					
$\Delta g:defense$	0.005	-0.17	-0.083	-0.081	0.004
	(0.09)	(0.069)*	(0.173)	(0.039)*	(0.074)
$\Delta g:defense$	0.18	0.15	-0.142	0.055	-0.23
(-1)	(0.09)*	(0.070)*	(0.177)	(0.040)	(0.075)
$\Delta g*:defense$	-0.004	0.011	0.038	-0.005	-0.013
$news$	(0.005)	(0.004)**	(0.010)**	(0.002)*	(0.004)**
$U(-1)$	0.11	0.15	0.38	-0.053	-0.095
	(0.05)*	(0.04)**	(0.011)**	(0.024)*	(0.046)*
$\Delta\tau(-1)$	-0.18	-0.15	-0.30	-0.033	0.12
	(0.07)**	(0.05)**	(0.133)*	(0.030)	(0.056)*
Yield Spread	-5.39	-3.49	-22.7	-4.80	-6.71
Squared	(6.57)	(5.16)	(13.1)	(2.92)	(5.54)
Sample: 1939-2006					
$\Delta g:defense$	-0.011	-0.115	-0.356	-0.009	-0.071
	(0.02)	(0.015)**	(0.043)**	(0.011)	(0.020)**
$\Delta g:defense$	0.11	0.038	0.096	-0.011	-0.027
(-1)	(0.02)**	(0.015)*	(0.044)*	(0.011)	(0.020)
$\Delta g*:defense$	0.004	0.012	0.034	-0.008	-0.002
$news$	(0.004)	(0.003)**	(0.008)**	(0.002)**	(0.004)
$U(-1)$	0.101	0.094	0.401	-0.030	-0.002
	(0.049)*	(0.036)**	(0.104)**	(0.025)	(0.048)
$\Delta\tau(-1)$	-0.008	-0.103	-0.067	-0.105	0.114
	(0.056)	(0.041)*	(0.112)	(0.028)**	(0.055)*
Yield Spread	1.13	-3.09	-20.3	-6.50	-8.04
Squared	(7.54)	(5.57)	(15.9)	(3.91)	(7.40)
Sample: 1930-2006					
$\Delta g:defense$	-0.001	-0.110	-0.34	-0.016	-0.074
	(0.036)	(0.017)**	(0.049)**	(0.016)	(0.019)**
$\Delta g:defense$	0.11	0.036	0.087	-0.003	-0.024
(-1)	(0.037)**	(0.017)*	(0.051)	(0.016)	(0.020)
$\Delta g*:defense$	-0.0004	0.011	0.035	-0.010	-0.002
$news$	(0.0065)	(0.003)**	(0.009)**	(0.003)**	(0.003)
$U(-1)$	0.17	0.081	0.30	0.041	0.006
	(0.045)**	(0.020)**	(0.060)**	(0.019)*	(0.023)
$\Delta\tau(-1)$	-0.060	-0.112	-0.010	-0.111	0.113
	(0.097)	(0.044)*	(0.130)	(0.042)**	(0.051)*
Yield Spread	-42.3	-12.9	-39.9	-4.91	-1.12
Squared	(5.62)	(2.58)**	(7.56)**	(2.42)*	(2.95)

Table 9: Replication of Barro and Redlick (2011), Table 7, Columns 1-4

	(1)	(2)	(3)	(4)
$\Delta g : defense$	0.67	0.53	0.66	0.49
	(0.26)**	(0.26)*	(0.26)**	(0.27)
$\Delta g : defense$	0.007	-0.23	-0.05	0.13
(-1)	(0.26)	(0.26)	(0.27)	(0.27)
$\Delta g* : defense$	0.025	0.03	0.027	0.018
$news$	(0.015)	(0.015)	(0.015)	(0.015)
$U(-1)$	0.52	0.51	0.48	0.49
	(0.16)**	(0.17)**	(0.16)	(0.16)
$\Delta\tau(-1)$	-0.53		0.43	-0.67
	(0.20)		(0.22)*	(0.21)**
$\Delta\tau$				0.38
				(0.22)
Romers: exogenous		-1.08	-0.56	
$[\Delta tax/Y(-1)](-1)$		(0.53)	(0.58)	
Yield Spread	-47.2	-43.4	-41.8	-40.1
Squared	(18.9)*	(20.3)*	(19.7)	(19.4)*
p-value: τ	0.01		0.049	0.006
p-value: Romers		0.042	0.331	
p-value: all tax vars	0.01	0.042	0.016	0.006

Table 10: Replication of Barro and Redlick (2011), Table 7, Columns 5-8

	(1)	(2)	(3)	(4)
$\Delta g : defense$	0.53	0.71	0.72	0.17
	(0.26)*	(0.28)*	(0.27)**	(0.20)
$\Delta g : defense$	-0.23	-0.21	-0.03	0.25
(-1)	(0.26)	(0.26)	(0.26)	(0.20)
$\Delta g* : defense$	0.029	0.016	0.02	-0.004
$news$	(0.015)*	(0.016)	(0.015)	(0.11)
$U(-1)$	0.51	0.49	0.49	0.30
	(0.17)**	(0.17)	(0.16)**	(0.13)*
$\Delta\tau(-1)$			-0.45	-0.49
			(0.23)*	(0.14)**
Romers: exogenous	-1.08			
$[\Delta tax/Y(-1)](-1)$	(0.54)			
Romers: exogenous	-0.03			
$[\Delta tax/Y(-1)]$	(0.51)			
$[\Delta(fedrev.)/Y(-1)](-1)$		-0.46	-0.17	
		(0.25)	(0.28)	
$[\Delta(fedrev.)/Y(-1)]$				1.28
				(0.22)**
Yield Spread	-42.9	-64.9	-52.5	-27.6
Squared	(20.3)*	(19.4)**	(19.7)**	(16.0)
p-value: τ				0.001
p-value: Romers	0.124		0.046	
p-value: fed. revenue		0.066	0.53	0.000
p-value: all tax vars	0.124	0.066	0.021	0.000

Baumeister and Peersman (2013)

We use the code from the American Economic Journal: Macroeconomics website to replicate Baumeister and Peersman (2013). This paper requires a local version of Matlab, but we were unable to verify the version of Matlab used by the authors. We conduct the estimation in Matlab R2013a (Windows). We take the key results as their Figures 1 and 4, which correspond to a "decline in the shortrun price elasticity of oil demand since the mid-1980s" (abstract), implying that oil production shocks have a greater effect on GDP than oil price shocks. With assistance from Christiane Baumeister, we were able to replicate their Figure 1 and their Figure 4 with some minor differences, shown in Figures 5 and 6.

Figure 5: Replication of Baumeister and Peersman (2013) Figure 1

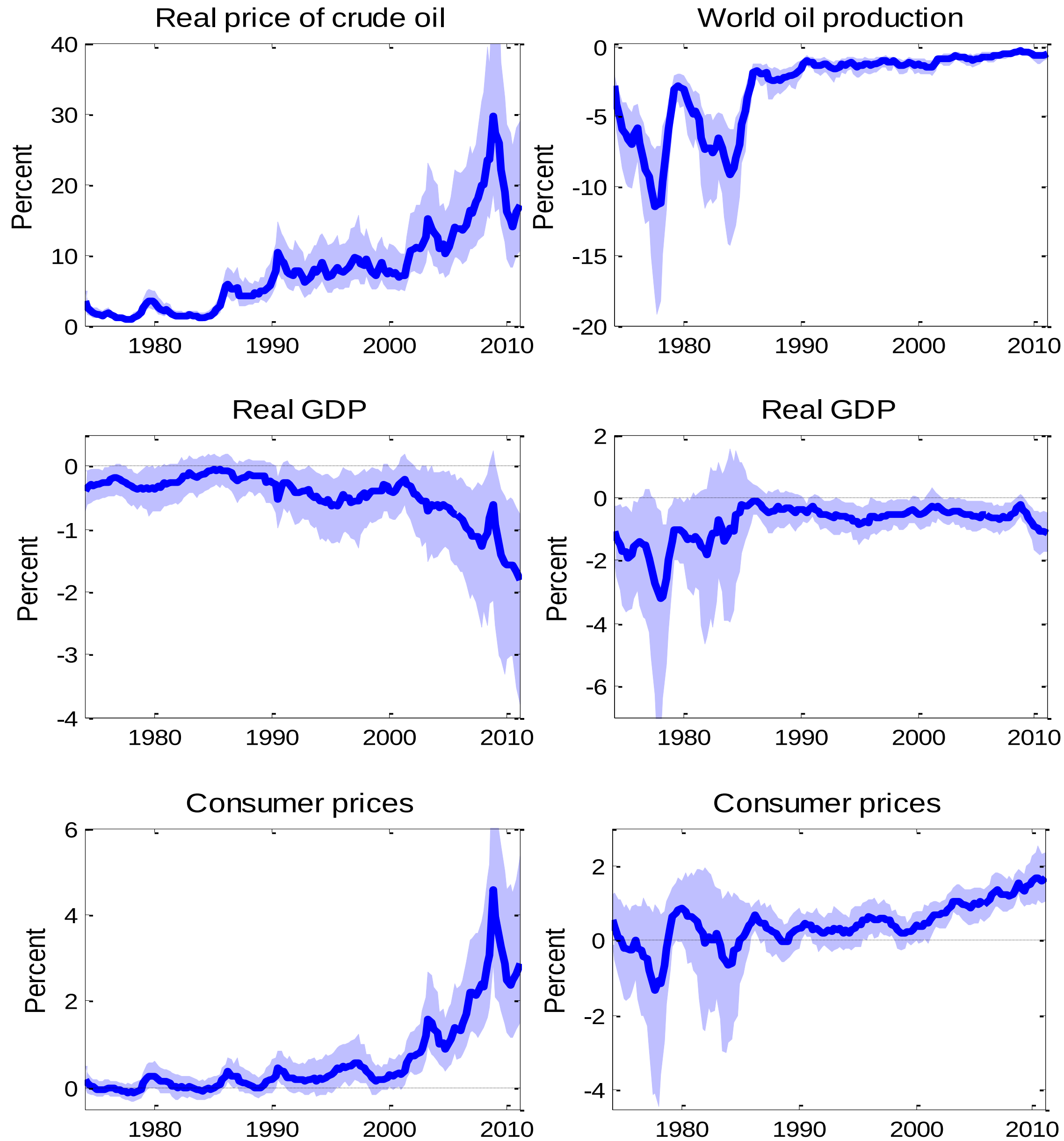
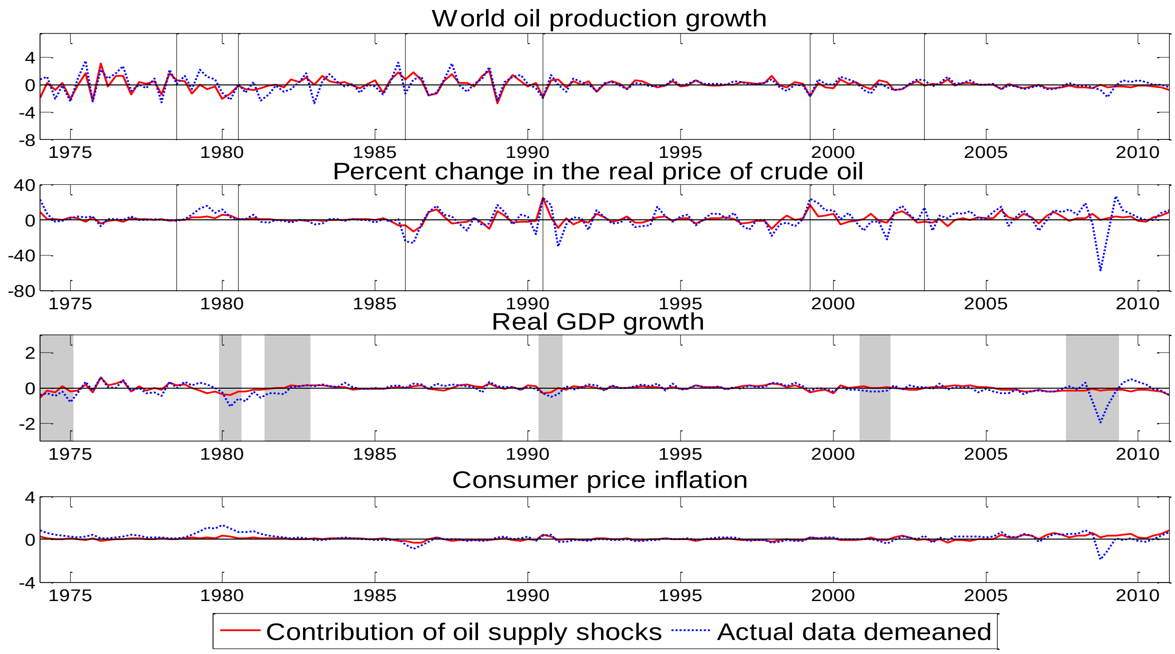


Figure 6: Replication of Baumeister and Peersman (2013) Figure 4



Canova and Gambetti (2010)

We believe the key results are Tables 5 and 6. The authors outline their key empirical results on page 184 of their article in the last 2 full paragraphs of the introduction. We were able to replicate both tables, but with a few discrepancies and while supplementing the files from the AEJ: Macro's website with our own RATS code using RATS 7.10 (Linux). We alter the replication code in the following four ways: (1) change the time period of the estimation to match the reported 186 observations in the paper, (2) change the VAR lags to 4 lags to match the description on page 195 of the paper, (3) write code to calculate variances in Table 6, and (4) add code to create tabular output. Tables 11 to 12 show our replication results.

Table 11: Replication of Canova and Gambetti (2010) Table 5

	Sample							
	60Q1-79Q2	60Q1-80Q2	60Q1-81Q2	60Q1-82Q2	79Q3-05Q4	80Q3-05Q4	81Q3-05Q4	82Q3-05Q4
Panel A. With Michigan Expectations								
ΔGDP	0.74	0.67	0.67	0.81	0.45	0.48	0.74	0.82
π	0.00	0.01	0.00	0.00	0.01	0.00	0.27	0.05
R	0.12	0.04	0.15	0.27	0.00	0.00	0.00	0.16
Panel B. With Term Structure Expectations								
ΔGDP	0.76	0.89	0.60	0.46	0.03	0.04	0.11	0.67
π	0.58	0.53	0.11	0.01	0.00	0.00	0.59	0.25
R	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.03

Table reports p-values of an F-test that the coefficients on the expectation variable are equal to zero in a 4-lag vector autoregression that includes the growth rate of GDP (ΔGDP), inflation (π), and the nominal interest rate (R) for the subsamples listed.

Table 12: Replication of Canova and Gambetti (2010) Table 6

	Sample							
	60Q1-79Q2	60Q1-80Q2	60Q1-81Q2	60Q1-82Q2	79Q3-05Q4	80Q3-05Q4	81Q3-05Q4	82Q3-05Q4
Panel A. With Michigan Expectations								
ΔGDP	0.87	0.87	0.89	1.12	0.62	0.60	0.60	0.39
π	0.08	0.09	0.10	0.10	0.05	0.05	0.04	0.04
R	0.51	0.76	1.50	1.98	0.96	0.95	0.57	0.19
Panel B. With Term Structure Expectations								
ΔGDP	0.87	0.89	0.89	1.09	0.56	0.54	0.52	0.35
π	0.10	0.10	0.11	0.11	0.05	0.05	0.04	0.04
R	0.44	0.53	1.03	1.36	0.65	0.65	0.47	0.15
Panel C. Without Inflation Expectations								
ΔGDP	0.90	0.90	0.93	1.14	0.63	0.61	0.57	0.36
π	0.11	0.11	0.12	0.13	0.06	0.06	0.04	0.04
R	0.58	0.90	1.66	2.13	1.17	1.08	0.51	0.18

Table reports the variances of reduced-form shocks in a 4-lag vector autoregression that includes the growth rate of GDP (ΔGDP), inflation (π), the nominal interest rate (R), and an expectations variable in panels A and B for the subsamples listed.

Carey and Shore (2013)

We take the key result as Table 4, namely that cross-sectional variance of income is higher in expansions (phrased in their abstract as "income volatility is higher in good state times than in bad”). Using the data and code files from ReStat’s website, we replicate their Table 4 almost exactly without assistance using Stata 13.0 (Windows).

Table 13: Replication of Carey and Shore (2013) Table 4

		Controls			
		None	Education	Education and % Black	Education, % Black, and Income
βx(NBER recession?)	Cumulative	-0.0292	-0.0270	-0.0283	-0.0270
		(0.0045)	(0.0045)	(0.0046)	(0.0046)
	Recent	-0.0036	-0.0014	-0.0005	-0.0021
		(0.0043)	(0.0042)	(0.0042)	(0.0043)
βx(Negative National Growth?)	Cumulative	-0.0189	-0.0158	-0.0161	-0.0160
		(0.0038)	(0.0038)	(0.0039)	(0.0040)
	Recent	-0.0124	-0.0103	-0.0114	-0.0113
		(0.0045)	(0.0044)	(0.0046)	(0.0046)
βx(% of Year in Recession)	Cumulative	-0.0365	-0.0321	-0.0350	-0.0330
		(0.0050)	(0.0050)	(0.0051)	(0.0051)
	Recent	-0.0082	-0.0057	-0.0032	-0.0028
		(0.0078)	(0.0077)	(0.0077)	(0.0080)
βx(National GDP growth)	Cumulative	-0.1144	-0.1439	-0.1946	-0.1603
		(0.0818)	(0.0813)	(0.0831)	(0.0862)
	Recent	0.0350	0.0189	-0.0428	-0.0244
		(0.0802)	(0.0800)	(0.0814)	(0.0829)
βx(Demeaned squared growth)	Cumulative	-11.9832	-10.7285	-11.0711	-10.5387
		(1.5016)	(1.4999)	(1.5210)	(1.5501)
	Recent	-15.4986	-14.3653	-15.0679	-14.3661
		(2.0533)	(2.0451)	(2.0694)	(2.1023)

Each value is an OLS regression with standard errors in parentheses, with the dependent variable as cross-sectional income variance. See text of Carey and Shore (2013) for additional details.

Chen, Curdia, and Ferrero (2012)

We take the key results as Figures 2 to 5. This paper was not subject to a data and code replication policy, but Vasco Curdia provided a working replication data and code files that we were able to use to replicate these figures exactly with Matlab R2010a (Linux).

Clark and McCracken (2010)

We take the key results as the panels for GDP growth in their Tables 3 to 5. In their Table 3, the relative root mean squared errors (RMSEs) of the forecasting models that the authors consider worsen relative to the forecasts from the univariate benchmarks going from the 1970-1984 sample to the 1985-2005 sample. The optimal GDP forecast, when gauged by RMSEs, comes from their Bayesian vector autoregressions or first-differenced vector autoregressions. Similar results hold in their Tables 4 and 5.

We are able to replicate these results almost exactly, with the exception of the rows in these tables with Bayesian model averaging, which are approximately the same as in the published paper. We display our replication results where we do not find an exact match in Tables 14 to 16. We use RATS 7.10 (Linux) to replicate their results.

Table 14: Replication of Clark and McCracken (2010) Table 3

	Sample Period					
Forecast Method	1970-1984			1985-2005		
	$h = 0Q$	$h = 1Q$	$h = 1Y$	$h = 0Q$	$h = 1Q$	$h = 1Y$
BMA: AIC	1.007	0.959	0.884	1.111	1.124	1.095
BMA: BIC	0.946	0.909	0.964	1.047	1.039	0.899
BMA: PIC	0.902	0.838	0.852	1.107	1.112	1.005

Rows displayed are those from the GDP forecast panel of Table 3 of Clark and McCracken (2010) where we could not match the published results exactly.

Table 15: Replication of Clark and McCracken (2010) Table 4

Forecast Method	Sample Period					
	1970-1984			1985-2005		
	$h = 0Q$	$h = 1Q$	$h = 1Y$	$h = 0Q$	$h = 1Q$	$h = 1Y$
BMA: AIC	0.921	0.875	0.909	1.088	1.124	1.126
BMA: BIC	0.959	0.867	0.887	1.055	1.078	1.025
BMA: PIC	0.878	0.807	0.826	1.072	1.101	1.089

Rows displayed are those from the GDP forecast panel of Table 4 of Clark and McCracken (2010) where we could not match the published results exactly.

Table 16: Replication of Clark and McCracken (2010) Table 5

Forecast Method	Sample Period					
	1970-1984			1985-2005		
	$h = 0Q$	$h = 1Q$	$h = 1Y$	$h = 0Q$	$h = 1Q$	$h = 1Y$
BMA: AIC	1.053	1.096	1.245	1.402	1.277	1.117
BMA: BIC	0.958	1.003	1.038	1.340	1.235	1.069
BMA: PIC	0.979	0.977	1.076	1.130	1.078	1.007

Rows displayed are those from Table 5 of Clark and McCracken (2010) where we could not match the published results exactly.

Corsetti, Meier, and Müller (2012)

We take the key result to be the panels for output and debt in their Figures 1 and 2, as debt is scaled by GDP. According to the paper's introduction, an "increase in government spending causes a substantial rise in aggregate output... a positive spending shock triggers a sizable buildup of public debt, followed over time by a decline of government spending below trend." We were able to replicate these panels exactly with assistance from the authors using code from ReStat's website in Matlab R2008a (Windows).

D'Agostino and Surico (2012)

We take the key results as the bottom panel of Figure 1, the right-hand side panels of Figure 2, and all of Figure 4. We downloaded the replication files from ReStat's website and were able to replicate these figures exactly without assistance using Matlab R2010a (Linux).

Den Haan and Sterk (2011)

This paper was not subject to a data and code availability policy. However, we were able to download the data from Wouter Den Haan's personal site on December 1, 2013 (www.wouterdenhaan.com/data.htm), and Vincent Sterk provided us with the code needed to replicate their results. We take the key figures as Figures 3, 4, 6, and 7. Our replication results are close to the published version using Matlab R2010a (Linux).

Figure 7: Replication of Den Haan and Sterk (2011) Figure 3

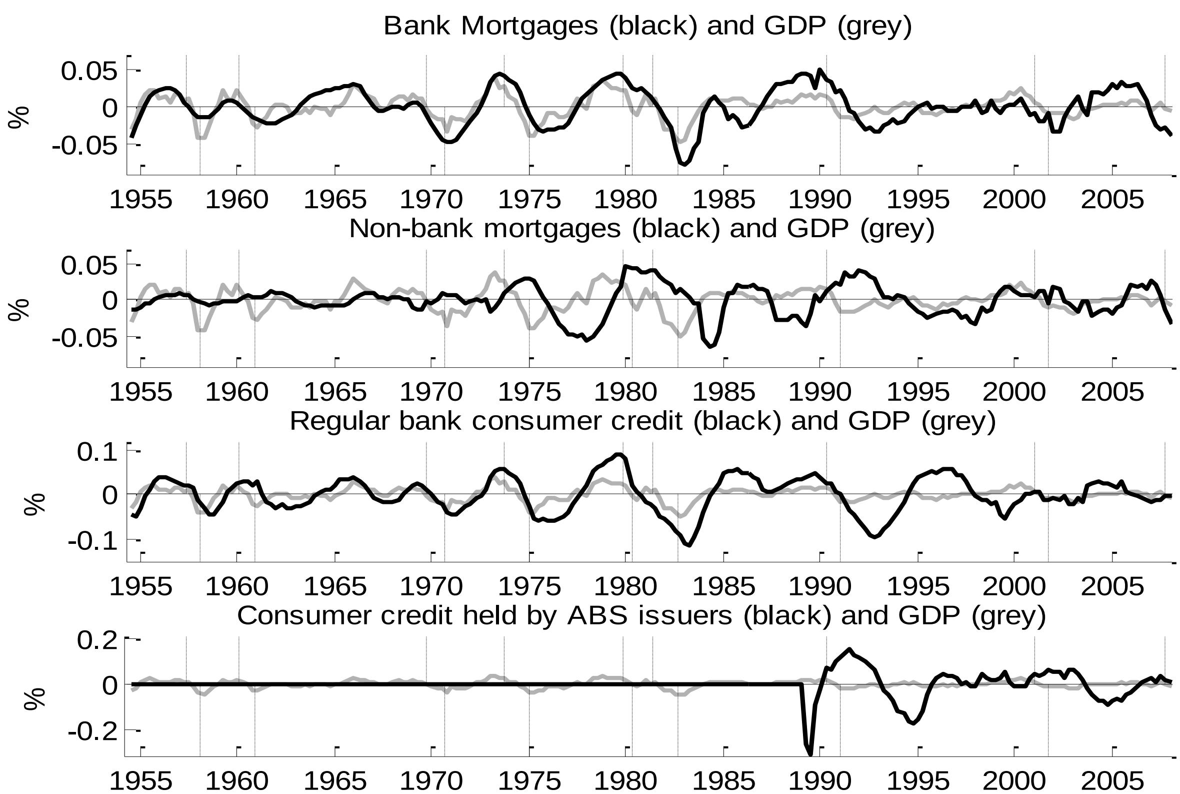


Figure 8: Replication of Den Haan and Sterk (2011) Figure 4

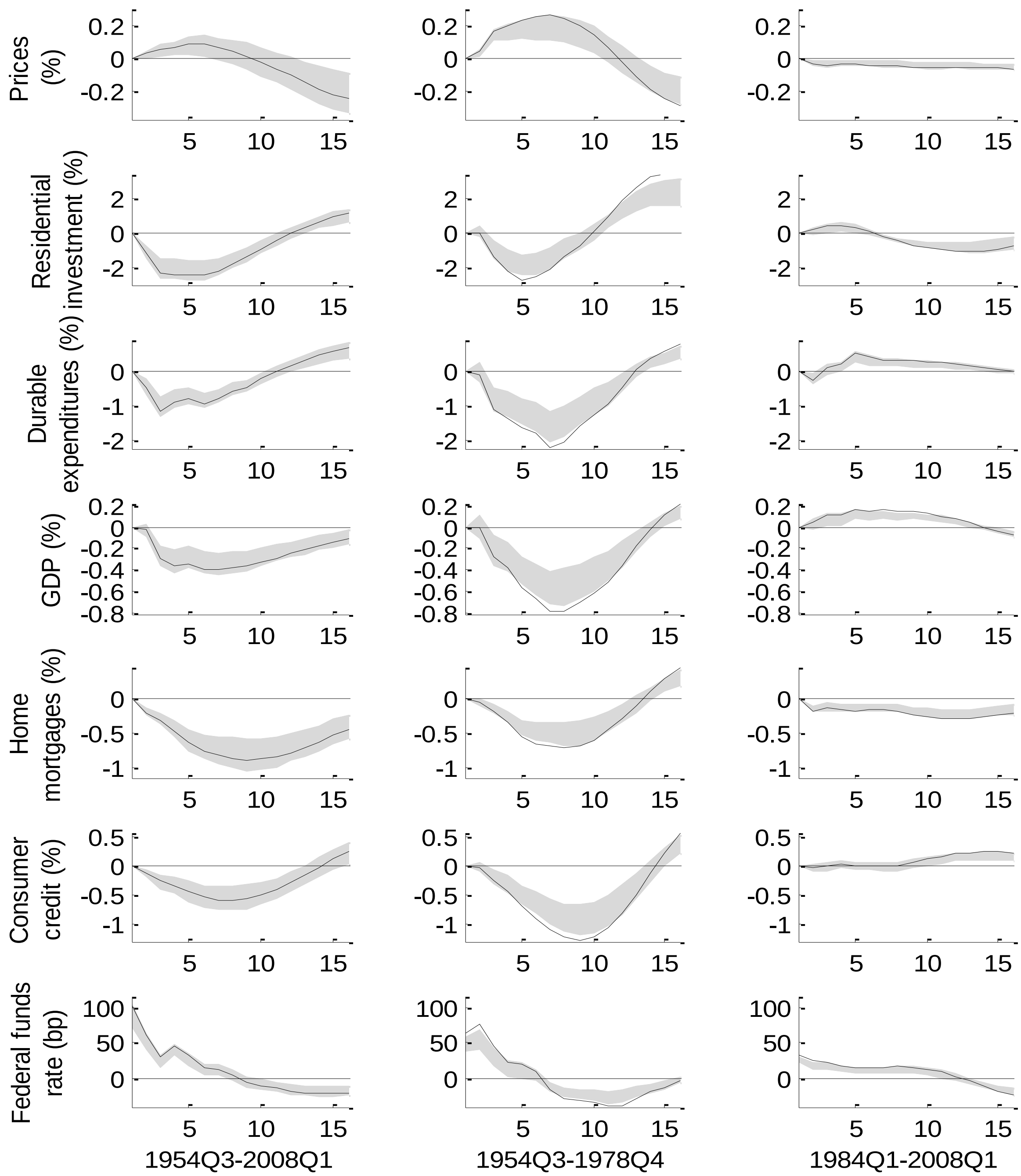


Figure 9: Replication of Den Haan and Sterk (2011) Figure 6

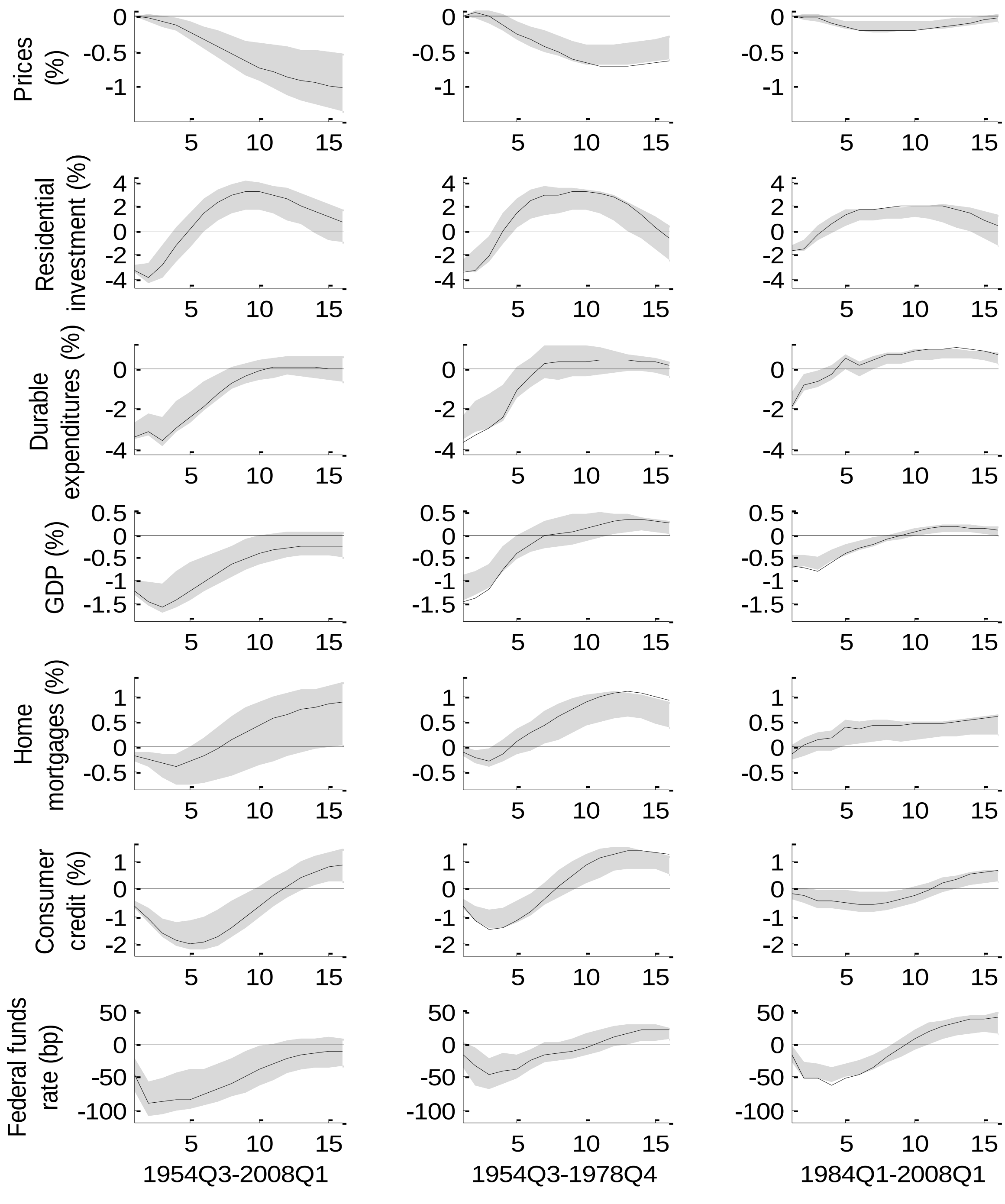
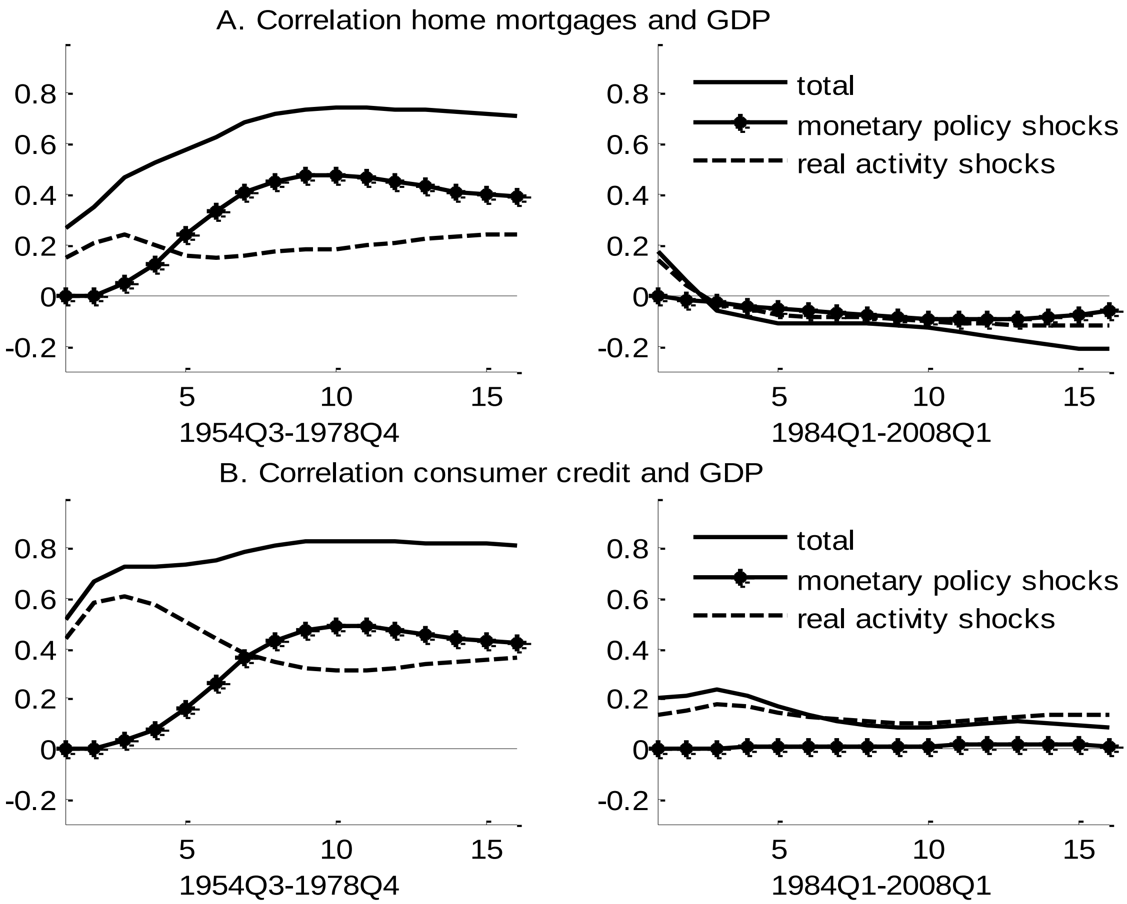


Figure 10: Replication of Den Haan and Sterk (2011) Figure 7



Favero and Giavazzi (2012)

We interpret the key results as Figures 5 to 6, with Figures 3 to4 as necessary conditions for the results in Figures 5 to 6. We were able to replicate Figures 5 and 6 exactly without help from the authors using EViews 8 (Windows).

Gabaix (2011)

We take the key results as Figure 2 and Tables 1 to 7, which we are able to replicate exactly using the data and code from Econometrica's website without assistance from the author using R 3.1.0 (Linux).

Hansen, Lunde, and Nason (2011)

We take the key result as Table 7, where the Hansen, Lunde, and Nason (2011) model confidence set procedure selects different variants of the Taylor (1993) rule. We are able to replicate this table exactly with code from Econometrica's website without help from the authors using OX 6.30 (Windows).

Heutel (2012)

We take the key result as Table 1, which demonstrates GDP elasticities of between 0.5 and 0.9. Using the replication files from the RED's website, we were able to replicate the table exactly without the author's assistance using Stata 13.0 (Linux).

Inoue and Rossi (2011)

We take the key results as Tables 1 and 3, with their Table 2 being additional motivation for their analysis of time-varying structural parameters. From their Table 1, Inoue and Rossi (2011) conclude that most of the parameters in their New Keynesian model are unstable (hypothesis tests generally reject the null of stable parameters), using both the Andrews (1993) QLR stability test and the Inoue and Rossi (2011) estimate of the set of stable parameters (ESS) procedure. Their Table 3 shows that the contribution to the standard deviations of inflation and output from the model parameters are often of opposite signs, so the net effect on the standard deviation of output and inflation is mitigated. Using the replication files from ReStat's website, we were able to qualitatively match the results from their Tables 1 and 3, but there are some minor differences using Matlab R2010a (Linux).

Table 17: Replication of Inoue and Rossi (2011) Table 1

Model Parameters	Individual p-Value	ESS p-Value
ρ_e	0	0
σ_ν	0	0
α	0	0
σ_a	0	0
σ_π	0	0
ρ_a	0	0
γ	0	0
ψ	0	0.01
ρ_{gy}	0	0
σ_e	0	0
ρ_υ	0	0
ρ_π	0	0
σ_z	1	1

Set of stable parameters (90% probability level): S = $\{\sigma_z\}$. This table reports p-values of the QLR stability test (Andrews, 1993) on individual parameters, labeled "Individual p-value," and the p-values of each step of the Inoue and Rossi (2011) ESS procedure, labeled "ESS p-value."

Table 18: Replication of Inoue and Rossi (2011) Table 3

Parameter:	Output	Inflation	Interest Rate
No change: (actual S.D.)	0.89	0.48	0.30
Unstable Parameters	% Contribution to Change		
ρ_e	7%	10%	-1%
σ_ν	71%	35%	40%
α	-2%	12%	1%
σ_a	-22%	-4%	-104%
σ_π	4%	15%	35%
ρ_a	25%	2%	94%
γ	20%	0%	18%
ψ	0%	0%	0%
ρ_{gy}	-43%	1%	24%
σ_e	-2%	-5%	-1%
ρ_υ	6%	5%	-15%
ρ_π	-13%	-23%	5%
Stable Parameters:			
σ_z	49%	53%	3%
All change: (actual S.D.)	1.45	0.92	0.39

Set of stable parameters (90% probability level): S = $\{\sigma_z\}$. This table shows the percentage contribution to the increase or decrease in the volatilities of output, inflation, and the interest rate by progressively allowing each parameter to be time varying, ordered according to the p-values of the QLR stability test (Andrews, 1993).

Ireland (2009)

We take the key table as Table 2, which we were able to replicate using code and data from the AER's website without help from Peter Ireland using Matlab R2010a (Linux).

Kilian (2009)

We take the key figure as Figure 5, namely the responses of GDP to oil supply shocks. We were able to replicate this figure using data and code from the AER's website without assistance from the author in Matlab R2013a (Windows).

Kormilitsina (2011)

We interpret the key result of this paper as Figure 2 of the paper, which contrasts optimal policy derived from Kormilitsina's model to the actual policy. We were able to replicate this figure exactly with assistance from the author in adjusting the replication code provided from RED's website. We used Fortran f90 (Linux) and Matlab R2012a (Windows).

Krishnamurthy and Vissing-Jorgensen (2012)

We take the key results of this paper to be Tables 1 to 2 of the paper, specifically the coefficient on debt to GDP in both tables. The EDF variable in Krishnamurthy and Vissing-Jorgensen (2012) is proprietary, so we replicate the columns in Tables 1 to2 that do not use the EDF variable. We were able to replicate the tables exactly without complication and independent of the authors with the files from the JPE in Matlab R2013a (Windows).

Mavroeidis (2010)

Key results are Figures 1 and 2. We were able to replicate these figures exactly without assistance from the author using data and code from the AER's website in OX 6.30 (Linux).

Mertens and Ravn (2011)

We take the key results as Figures 1, 3, and 4. We were able to replicate these figures exactly without assistance from the authors using data and code from the RED's website using Matlab R2013a (Windows).

Mertens and Ravn (2013)

We interpret the key result to be from page 1228, "perhaps the most important result in this paper is that the estimated short run output effects of changes in average tax rates are large," which corresponds to Figures 2 and 3. We were able to use data and code from the AER's website and replicate these figures closely without help from the authors using Matlab R2013a (Windows).

Figure 11: Replication of Mertens and Ravn (2013) Figure 2

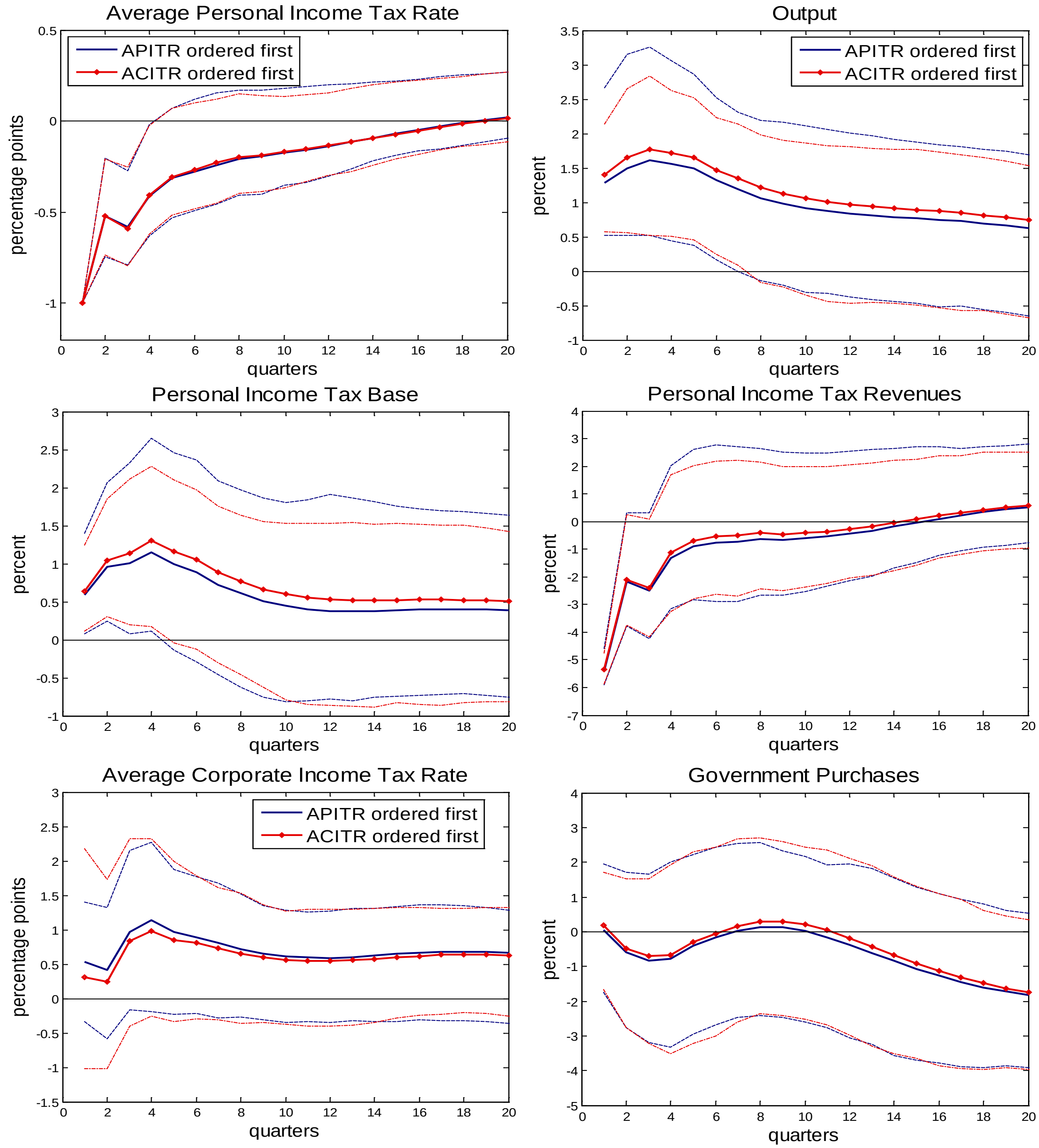
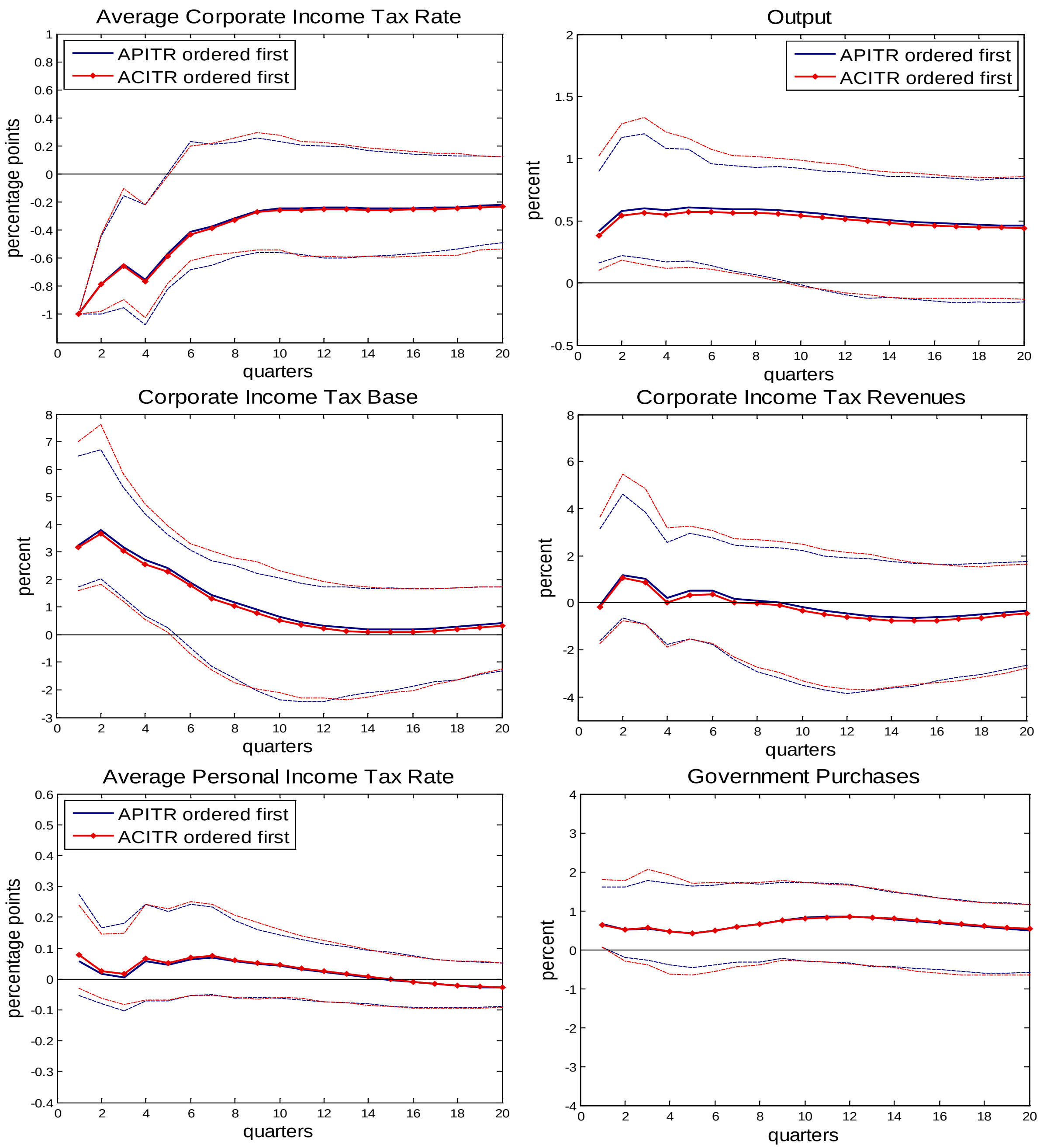


Figure 12: Replication of Mertens and Ravn (2013) Figure 3



Morley and Piger (2012)

We take the key result from this paper as Figure 3, namely that their model-averaged measure of the business cycle displays an asymmetric shape, with the variance of output in a recession being greater than the variance of output during an expansion. We obtain the replication files from ReStat's website. The authors specify that they ran their programs in Gauss 10 on Mac OS X in their readme (one of the few papers that specify a software version-operating system combination), but we did not encounter any issues running their programs on our version of Gauss (version 9.0.2 for Linux). We replicate their Figure 3 closely with help from the authors.

Figure 13: Replication of Morley and Piger (2012) Figure 3

NBER recessions are shaded.

Nakov and Pescatori (2010)

We take the key result as Table 5, although Tables 3 to 4 and 6to 7 lend supporting evidence to Table 5, namely that "Around half of the reduced volatility of inflation [since the mid-1980s] is explained by better monetary policy alone, and 57% of the reduced volatility of GDP growth is attributed to smaller TFP [total factor productivity] shocks. Oil related effects explain around a third" (Nakov and Pescatori (2010) abstract). Our estimates of the posterior distribution in their Table 3 are slightly off, but the qualitative results of the paper still hold. We perform the estimation in Matlab R2012a (Linux).

Table 19: Replication of Nakov and Pescatori (2010) Table 3, Posterior Parameters

	Mean	Standard Deviation	Mode
1970-1983			
θ	0.672	0.068	0.622
ψ	1.103	0.224	0.897
ϕ_i	0.579	0.079	0.537
ϕ_π	1.246	0.359	2.224
ϕ_y	0.557	0.108	0.531
ρ_a	0.942	0.017	0.969
ρ_b	0.894	0.035	0.896
ρ_z	0.912	0.032	0.927
ρ_ω	0.904	0.031	0.937
$100\sigma_a$	0.014	0.001	0.013
$100\sigma_b$	0.031	0.006	0.022
$100\sigma_z$	0.198	0.024	0.213
$100\sigma_\omega$	0.374	0.068	0.307
$100\sigma_r$	0.004	0.001	0.005
1984-2007			
θ	0.471	0.063	0.477
ψ	1.065	0.238	1.009
ϕ_i	0.675	0.059	0.691
ϕ_π	3.193	0.295	3.099
ϕ_y	0.533	0.098	0.539
ρ_a	0.978	0.010	0.983
ρ_b	0.952	0.015	0.951
ρ_z	0.881	0.033	0.882
ρ_ω	0.953	0.018	0.960
$100\sigma_a$	0.006	0.000	0.006
$100\sigma_b$	0.023	0.005	0.019
$100\sigma_z$	0.149	0.016	0.152
$100\sigma_\omega$	0.257	0.047	0.234
$100\sigma_r$	0.002	0.000	0.002

Table 20: Replication of Nakov and Pescatori (2010) Table 4

	1970-1983		1984-2007		Volatility Reduction	
	Data	Model	Data	Model	Data	Model
Inflation	1.20	1.64	0.52	0.72	57%	56%
GDP Growth	0.57	0.61	0.25	0.25	57%	58%
Interest Rate	0.88	0.89	0.57	0.43	35%	51%
Real Oil Price	18.96	16.62	13.00	12.02	31%	28%

This table displays the second moments of the data and the Nakov and Pescatori (2010) model, with the implied volatility reduction from the first period to the second period in percent.

Table 21: Replication of Nakov and Pescatori (2010) Table 5

	Oil		Monetary Policy			
	Share	Shocks	Rule	Shocks	TFP Shock	Other Factors
Inflation	32%	16%	39%	11%	2%	-1%
GDP Growth	18%	11%	0%	4%	57%	10%
Interest Rate	12%	3%	37%	4%	8%	37%
Real Oil Price	-3%	101%	0%	0%	0%	1%

This table shows the Nakov and Pescatori (2010) model-implied percent contributions to reduced volatility by changing parameters from their estimated pre-1984 values to their estimated values for 1984 and later. Positive numbers indicate a percent volatility reduction. TFP = total factor productivity.

Table 22: Replication of Nakov and Pescatori (2010) Table 6

		1970-1983	1984-2007	Counterfactual s_0
Elasticity of oil in Production	s_0	0.047	0.026	0.026
Common slope coefficient	λ	0.668	1.748	0.664
Oil markup pass-through	$s_0\lambda$	0.032	0.046	0.018
Oil markup volatility	$std(\hat{\nu}_t)$	0.238	0.175	0.233
Oil markup persistence	$\rho(\hat{\nu}_t)$	0.934	0.939	0.933
Output gap coefficient	$(1-s_0)\lambda$	0.637	1.702	0.647
Output gap volatility	$std(\hat{y}_t)$	0.012	0.005	0.007
Output gap persistence	$\rho(\hat{y}_t)$	0.886	0.913	0.779

This table displays estimates of Phillips curve parameters and counterfactual Phillips curve parameters under the scenario of a 44% reduction in the elasticity of oil in production.

Table 23: Replication of Nakov and Pescatori (2010) Table 7

	US Shocks			Oil Shocks	
	Real		Nominal		
	$\hat{a}$	$\hat{b}$	$\hat{r}$	$\hat{z}$	$\hat{\omega}$
pre-1984					
Inflation	2.74	33.34	15.69	12.23	36.01
GDP Growth	69.09	1.10	5.00	22.68	2.13
Interest Rate	8.70	69.31	4.44	0.27	17.27
Real oil price	0.05	0.00	0.04	87.09	12.81
post-1984					
Inflation	0.80	37.45	41.49	1.18	19.08
GDP Growth	77.83	0.42	2.12	17.49	2.14
Interest Rate	3.29	87.42	1.10	1.08	7.12
Real oil price	0.02	0.00	0.01	87.71	12.26

This table displays the variance decomposition of inflation, GDP growth, the interest rate, and the real oil price to various shocks. $\hat{a}$: TFP shock. $\hat{b}$: discount factor shock. $\hat{r}$: inflation shock. $\hat{z}$: oil-sector technology shock. $\hat{\omega}$: fringe oil producers capacity shock.

Ramey (2011)

We take the key results from this paper as Figures 10 and 12. The QJE does not have a data and code replication policy. We downloaded the replication files from Valerie Ramey's personal website. We were able to replicate Ramey (2011)'s results exactly using these replication files using Stata 13.1 (Windows).

Reis and Watson (2010)

We take the key results to be their Tables 4 and 5. In their Table 4, panels A-C there exists a significant association between GDP and PCE inflation because these panels only control for absolute price changes. Their Table 4, panels D-E control for relative price changes, and the association between GDP and PCE inflation disappears. Similarly, in their Table 5, controlling for relative prices in panels B-D removes the association between pure inflation and GDP. We ran into difficulties running the code for Reis and Watson (2010), which we suspect was partly caused by running the code on a different version of Gauss. We have access to Gauss 9.0.2 for Linux, but were unsure which versions the authors used. We were able to replicate these tables exactly without assistance from the authors.

Romer and Romer (2010)

We take the key figures as their Figures 6, 7, and 9. Due to data constraints, we only replicate panel B of Figure 7, although we have the data to replicate all of their Figures 6 and 9. We were able to replicate their Figures 6, 9, and panel B of their Figure 7 using code and data from the AER's website without assistance from the authors using RATS 7.10 (Linux).

Schmitt-Grohé and Uribe (2011)

We take the key result as Table 7. We used the data and code files from the RED's website, but the files were missing a function needed to create Table 7. On request, the authors provided us with some code that we modified to create Table 7, which we were able to replicate exactly using Matlab R2013a (Windows).

Schmitt-Grohé and Uribe (2012)

We take the key results as Figure 2, Table 3, Table 5, and Table 6. We used the replication files on Econometrica's website, which produced the author's results without assistance using Matlab R2008a (Windows).