Online Appendix: Replication Results

Is Economics Research Replicable? Sixty Published Papers From Thirteen Journals Say "Often Not"

This Online Appendix contains the dependent variables, the text of each paper that we used to identify the key results, and the descriptions of our replication attempts for each one of the 67 papers. We do not display the published results for the papers that we were able to exactly replicate. We also do not display results for the papers that we were unable to replicate and were unable to obtain any output. For papers that we were able to qualitatively, but not exactly, replicate and for the papers that we were unable to replicate but were able to obtain some output, we include both the original published results and our replication results. For a given replication, we occasionally supplemented the author-provided files with FAME code (version 10.2).

7.1 Alexopoulos (2011), "Read All About it!! What Happens Following a Technology Shock?"

Dependent variables: TFP, real GDP per hour, real investment.

Key Results: "Changes in information technology are found to be important sources of economic fluctuations..."(abstract)

This paper used proprietary data, so we were unable to replicate any of the key results. We classified this paper as "failed due to proprietary data."

7.2 Alexopoulos and Cohen (2011), "Volumes of Evidence: Examining Technical Change in the Last Century Through a New Lens"

Dependent variables: TFP, real GDP.

Key Results: "A strong, causal relationship is found to exist between these [technological innovation] indicators and changes in TFP and output per capita." (abstract)

This paper used proprietary data, so we were unable to replicate any of the key results. We classified this paper as "failed due to proprietary data."

7.3 Ang, Boivin, Dong, and Loo-Kung (2011), "Monetary Policy Shifts and the Term Structure"

Dependent variables: Various yields, including the 3-month T-bill.

Key Results: "The monetary policy loading on the [Taylor Rule's] output gap has averaged around 0.4 and has not changed very much over time. The overall response of the yield curve to output gap components is relatively small." (abstract)

We only had the data but not the code for this paper. Since we were unable to replicate any of the key results, we classified this paper as "failed due to missing public data or code."

7.4 Arestis, Chortareas, and Tsoukalas (2010), "Money and Information in a new Neoclassical Synthesis Framework"

Dependent variables: log of real GDP per capita, log of the implicit GDP deflator, Federal Funds Rate, 3-month T-bill rate, ln(real M2 per capita), ln(nominal M2 per capita).

Key Results: "money has information value... reflected in higher precision in terms of unobserved model concepts such as the natural rate of output." (abstract)

We had neither data or code for this paper. Since we were unable to replicate any of the key results, we classified this paper as "failed due to missing public data or code."

7.5 Auerbach and Gorodnichenko (2012, 2013), "Measuring the Output Responses to Fiscal Policy"

Dependent variables: Real natural logs of: Government purchases, Government receipts, GDP.

Key Results: "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." (abstract).

We took the key results of Auerbach and Gorodnichenko (2012) as the top panel of Table 1, and Figures 2, 4, 5, and 7. We attempted 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 were qualitatively similar to the published paper, but the results did not match exactly with Matlab R2013a (Windows). For Table 1, we obtained many point estimates that matched Auerbach and Gorodnichenko (2013) exactly, but a few of the other estimates were a bit different. For example, we estimated the cumulative point estimate for defense spending under a recession to be 2.18, while Auerbach and Gorodnichenko (2013) obtained 3.69. Similarly, we were able to match a majority of the figures exactly, but there were some differences in the magnitudes. We were able to replicate the paper without assistance from the authors, so we classified this replication as successful.

Table 2: Auerbach and Gorodnichenko (2012) Table 1, Top Panel

	Max Point Estimate	Standard Error	Cumulative Point Estimate	Standard Error
Total Spending				
Linear	0.87	0.29	0.58	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	0.66	0.21	-1.03	0.25
Recession	5.28	0.91	3.69	0.83
Nondefense Spending				
Linear	1.69	0.08	2.08	0.15
Expansion	1.21	0.16	1.17	0.15
Recession	1.22	0.29	1.34	0.31
Consumption Spending				
Linear	0.82	0.28	0.89	0.29
Expansion	0.12	0.13	-0.16	0.11
Recession	2.28	0.64	1.37	0.35
Investment Spending				
Linear	2.07	0.60	2.75	0.60
Expansion	2.82	0.26	1.94	0.17
Recession	2.79	0.52	4.26	0.46

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

Table 3: 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. We found a smaller multiplier (both maximum and cumulative) for defense spending in recessions, but otherwise our replication was similar to the reported results.

Figure 4: Auerbach and Gorodnichenko (2013) Figure 2

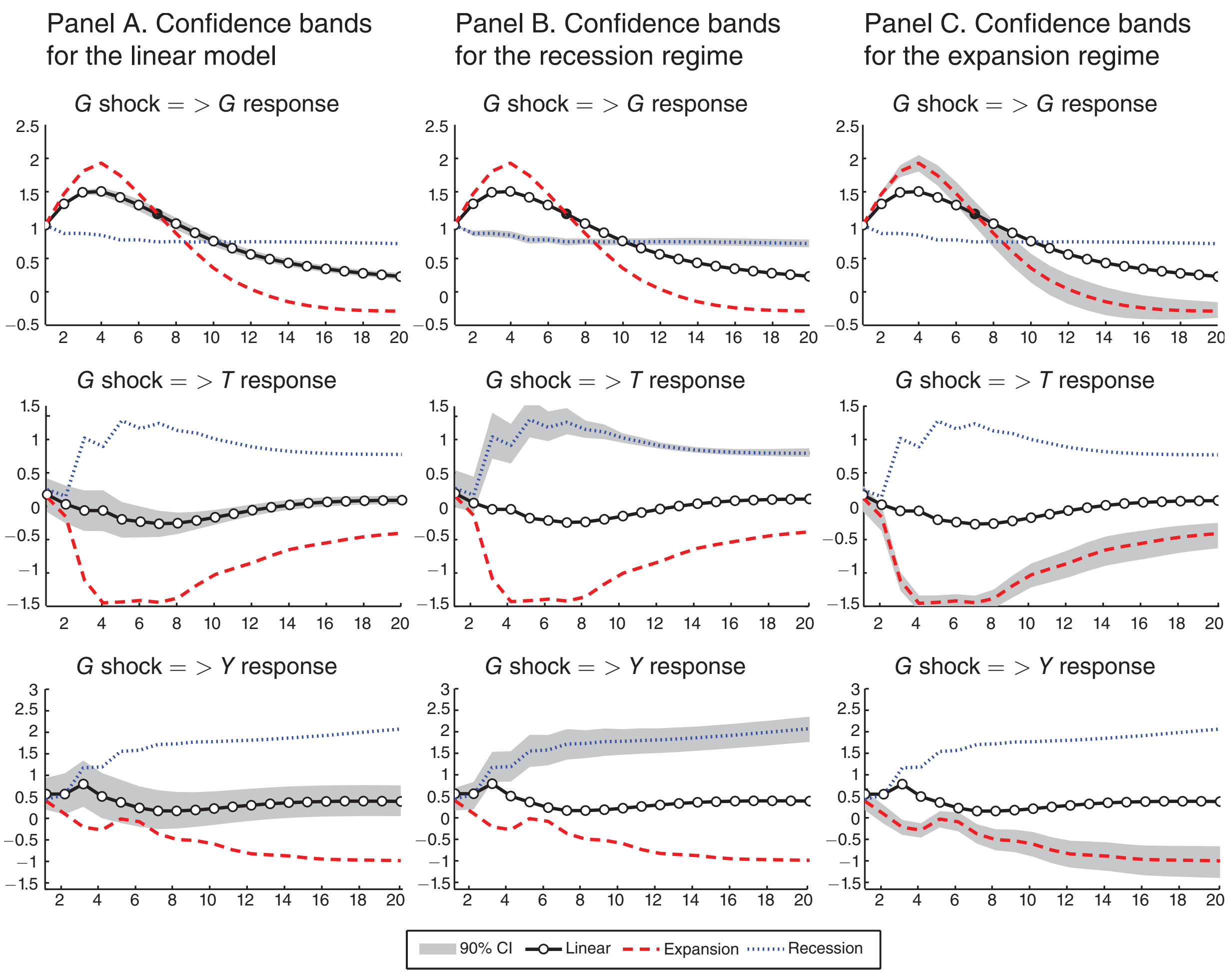
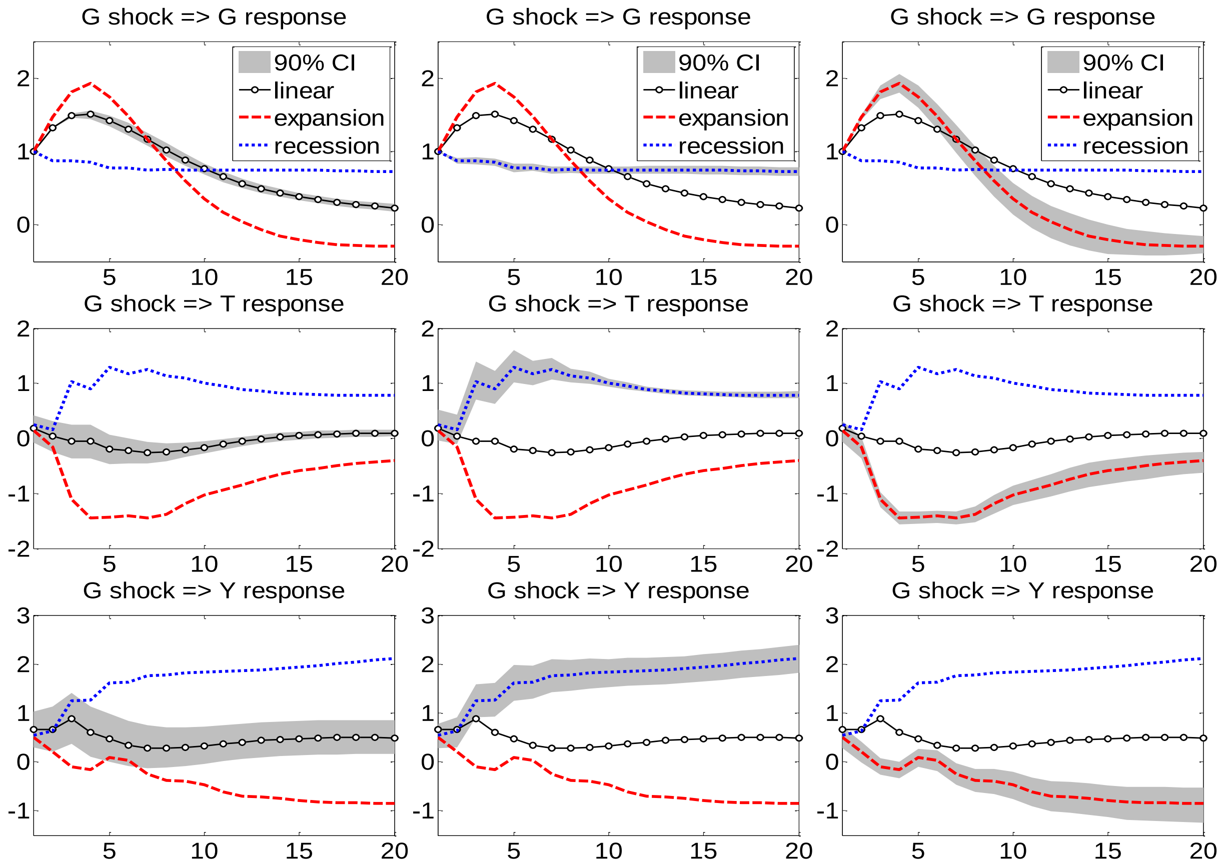
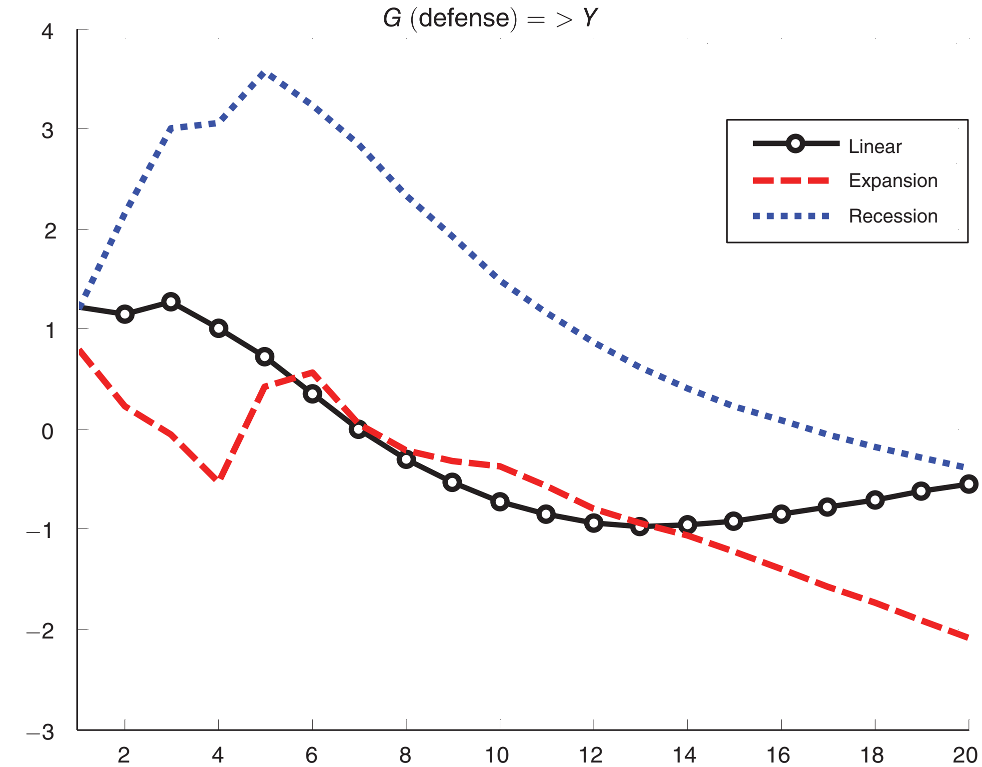


Figure 5: Replication of Auerbach and Gorodnichenko (2013) Figure 2



Our replication was close to Auerbach and Gorodnichenko (2013).

Figure 6: Auerbach and Gorodnichenko (2012) Figure 4



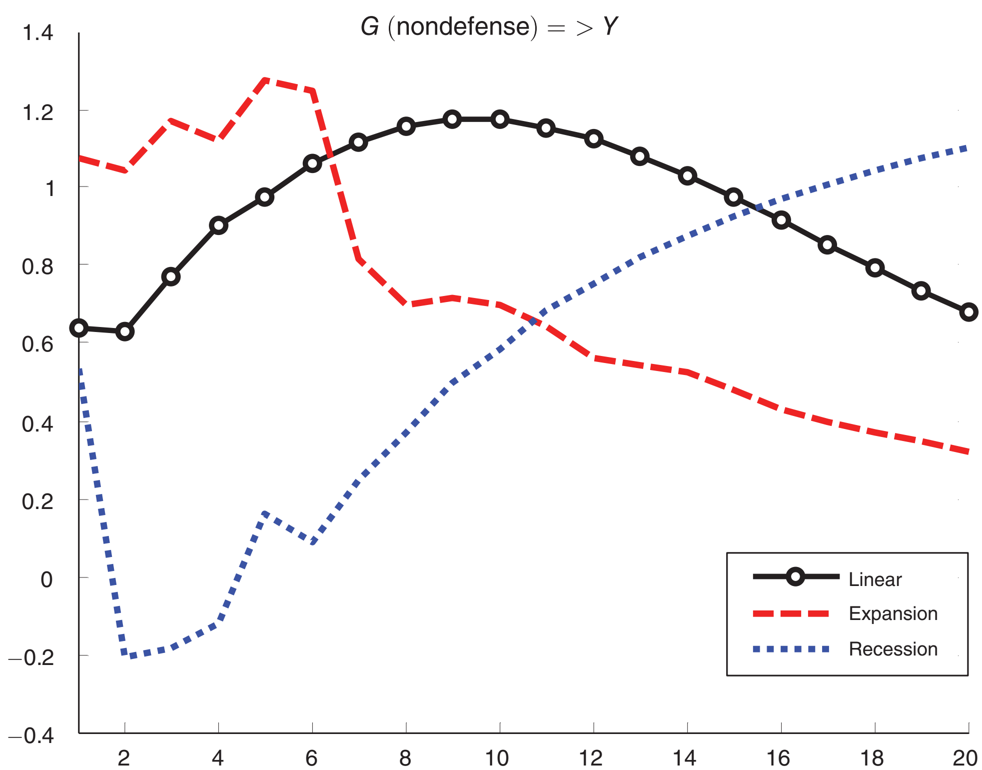
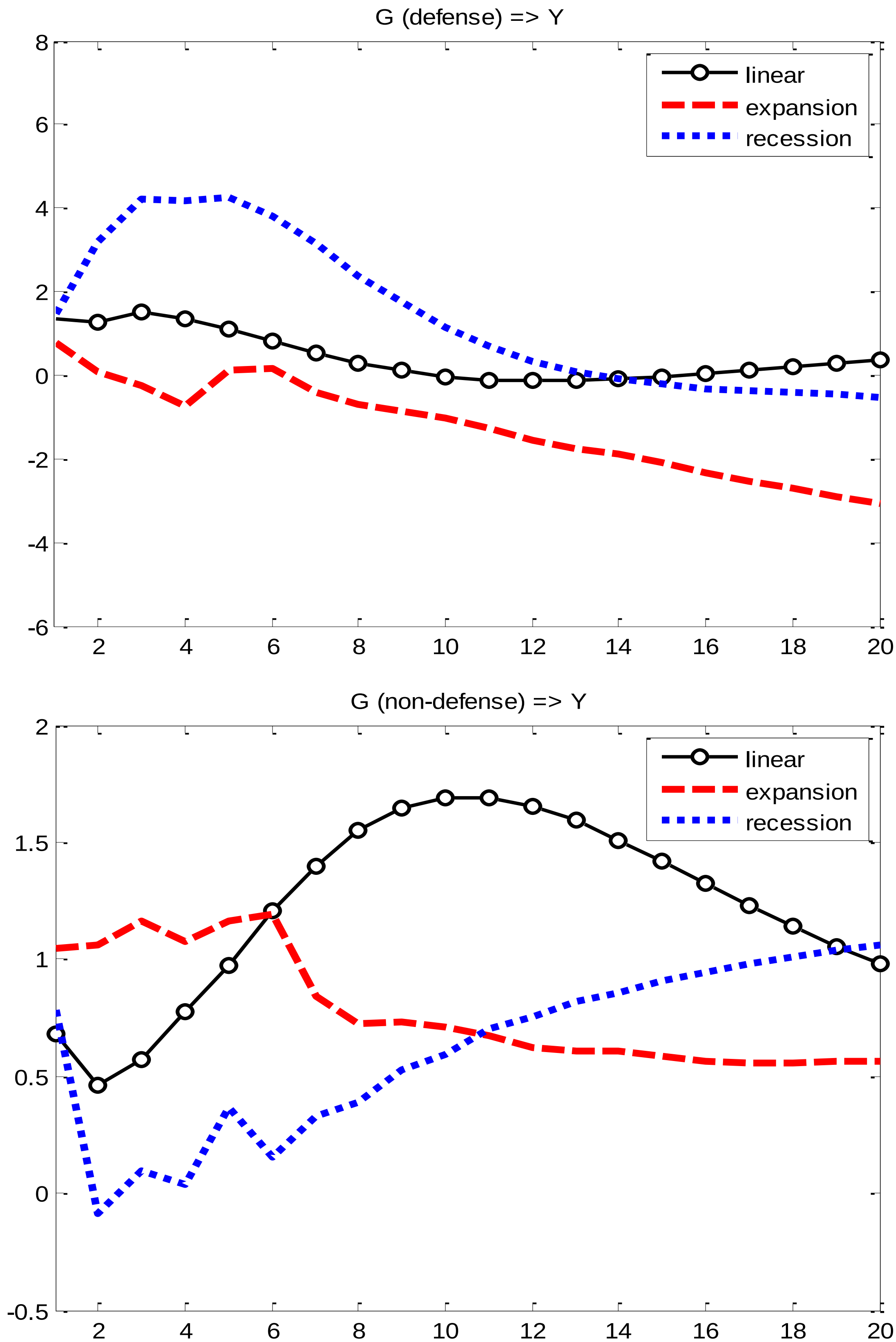


Figure 7: Replication of Auerbach and Gorodnichenko (2012) Figure 4



Our replication was close to Auerbach and Gorodnichenko (2012).

Figure 8: Auerbach and Gorodnichenko (2012) Figure 5

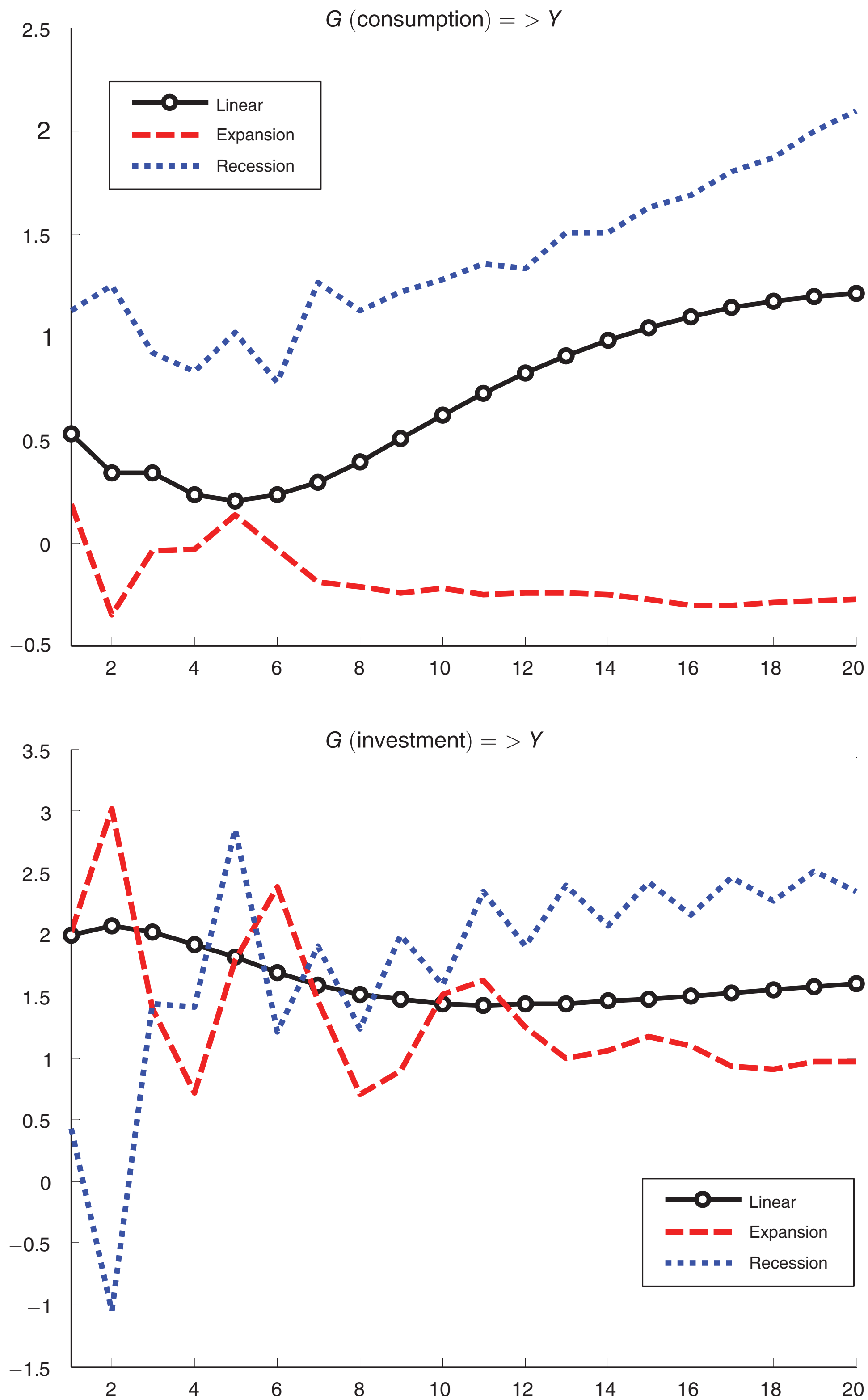
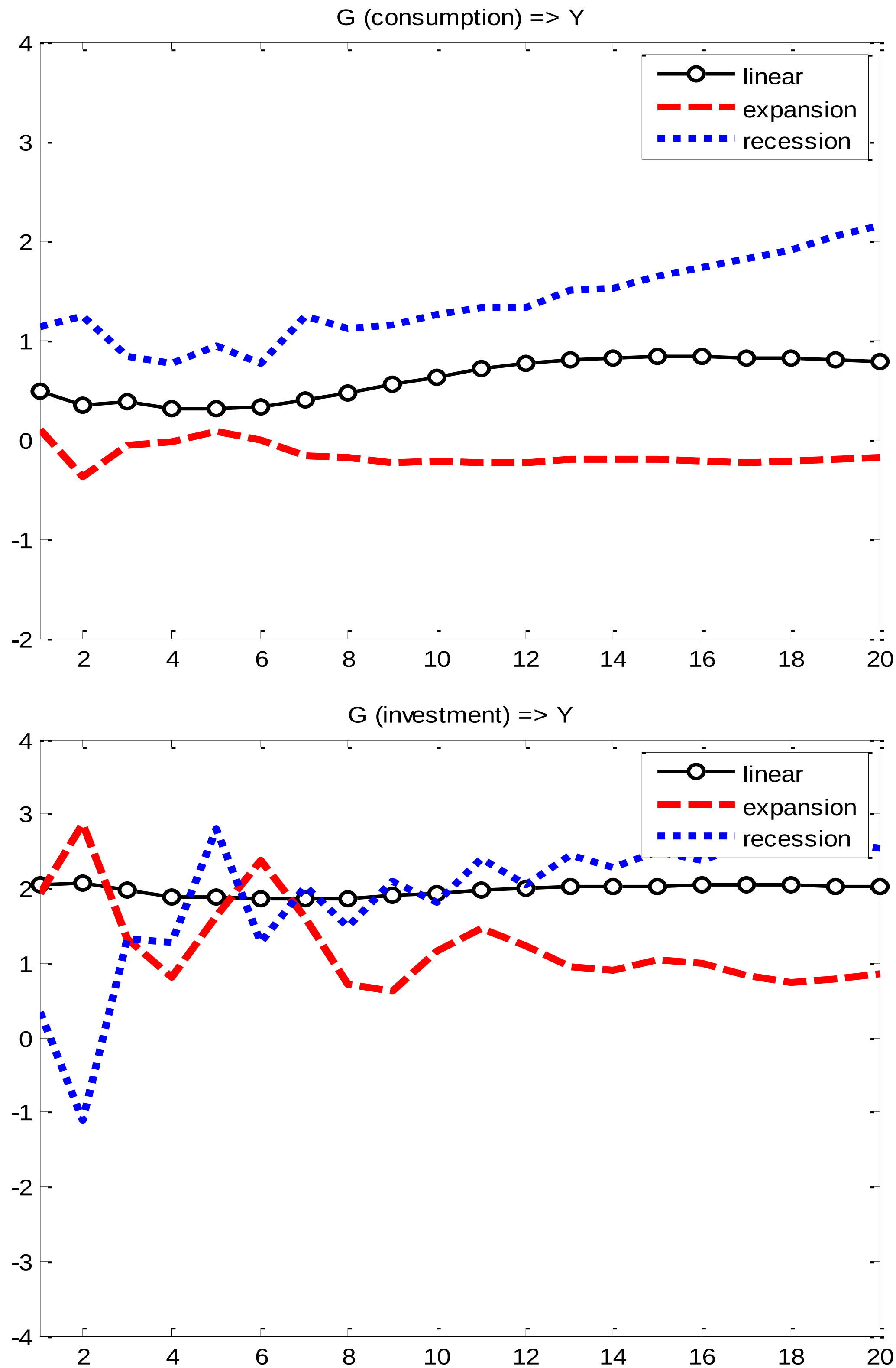


Figure 9: Replication of Auerbach and Gorodnichenko (2012) Figure 5



Our replication was close to Auerbach and Gorodnichenko (2012).

Figure 10: Auerbach and Gorodnichenko (2012) Figure 7

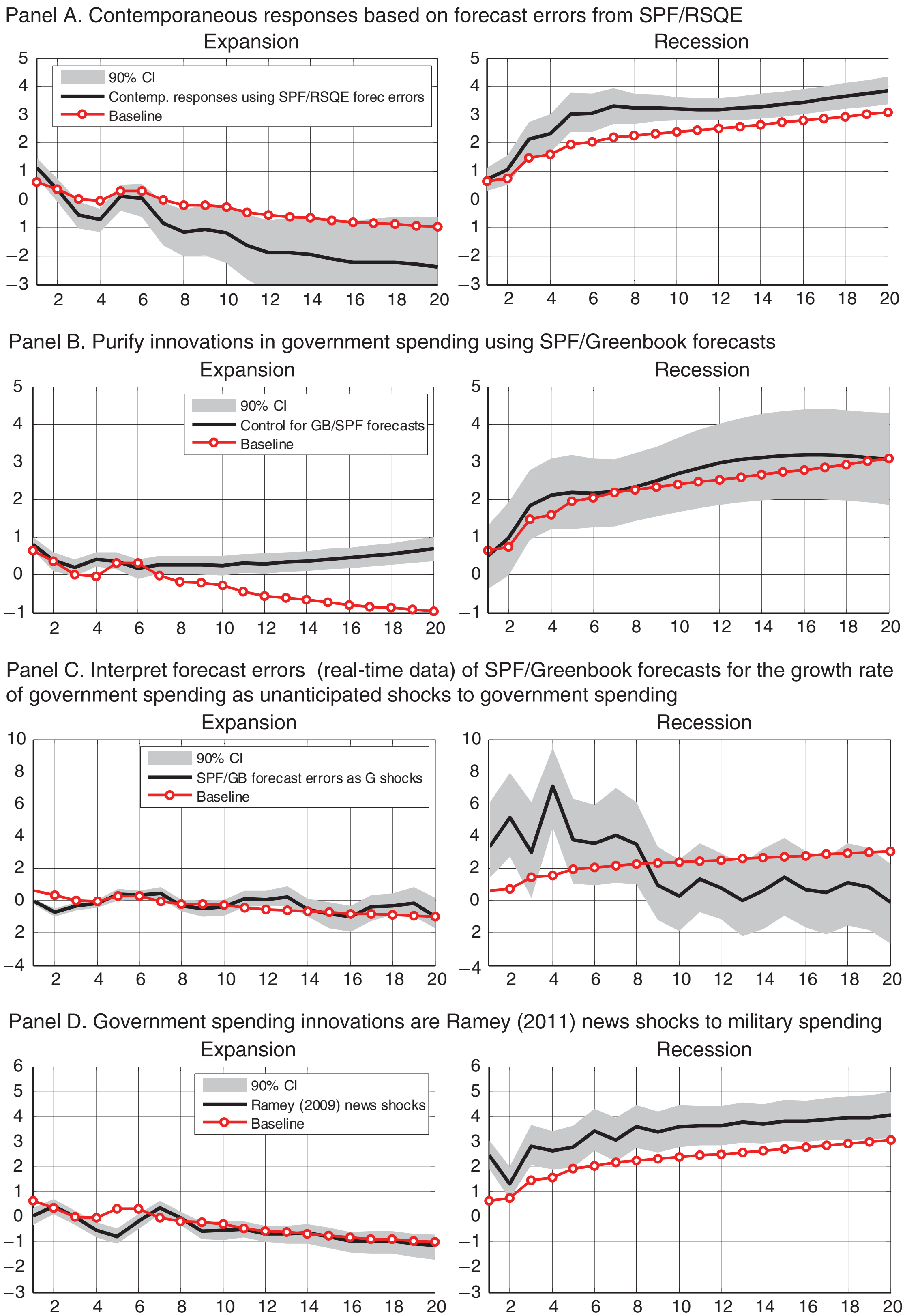
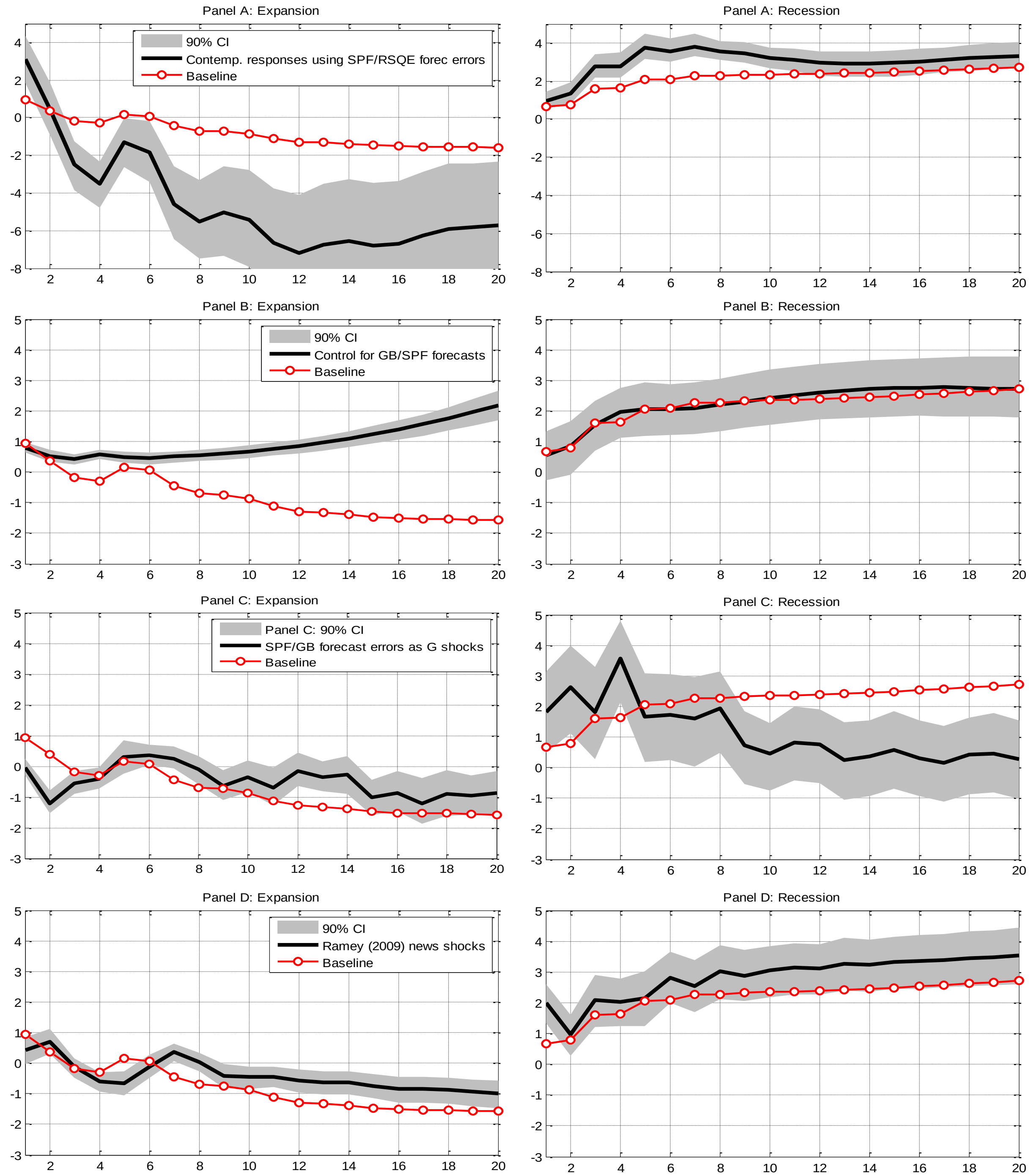


Figure 11: Replication of Auerbach and Gorodnichenko (2012) Figure 7



In Panels A and B our replication showed bigger long-run effects with larger standard errors than Auerbach and Gorodnichenko (2012). In the recession box of Panel C, our replicated effect was smaller on impact. Our replication was similar to Auerbach and Gorodnichenko (2012) for the remainder of the Figure.

7.6 Bai and Wang (2011), "Conditional Markov Chain and its Application in Economic Time Series Analysis"

Dependent variables: log-differences of GDP and employment. Estimated mean and variance state probabilities for GDP and employment.

Key Results: "we find that there is a volatility change [of GDP and employment] at around the first quarter of 1984" (introduction, pg. 716)

We only had the data but not the code for this paper. Since we were unable to replicate any of the key results, we classified this paper as "failed due to missing public data or code."

7.7 Bansak, Graham, and Zebedee (2012), "Business Cycles and Gender Diversification: An Analysis of Establishment-Level Gender Dissimilarity"

Dependent variables: Establishment employment gender-dissimilarity index.

Key Results: "although gender-related occupational segregation has declined over the years 1966-2010, economic downturns interrupt this trend." (pg. 561)

This paper used proprietary data, so we were unable to replicate any of the key results. We classified this paper as "failed due to proprietary data."

7.8 Barro and Redlick (2011), "Macroeconomic Effects from Government Purchases and Taxes"

Dependent variables: Growth of real GDP per capita.

Key Results: "all estimated [spending] multipliers are significantly less than one." (abstract)

We took the key results of Barro and Redlick (2011) as Tables 2, 3, 5, and 7. Robert Barro provided us with an EViews file, which contained all of the data and some of the specifications reported in the paper, 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 were able to replicate a vast majority of their point estimates exactly, except for column (8) of their Table 7, where we found slightly different estimates. Our standard errors were slightly different than the published estimates across all specifications. We classified this replication as successful.

Table 4: 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	0.47	0.98
	(0.27)*	(0.06)**	(0.08)**	(0.08)**	(0.08)**	(0.65)
$\Delta g : defense$	0.01	0.20	0.21	0.25	0.16	-0.54
(-1)	(0.28)	(0.06)**	(0.09)*	(0.08)**	(0.08)	(0.56)
$\Delta g* : defense$	0.026	0.039	0.034	0.034	0.034	-0.12
$news$	(0.016)	(0.011)**	(0.015)*	(0.014)*	(0.017)	(0.112)
$U(-1)$	0.50	0.58	0.61	0.58	0.47	0.51
	(0.17)**	(0.14)**	(0.10)**	(0.10)**	(0.10)	(0.18)**
$\Delta\tau(-1)$	-0.54	-0.16	-0.26	-0.52	-0.19	-0.48
	(0.21)**	(0.16)	(0.22)	(0.23)*	(0.25)	(0.22)*
Yield Spread	-43.9	-37.8	-101.5	-103.4	-73.6	-43.1
Squared	(20.7)*	(22.0)	(12.8)**	(12.4)**	(12.2)**	(21.8)*
p-value, defense variables	0.03	0.00	0.00	0.00	0.00	0.47

Table 5: Replication of Barro and Redlick (2011) Table 2

	Starting date					
	(1) 1950	(2) 1939	(3) 1930	(4) 1930 (w/o 1949)	(5) 1917	(6) 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	-43.1
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

We found similar point estimates and slightly smaller standard errors than Barro and Redlick (2011).

Table 6: 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
	(0.27)**	(0.08)**	(0.32)	(0.10)**	(0.24)**	(0.26)
$\Delta g : defense$	-0.13	0.21	0.08	0.18	-0.36	0.02
(-1)	(0.27)	(0.09)*	(0.28)	(0.09)*	(0.25)	(0.26)
$\Delta g* : defense$	0.040	0.036	0.028	0.033	0.014	0.016
$news$	(0.016)**	(0.016)*	(0.016)	(0.015)*	(0.013)	(0.014)
$U(-1)$	0.64	0.60	0.43	0.62	0.26	0.55
	(0.17)**	(0.11)**	(0.18)**	(0.10)**	(0.16)*	(0.16)**
$\Delta\tau(-1)$	-0.45	-0.25	-0.56	-0.25	-0.26	-0.38
	(0.20)*	(0.23)	(0.21)**	(0.22)	(0.19)	(0.20)*
Yield Spread	-31.2	-100.9	-28.4	-102.3	-38.9	-21.6
Squared	(20.0)	(13.3)**	(25.4)	(13.0)**	(18.1)*	(20.5)
$\Delta g : nondefense$	2.65	0.12				
	(0.93)**	(0.63)				
$\Delta(transfers)$			-1.53	0.64		
			(0.92)	(0.68)		
$\Delta(GMsales)$					3.66	
					(0.86)**	
$\Delta(GEsales)$						17.6
						(4.7)**

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)**

We found similar point estimates and slightly smaller standard errors than Barro and Redlick (2011).

Table 8: 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	-0.171	-0.083	-0.081	0.004
	(0.093)	(0.073)*	(0.185)	(0.041)	(0.079)
$\Delta g : defense$	0.179	0.147	-0.142	0.055	-0.231
(-1)	(0.095)	(0.075)*	(0.189)	(0.042)	(0.080)**
$\Delta g* : defense$	-0.0035	0.0106	0.0377	-0.0055	-0.0135
$news$	(0.0053)	(0.0041)**	(0.0105)**	(0.0023)*	(0.0044)**
$U(-1)$	0.112	0.145	0.382	-0.053	-0.095
	(0.058)	(0.045)**	(0.115)**	(0.026)*	(0.049)*
$\Delta\tau(-1)$	-0.184	-0.145	-0.30	-0.033	0.122
	(0.071)**	(0.056)**	(0.142)*	(0.032)	(0.060)*
Yield Spread	-5.4	-3.5	-22.7	-4.80	-6.7
Squared	(7.0)	(5.5)	(13.9)	(3.1)	(5.0)
Sample: 1939-2006					
$\Delta g : defense$	-0.011	-0.115	-0.356	-0.009	-0.071
	(0.022)	(0.016)**	(0.045)**	(0.011)	(0.021)**
$\Delta g : defense$	0.107	0.038	0.096	-0.011	-0.027
(-1)	(0.022)**	(0.016)*	(0.046)*	(0.011)	(0.022)
$\Delta g* : defense$	0.0044	0.0116	0.0341	-0.0082	-0.0023
$news$	(0.004)	(0.003)**	(0.0084)**	(0.0021)**	(0.0039)
$U(-1)$	0.101	0.094	0.401	-0.030	-0.002
	(0.052)	(0.038)**	(0.109)**	(0.027)	(0.051)
$\Delta\tau(-1)$	-0.008	-0.103	-0.067	-0.105	0.114
	(0.059)	(0.043)*	(0.124)	(0.030)**	(0.058)*
Yield Spread	1.1	-3.1	-20.3	-6.5	-8.0
Squared	(8.0)	(5.9)	(16.8)	(4.1)	(7.8)
Sample: 1930-2006					
$\Delta g : defense$	-0.001	-0.110	-0.34	-0.016	-0.074
	(0.038)	(0.017)**	(0.051)**	(0.016)	(0.020)**
$\Delta g : defense$	0.11	0.036	0.087	-0.003	-0.024
(-1)	(0.040)**	(0.018)*	(0.053)	(0.017)	(0.021)
$\Delta g* : defense$	-0.0004	0.0113	0.0353	-0.0096	-0.002
$news$	(0.0068)	(0.0031)**	(0.0092)**	(0.003)**	(0.0036)
$U(-1)$	0.17	0.082	0.30	0.041	0.006
	(0.047)**	(0.021)**	(0.063)**	(0.020)*	(0.025)
$\Delta\tau(-1)$	-0.060	-0.112	-0.010	-0.111	0.113
	(0.101)	(0.047)*	(0.136)	(0.044)**	(0.053)*
Yield Spread	-42.3	-12.9	-39.9	-4.9	-1.1
Squared	(5.9)**	(2.7)**	(7.9)**	(2.5)	(3.1)

Table 9: 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)

We found similar point estimates and slightly smaller standard errors than Barro and Redlick (2011).

Table 10: Barro and Redlick (2011), Table 7, Columns 1-4

	(1)	(2)	(3)	(4)
$\Delta g : defense$	0.67	0.53	0.66	0.61
	(0.28)*	(0.27)*	(0.28)**	(0.35)
$\Delta g : defense$	0.01	-0.23	-0.05	0.05
(-1)	(0.28)	(0.28)	(0.29)	(0.32)
$\Delta g* : defense$	0.025	0.029	0.027	0.023
$news$	(0.015)	(0.016)	(0.016)	(0.018)
$U(-1)$	0.51	0.51	0.48	0.50
	(0.17)**	(0.18)**	(0.17)**	(0.17)**
$\Delta\tau(-1)$	-0.53		0.43	-0.58
	(0.21)		(0.24)*	(0.28)**
$\Delta\tau$				0.12
				(0.47)
Romers: exogenous		-1.08	-0.56	
$[\Delta tax/Y(-1)](-1)$		(0.57)	(0.62)	
Yield Spread	-47.2	-43.4	-41.8	-44.4
Squared	(20.2)*	(21.7)*	(21.2)	(21.9)*
p-value: τ	0.015		0.074	0.039
p-value: Romers		0.063	0.37	
p-value: all tax vars	0.015	0.063	0.029	0.039

Table 11: 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

We found slightly tighter standard errors on the yield spread squared coefficient, but otherwise our replication was fairly close.

Table 12: Barro and Redlick (2011), Table 7, Columns 5-8

	(5)	(6)	(7)	(8)
$\Delta g : defense$	0.53	0.71	0.72	0.49
	(0.28)	(0.30)*	(0.29)**	(0.31)
$\Delta g : defense$	-0.23	-0.21	-0.03	0.10
(-1)	(0.28)	(0.28)	(0.29)	(0.26)
$\Delta g* : defense$	0.029	0.016	0.021	0.015
$news$	(0.016)*	(0.017)	(0.017)	(0.018)
$U(-1)$	0.51	0.49	0.49	0.43
	(0.18)**	(0.18)**	(0.18)**	(0.17)*
$\Delta\tau(-1)$			-0.45	-0.52
			(0.24)	(0.18)**
Romers: exogenous	-1.08			
$[\Delta tax/Y(-1)](-1)$	(0.58)			
Romers: exogenous	-0.03			
$[\Delta tax/Y(-1)]$	(0.55)			
$[\Delta(fedrev.)/Y(-1)](-1)$		-0.46	-0.17	
		(0.27)	(0.30)	
$[\Delta(fedrev.)/Y(-1)]$				0.46
				(0.53)
Yield Spread	-42.9	-64.9	-52.5	-37.4
Squared	(21.9)	(20.7)**	(21.3)**	(21.0)
p-value: τ				0.006
p-value: Romers	0.17		0.070	
p-value: fed. revenue		0.091	0.56	0.39
p-value: all tax vars	0.17	0.091	0.037	0.010

Table 13: Replication of Barro and Redlick (2011), Table 7, Columns 5-8

	(5)	(6)	(7)	(8)
$\Delta g : defense$	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

We found slightly tighter standard errors on the yield spread squared coefficient, and a few different point estimates in column (8) than Barro and Redlick (2011).

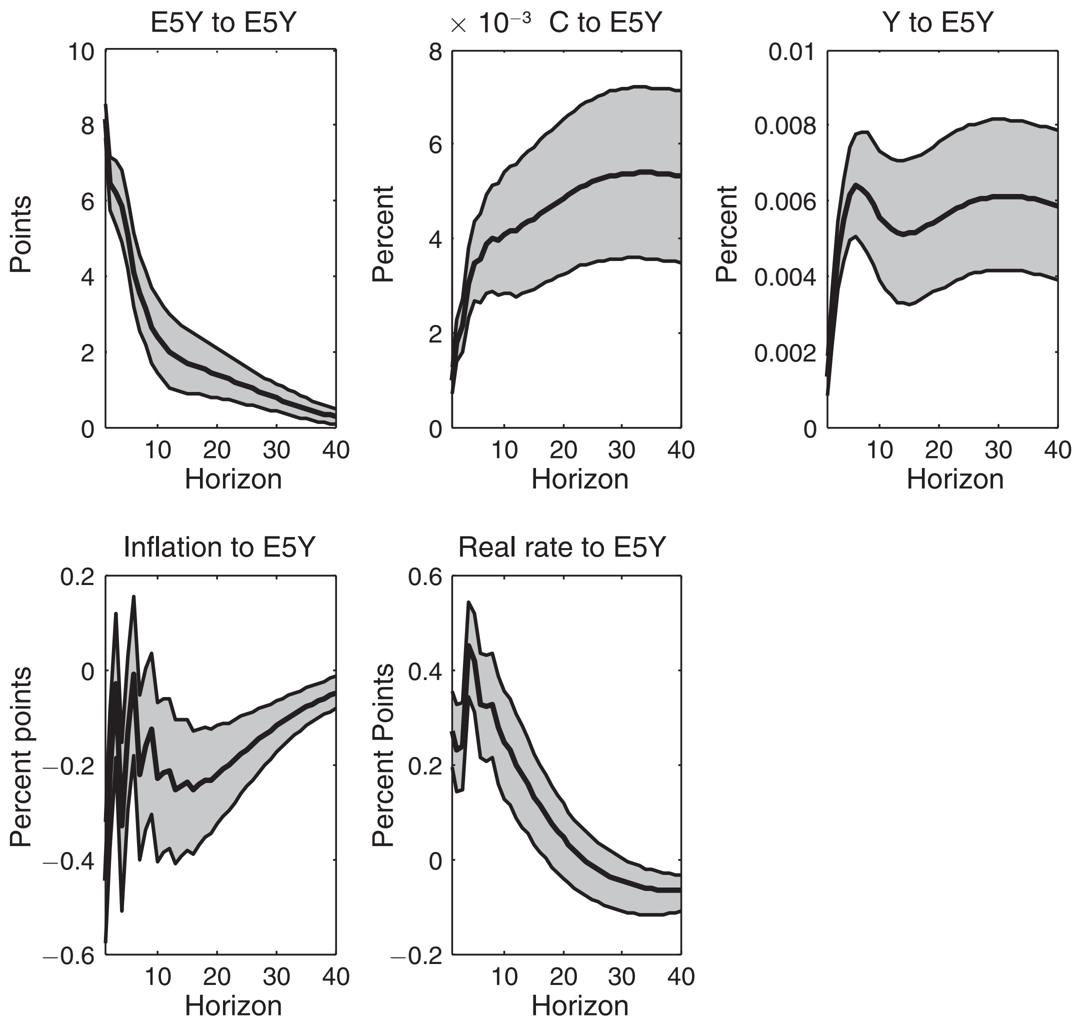
7.9 Barsky and Sims (2012), "Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence"

Dependent variables: Various; often expectations of economic conditions.

Key Results: "the relationship between confidence and subsequent activity is almost entirely reflective of the news..." (abstract)

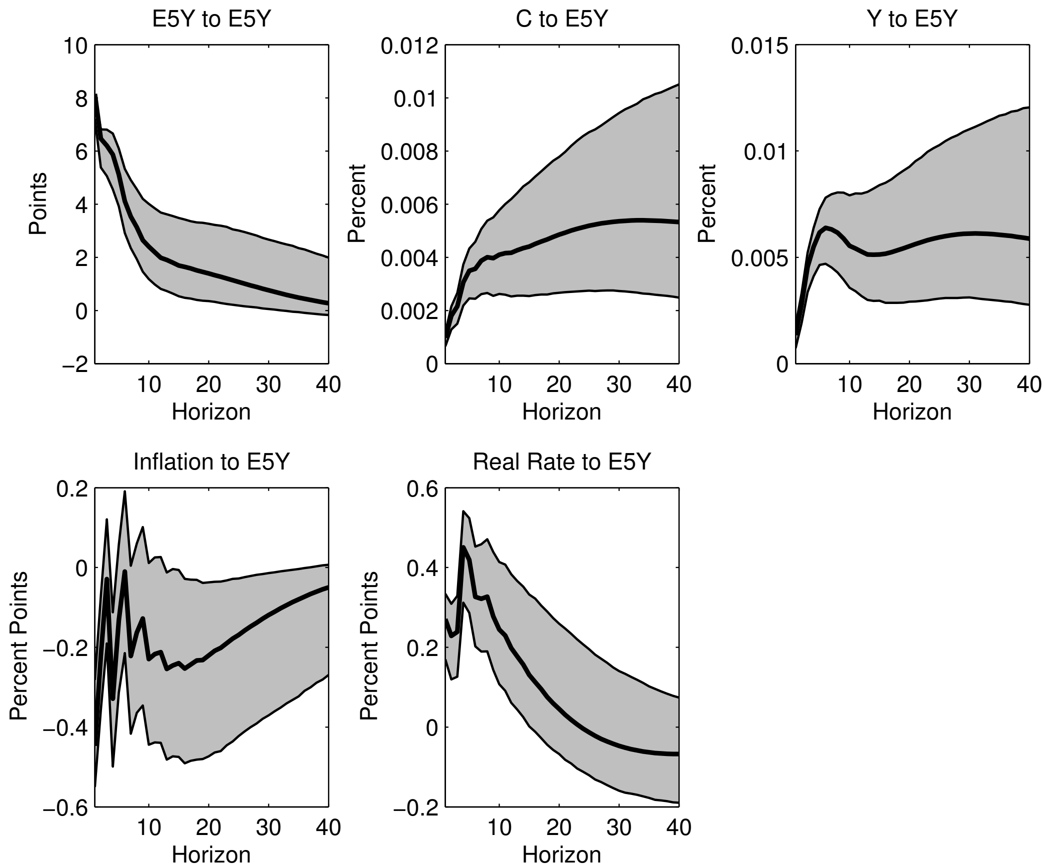
These result corresponded to their Table 3 and Figures 7, 8, 9, 12, and 13. We attempted to replicate the results using the code from the *American Economic Review*'s website. We were able to qualitatively replicate Figure 7 without help from the authors, but we were unable to produce any of the remaining results even with help from Eric Sims. The main issues were missing code dependencies belonging to an "AIM" package, missing code dependencies that we could not attribute to any particular library, and missing matrices, all of which led to Matlab crashing before it could produce the desired results. We also attempted to locate and substitute as many of the missing functions as possible, but Matlab would still crash before producing results. For these reasons, we classified this paper as "failed due to incorrect public data or code."

Figure 12: Barsky and Sims (2012) Figure 7



Impulse responses from a VAR.

Figure 13: Replication of Barsky and Sims (2012) Figure 7



Impulse responses from a VAR. Our replication of the point estimates was very close, but our confidence intervals were a bit wider at further horizons than those reported by Barsky and Sims (2012).

7.10 Baumeister and Peersman (2013), "Time-Varying Effects of Oil Supply Shocks on the US Economy"

Dependent variables: First-differenced natural logs of: Global Oil Production, Real US Refiners' Acquisition Cost of Imported Crude Oil, real GDP, CPI.

Key Results: "decline in the shortrun price elasticity of oil demand since the mid-1980s" (abstract)

We took the key results as Figures 1 and 4 from Baumeister and Peersman (2013), where the text implied that oil production shocks have a greater effect on GDP than oil price shocks. We used the code from the *American Economic Journal: Macroeconomics*' website. This paper required a local version of Matlab, but we were unable to verify the version of Matlab used by the authors. We conducted the estimation in Matlab R2013a (Windows). With assistance from Christiane Baumeister, we were able to exactly replicate their Figure 1, shown in our Figure 15. Also, we were able replicate their Figure 4 with some minor differences, shown in our Figure 17; our "Actual data demeaned" series had less volatility than the ones in Baumeister and Peersman (2013), while the other series were similar. We classified this replication as successful.

Figure 14: Baumeister and Peersman (2013) Figure 1

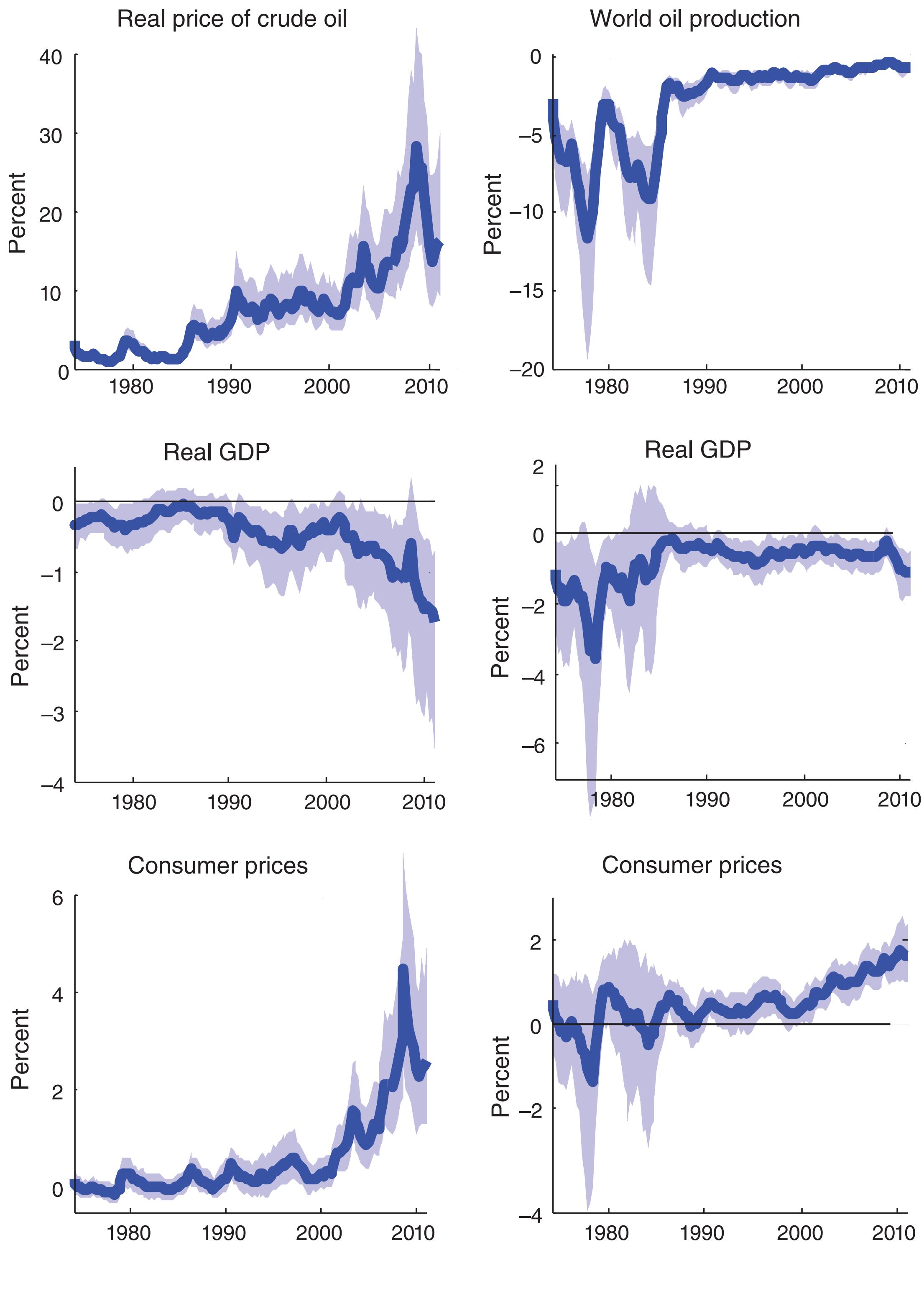
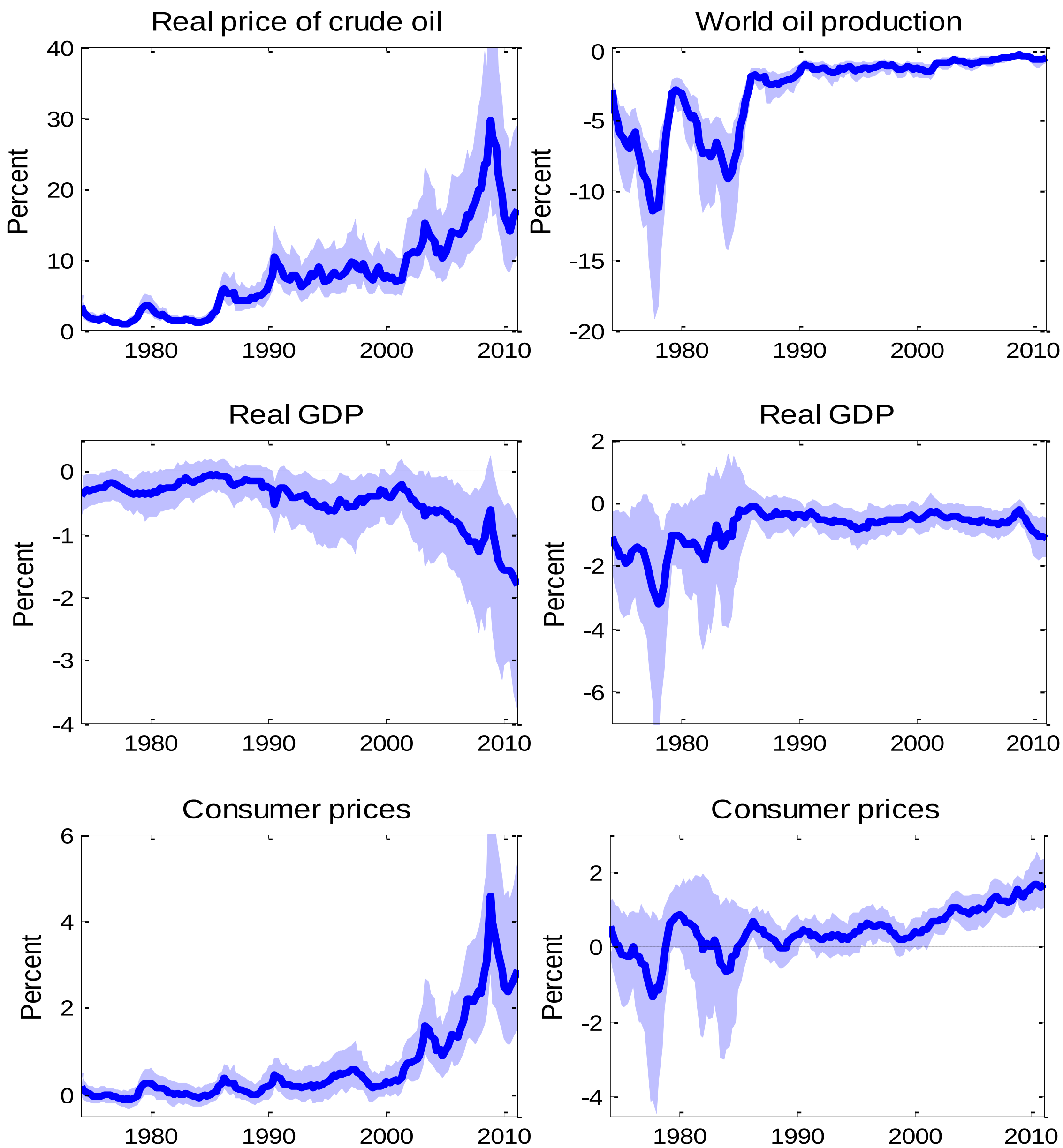


Figure 15: Replication of Baumeister and Peersman (2013) Figure 1



Our replication was very close to Baumeister and Peersman (2013).

Figure 16: Baumeister and Peersman (2013) Figure 4

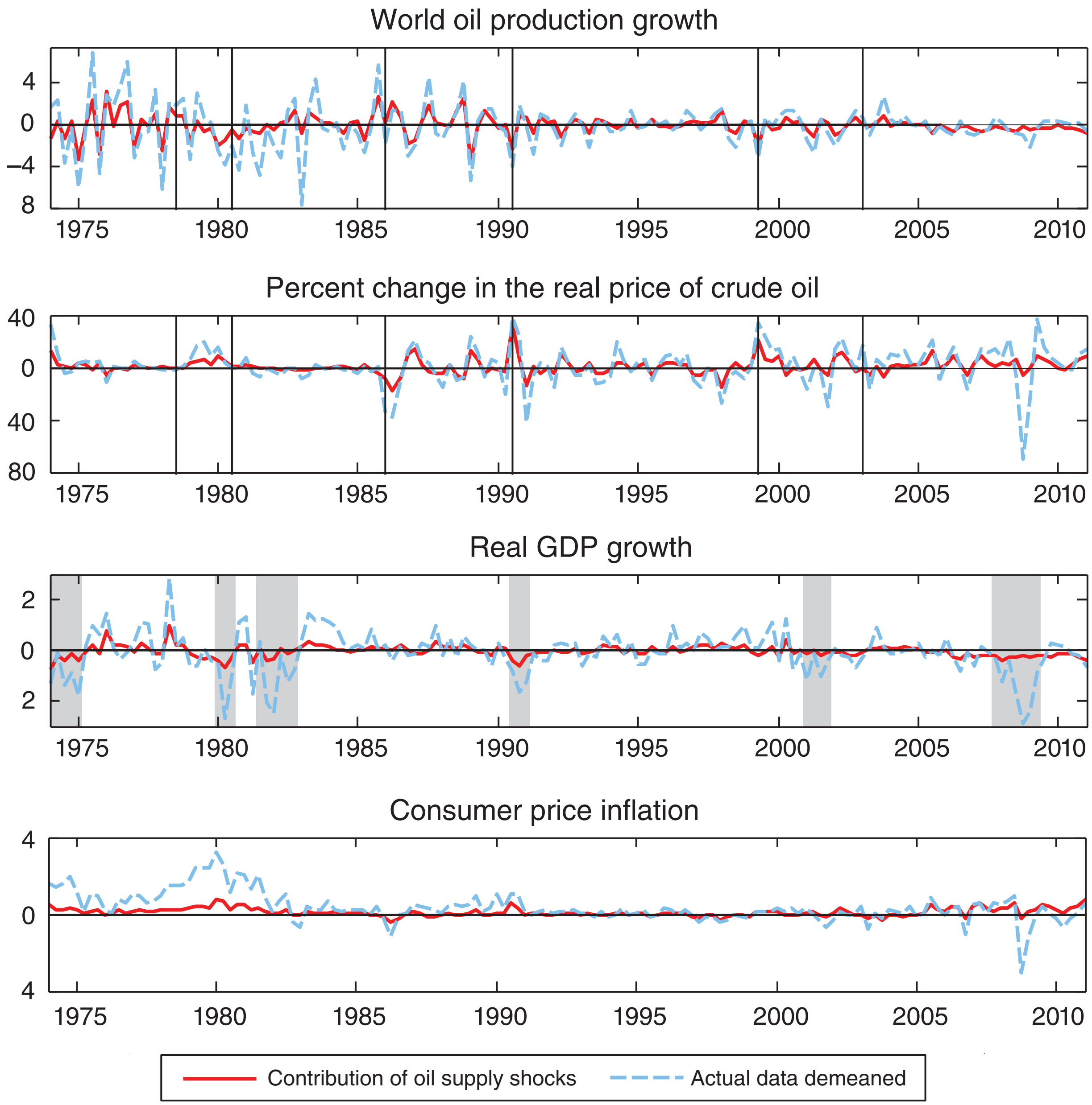
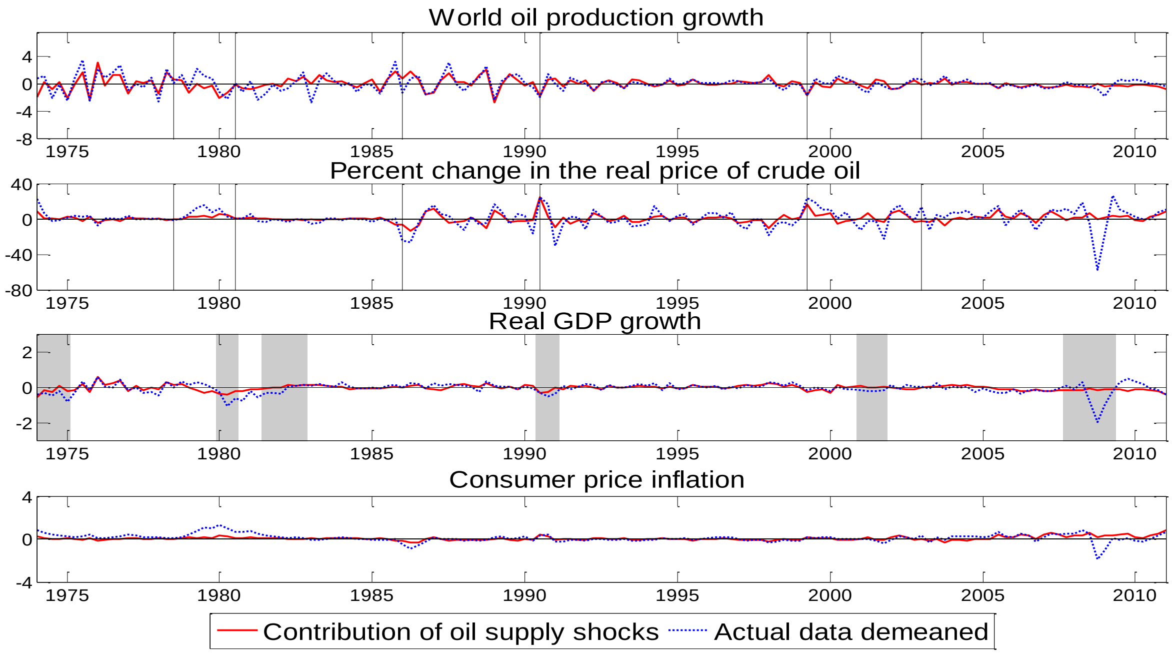


Figure 17: Replication of Baumeister and Peersman (2013) Figure 4



Our replicated decomposition showed less volatility than Baumeister and Peersman (2013), particularly for real GDP growth and consumer price inflation.

7.11 Bhattarai, Lee, and Park (2012), "Monetary-Fiscal Policy Interactions and Indeterminacy in Postwar US Data"

Dependent variables: Output measured as first-difference of the log of the sum of real per capita PCE nondurables, PCE services, government consumption, first-difference of the log of the GDP deflator, Federal Funds Rate, sum of current tax receipts plus contributions for government social insurance divided by output, market value of privately held gross federal debt divided by output.

Key Results: "pre-Volcker, an unanticipated increases in interest rates led to an increase in output and inflation, post-Volcker, it lead to a decrease in output and inflation. Moreover, while pre-Volcker, an unanticipated increase in the (lump-sum) tax revenues-to-output ratio led to a decline in output and inflation, post-Volcker, it had no effects on output or inflation." (pg. 173)

We had neither data or code for this paper, so we were unable to replicate any of the key results. We classified this paper as "failed due to missing public data or code."

7.12 Bianchi (2012), "Evolving Monetary / Fiscal Policy Mix in the United States"

Dependent variables: real GDP growth rate, GDP deflator, Federal Funds Rate, Debt to GDP, tax revenues to GDP.

Key Results: "the monetary/fiscal policy mix has evolved over time" (pg. 167) and "the behavior of the macroeconomy is similar across the three [policy] regimes, even if the behavior of the fiscal variables can differ..." (pg. 170)

We had neither data or code for this paper, so we were unable to replicate any of the key results. We classified this paper as "failed due to missing public data or code."

7.13 Bianchi (2013), "Regime Switches, Agents' Beliefs, and Post-World War II U.S. Macroeconomic Dynamics"

Dependent variables: Federal Funds Rate, real GDP per capita, PCE inflation, real gross investment on durables, real consumption of nondurables goods and services per capita.

Key Results: "If, in the 1970s, agents had anticipated the appointment of an extremely conservative Chairman, inflation would have been lower and the inflation-output trade-off more favourable. The large drop in inflation and output at the end of 2008 would have been mitigated if agents had expected the Federal Reserve to be exceptionally active in the near future." (abstract)

We only had the data but not the code for this paper. Because we were unable to replicate any of the key results, we classified this paper as "failed due to missing public data or code."

7.14 Bilbiie and Straub (2013), "Asset Market Participation, Monetary Policy Rules, and the Great Inflation"

Dependent variables: N/A

Key Results: "a change in the sign of the IS curve slope [starting with Volcker] and "the response of monetary policy changed from passive to active [starting with Volcker]" (introduction, pg. 377)

These key results corresponded to their Tables 2 and 3. Using data and code from R*eview of Economics and Statistics*' website, we were able to qualitatively replicate the results corresponding to the pre-Volcker and Volcker-Greenspan regimes, but our estimates were a bit different than the ones from Bilbiie and Straub (2013). However, our Dynare 4.4.2 program was unable to produce any of the results corresponding to the post-1984 regime after multiple troubleshooting attempts. For these reasons, we classified this paper as "failed due to incorrect public data or code."

Table 14: Bilbiie and Straub (2013) Table 2

Parameter	Pre-Volcker	Volcker-Greenspan	Post-1984
λ	0.50	0.18	0.20
	[0.40, 0.59]	[0.10, 0.25]	[0.14, 0.26]
φ	2.91	2.79	3.07
	[2.03, 3.68]	[1.97, 3.49]	[2.29, 3.86]
γ	0.46	0.50	0.38
	[0.06, 0.85]	[0.38, 0.62]	[0.26, 0.51]
IS-Slope	0.34	-0.60	-0.99
ϕ_π	0.40	1.87	1.63
	[0.23, 0.58]	[1.60, 2.15]	[1.30, 1.93]
ϕ_y	0.41	0.11	0.33
	[0.22, 0.62]	[0.01, 0.20]	[0.15, 0.50]
ϕ_r	0.84	0.64	0.66
	[0.77, 0.90]	[0.56, 0.73]	[0.55, 0.76]
$\pi*$	4.01	3.82	3.25
	[2.41, 5.61]	[2.42, 5.20]	[2.64, 3.87]
$r*$	1.36	2.87	2.44
	[0.58, 2.04]	[2.29, 3.46]	[1.89, 2.96]
ρ^g	0.65	0.86	0.82
	[0.54, 0.76]	[0.79, 0.93]	[0.77, 0.88]
ρ^z	0.66	0.75	0.62
	[0.41, 0.90]	[0.66, 0.84]	[0.51, 0.74]
σ^{ϵ_g}	0.32	0.20	0.21
	[0.20, 0.47]	[0.15, 0.25]	[0.14, 0.28]
σ^{ϵ_z}	0.98	0.88	0.83
	[0.84, 1.14]	[0.73, 1.03]	[0.66, 0.99]
σ^{ϵ_r}	0.17	0.27	0.14
	[0.15, 0.20]	[0.23, 0.32]	[0.13, 0.16]
ρ^{gz}	0.46	0.58	0.91
	[-0.15, 0.98]	[0.41, 0.76]	[0.85, 0.96]

Posterior means with 90% intervals in brackets.

Table 15: Replication of Bilbiie and Straub (2013) Table 2

Parameter	Pre-Volcker	Volcker-Greenspan	Post-1984
λ	0.50	0.17	
	[0.42, 0.58]	[0.08, 0.26]	
φ	2.66	1.77	
	[1.81, 3.50]	[1.21, 2.41]	
γ	0.62	0.42	
	[0.42, 0.82]	[0.30, 0.56]	
IS-Slope	0.24	-0.56	
ϕ_π	0.37	2.12	
	[0.15, 0.59]	[1.72, 2.51]	
ϕ_y	0.53	0.33	
	[0.26, 0.79]	[0.03, 0.64]	
ϕ_r	0.91	0.81	
	[0.86, 0.95]	[0.76, 0.87]	
$\pi*$	3.13	3.13	
	[1.52, 4.66]	[2.39, 3.84]	
$r*$	1.59	1.54	
	[0.73, 2.40]	[1.03, 2.02]	
ρ^g	0.62	0.82	
	[0.51, 0.74]	[0.75, 0.90]	
ρ^z	0.62	0.32	
	[0.48, 0.75]	[0.21, 0.42]	
σ^{ϵ_g}	0.196	0.16	
	[0.14, 0.24]	[0.12, 0.20]	
σ^{ϵ_z}	0.57	1.25	
	[0.46, 0.67]	[1.01, 1.47]	
σ^{ϵ_r}	0.13	0.14	
	[0.13, 0.14]	[0.13, 0.15]	
ρ^{gz}	0.00	0.01	
	[-0.63, 0.64]	[-0.60, 0.64]	

Posterior means with 90% intervals in brackets. The pre-Volcker and Volcker-Greenspan estimates are a bit different, but qualitatively similar to Bilbiie and Straub (2013) as the IS curve sign switches with Volcker-Greenspan. We were unable to replicate the post-1984 estimates.

Table 16: Bilbiie and Straub (2013) Table 3, Original Results and Replication

Sample	Original Log Data Density	Replication of Log Data Density
Pre-Volcker	-349.05	-344.43
Volcker-Greenspan	-363.98	-382.65

Our replication showed a poorer fit of the Bilbiie and Straub (2013) model under determinacy in the Volcker-Greenspan period than reported by the authors.

7.15 Canova and Gambetti (2010), "Do Expectations Matter? The Great Moderation Revisited"

Dependent variables: Federal Funds Rate, Year-to-year changes in real GDP and CPI, Various measures of inflation expectations.

Key Results: "lags of expectations are either always significant or always insignificant [in our VARs], and there is no clear switch over time in their importance" and "reduced-form variances...display similar features and little evidence of time-varying biases." (introduction, pg. 184-185)

These key results corresponded to Tables 5 and 6 in Canova and Gambetti (2010). We were able to replicate both tables, but with a few discrepancies and while supplementing the files from *American Economic Journal: Macroeconomics*' website with our own RATS code using RATS 7.10 (Linux). We altered the replication code in the following four ways: (1) changed the time period of the estimation to match the reported 186 observations in the paper, (2) changed the VAR lags to 4 lags to match the description on page 195 of the paper, (3) wrote code to calculate variances in Table 6, and (4) added code to create tabular output. Tables 18 to 20 show our replication results. In Table 18, we were able to exactly replicate around a quarter of the p-values while the other p-values were slightly different. For Table 20, we were able to exactly replicate around 10% of the variances, and the remainder of the variances differed slightly. For these reasons, we classified this replication as successful.

Table 17: 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	0.73	0.70	0.81	0.91	0.70	0.55	0.99	0.92
π	0.00	0.01	0.01	0.00	0.02	0.00	0.04	0.05
R	0.07	0.00	0.11	0.24	0.00	0.01	0.10	0.05
Panel B. With Term Structure Expectations								
ΔGDP	0.69	0.82	0.52	0.29	0.02	0.03	0.10	0.67
π	0.58	0.51	0.10	0.00	0.00	0.00	0.59	0.24
R	0.00	0.00	0.00	0.00	0.00	0.00	0.06	0.02

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 18: 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. Our replication was similar to the published results.

Table 19: 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	0.80	0.81	0.86	1.06	0.60	0.58	0.56	0.34
π	0.07	0.08	0.09	0.10	0.05	0.04	0.03	0.03
R	0.50	0.75	1.47	1.96	0.93	0.92	0.46	0.15
Panel B. With Term Structure Expectations								
ΔGDP	0.80	0.81	0.83	1.00	0.55	0.53	0.51	0.34
π	0.10	0.10	0.10	0.10	0.04	0.04	0.04	0.03
R	0.43	0.52	1.03	1.35	0.64	0.64	0.46	0.15
Panel C. Without Inflation Expectations								
ΔGDP	0.83	0.83	0.88	1.07	0.62	0.60	0.56	0.35
π	0.10	0.10	0.11	0.13	0.06	0.05	0.04	0.04
R	0.57	0.89	1.65	2.12	1.15	1.06	0.50	0.17

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.

Table 20: 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. Our replication was similar to the published results.

7.16 Carey and Shore (2013), "From the Peaks to the Valleys: Cross-State Evidence on Income Volatility Over the Business Cycle"

Dependent variables: Cross-sectional variance of excess log income.

Key Results: "income volatility is higher in good state times than in bad" (abstract)

We took the key result as Table 4. Using the data and code files from *Review of Economics and Statistics*' website, we replicated their Table 4's point estimates and OLS standard errors almost exactly without assistance using Stata 13.0 (Windows), although we were unable to replicate their bootstrapped standard errors. We classified this replication as successful.

Table 21: 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)
		[0.0067]	[0.0067]	[0.0070]	[0.0071]
	Recent	-0.0036	-0.0014	-0.0005	-0.0021
		(0.0043)	(0.0042)	(0.0042)	(0.0043)
		[0.0078]	[0.0078]	[0.0078]	[0.0078]
βx(Negative National Growth?)	Cumulative	-0.0189	-0.0158	-0.0161	-0.0160
		(0.0038)	(0.0038)	(0.0039)	(0.0040)
		[0.0057]	[0.0057]	[0.0060]	[0.0061]
	Recent	-0.0124	-0.0103	-0.0114	-0.0113
		(0.0045)	(0.0044)	(0.0046)	(0.0046)
		[0.0040]	[0.0041]	[0.0042]	[0.0043]
βx(% of Year in Recession)	Cumulative	-0.0365	-0.0321	-0.0350	-0.0330
		(0.0050)	(0.0050)	(0.0051)	(0.0051)
		[0.0088]	[0.0089]	[0.0092]	[0.0093]
	Recent	-0.0082	-0.0057	-0.0032	-0.0028
		(0.0078)	(0.0077)	(0.0077)	(0.0080)
		[0.0099]	[0.0098]	[0.0099]	[0.0102]
βx(National GDP growth)	Cumulative	-0.1133	-0.1439	-0.1935	-0.1593
		(0.0818)	(0.0813)	(0.0830)	(0.0861)
		[0.0768]	[0.0771]	[0.0803]	[0.0824]
	Recent	0.0346	0.0185	-0.0431	-0.0248
		(0.0802)	(0.0800)	(0.0814)	(0.0829)
		[0.090]	[0.0893]	[0.0915]	[0.0930]
βx(Demeaned squared growth)	Cumulative	-11.9832	-10.7351	-11.0769	-10.5451
		(1.5012)	(1.4995)	(1.5206)	(1.5496)
		[2.4095]	[2.420]	[2.4992]	[2.5691]
	Recent	-15.4747	-14.3463	-15.0460	-14.3461
		(2.0533)	(2.0411)	(2.0654)	(2.0984)
		[2.7565]	[2.7823]	[2.8527]	[2.9320]

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

Table 22: 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 OLS standard errors in parentheses, with the dependent variable as cross-sectional income variance. See text of Carey and Shore (2013) for additional details. We were unable to replicate Carey and Shore (2013)'s bootstrapped standard errors, but our point estimates and OLS standard errors were very close to their reported values.

7.17 Castelnuovo and Surico (2010), "Monetary Policy, Inflation Expectations and the Price Puzzle"

Dependent variables: log-deviation of real GDP with respect to CBO's potential output, GDP Deflator, Federal Funds Rate.

Key Results: "[in our VARs] the positive response of prices to a monetary policy shock is historically limited to the sub-samples that are typically associated with a weak interest rate response to inflation." (abstract)

We only had the code but not the data for this paper. Because we were unable to replicate any of the key results, we classified this paper as "failed due to missing public data or code."

7.18 Chen, Curdia, and Ferrero (2012), "The Macroeconomic Effects of Large-scale Asset Purchase Programmes"

Dependent variables: Federal Funds Rate, Term Premium, 10-year yield, log-difference of core PCE price index, log-difference real GDP per capita, level of real GDP per capita.

Key Results: "GDP growth increases by less than a third of a percentage point and inflation barely changes relative to the absence of intervention." (abstract)

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

7.19 Clark and McCracken (2010), "Averaging Forecasts from VARs with Uncertain Instabilities"

Dependent variables: Output, Inflation, Short-Term Interest Rate (various measures for all three).

Key Results: "The best forecast is a simple average of projections from a univariate model and a VAR using detrended inflation and interest rates. At the other extreme, forecasts based on OLS-type combination and factor model-based combination rank among the worst." (introduction, pg. 6)

We took the key results as the panels for GDP growth in their Tables 3 to 5. We used the data and code from the *Journal of Applied Econometrics* data archives. In their Table 3, the relative root mean squared errors (RMSEs) of the forecasting models that the authors considered worsened 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, came from their Bayesian vector autoregressions or first-differenced vector autoregressions. Similar results hold in their Tables 4 and 5.

We were able to replicate these results almost exactly with RATS 7.10 (Linux), with the exception of the rows in these tables with Bayesian model averaging, which were approximately the same as in the published paper. We display our replication results where we did not find an exact match in Tables 24, 26, and 28. This replication was classified as successful.

Table 23: Clark and McCracken (2010) Table 3

Forecast Method	Sample Period					
	1970-1984			1985-2005		
	$h = 0Q$	$h = 1Q$	$h = 1Y$	$h = 0Q$	$h = 1Q$	$h = 1Y$
BMA: AIC	1.008	0.966	0.901	1.120	1.106	0.978
BMA: BIC	0.946	0.909	0.964	1.047	1.038	0.8998
BMA: PIC	0.902	0.837	0.849	1.100	1.087	0.926

Rows displayed are those from the GDP growth forecast panel of Table 3 of Clark and McCracken (2010) where we could not match the published results exactly. Our replication was very close to Clark and McCracken (2010).

Table 24: Replication of Clark and McCracken (2010) Table 3

Forecast Method	Sample Period					
	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 growth forecast panel of Table 3 of Clark and McCracken (2010) where we could not match the published results exactly.

Table 25: 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	0.879	0.914	1.154	1.190	1.147
BMA: BIC	0.959	0.867	0.887	1.055	1.077	1.024
BMA: PIC	0.873	0.796	0.812	1.120	1.148	1.132

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

Table 26: 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 growth forecast panel of Table 4 of Clark and McCracken (2010) where we could not match the published results exactly. Our replication was very close to Clark and McCracken (2010).

Table 27: 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	1.007	1.054	1.259	1.434	1.332	1.138
BMA: BIC	0.958	1.003	1.038	1.316	1.194	0.995
BMA: PIC	1.042	1.053	1.139	1.332	1.250	1.067

Root mean squared ratios of different forecasting methods, relative to a univariate benchmark. BMA = Bayesian Model Average. Rows displayed are those from Table 5 of Clark and McCracken (2010) where we could not match the published results exactly.

Table 28: 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

Root mean squared (RMSE) ratios of different forecasting methods, relative to a univariate benchmark. BMA = Bayesian Model Average. Rows displayed are those from Table 5 of Clark and McCracken (2010) where we could not match the published results exactly. Our replicated BMA: PIC RMSE ratios were lower than those reported by Clark and McCracken (2010), while our replicated BMA: AIC and BMA: BIC RMSE ratios were similar to Clark and McCracken (2010).

7.20 Clements and Galvão (2009), "Forecasting US Output Growth Using Leading Indicators: An Appraisal Using MIDAS Models"

Dependent variables: GDP Growth, Using End-of-Sample Vintages, Real-Time Vintages, and Final-Vintage Data (Koenig, Dolmas, and Piger, 2003).

Key Results: "MIDAS is a useful vehicle for combining a small group of indicators for forecasting," "information on the current quarter improves forecasts," "combination in modelling with MIDAS is better than combination of forecasts when predicting the direction of change of output growth," "real-time vintage data... improves forecast performance," and "evidence of the predictive ability of the indicators... is stronger when the aim is to forecast final data" (introduction, pg 1188).

We only had the data but not the code for this paper, so we were unable to replicate any of the key results. We classified this paper as "failed due to missing public data or code."

7.21 Clements and Galvão (2013), "Real-Time Forecasting of Inflation and Output Growth with Autoregressive Models in the Presence of Data Revisions"

Dependent variables: GDP growth, PCE deflator, GDP deflator, Using End-of-Sample Vintages, Real-Time Vintages (Koenig, Dolmas, and Piger, 2003).

Key Results: "we find improvements in root mean square forecasting error of 2-4% when forecasting output growth and inflation with univariate models, and of 8% with multivariate models [using 'lightly revised' data]" and "multiple-vintage models...require large estimation samples to deliver competitive forecasts." (abstract)

We took the key results of Clements and Galvão (2013) as Tables IV and V. The data archive from *Journal of Applied Econometrics* only contained the data series used in their paper. We used code and data provided by Ana Galvão through personal communication.

We were able to exactly replicate a majority of the qualitative results with help from the authors, but there were some differences. In particular, we had some initial issues replicating the results from their Table V using the 1979 Q1 data for the KK model and to some extent the ADL model. After discussions with Ana Galvão, we were advised to change the sample slightly to add more observations for some models to produce the results in the following tables. Otherwise, using the original settings that came with the code, the code would either fail to display any results or it would display an error message regarding the positive definiteness of a matrix; the failed estimates are marked as "N/A" in the tables. With this change, we were able to qualitatively replicate a vast majority of the remaining entries, although we still failed to obtain a few estimates. For these reasons, we classified this replication as successful.

Table 29: Clements and Galvão (2013) Table IV

(A) Forecasting output growth

	$h=1$			$h=4$		
Model	y_{T+1}^{T+1+1}	y_{T+1}^{T+15}	y_{T+1}^{09Q1}	y_{T+4}^{T+4+1}	y_{T+4}^{T+4+14}	y_{T+4}^{09Q1}
AR(4)	0.998	1.008	1.001	0.970	0.978	0.975
VB-VAR, $q=5$	0.995	0.986	0.994	0.967	0.983	0.984
VB-VAR, $q=14$	1.044	1.028	0.999	1.017	1.025	1.001
KK, $q=5$, $p=1$	1.040	1.052	1.034	0.976	0.975	0.982
KK, $q=14$, $p=1$	1.065	1.085	1.056	0.972	0.972	0.972

(B) Forecasting GDP deflator inflation

	$h=1$			$h=4$		
Model	y_{T+1}^{T+1+1}	y_{T+1}^{T+15}	y_{T+1}^{09Q1}	y_{T+4}^{T+4+1}	y_{T+4}^{T+4+14}	y_{T+4}^{09Q1}
AR(4)	0.962	0.992	1.018	0.956	1.034	1.035
VB-VAR, $q=5$	0.988	0.987	0.993	1.034	1.074	1.070
VB-VAR, $q=14$	0.955	0.987	1.011	0.973	1.058	1.056
KK, $q=5$, $p=4$	0.997	1.008	1.026	1.057	1.096	1.098
KK, $q=14$, $p=4$	1.047	1.056	1.147	1.221	1.338	1.334

(C) Forecasting PCE deflator inflation

	$h=1$			$h=4$		
Model	y_{T+1}^{T+1+1}	y_{T+1}^{T+15}	y_{T+1}^{09Q1}	y_{T+4}^{T+4+1}	y_{T+4}^{T+4+14}	y_{T+4}^{09Q1}
AR(4)	0.974	0.992	0.996	0.964	0.990	0.992
VB-VAR, $q=5$	1.000	1.012	1.009	1.017	1.027	1.022
VB-VAR, $q=14$	1.068	1.063	1.071	1.054	1.056	1.052
KK, $q=5$, $p=4$	1.001	1.011	1.009	1.003	1.014	1.015
KK, $q=14$, $p=4$	1.008	1.018	1.039	1.113	1.157	1.170

Table 30: Replication of Clements and Galvão (2013) Table IV

(A) Forecasting output growth

	$h=1$			$h=4$		
Model	y_{T+1}^{T+1+1}	y_{T+1}^{T+15}	y_{T+1}^{09Q1}	y_{T+4}^{T+4+1}	y_{T+4}^{T+4+14}	y_{T+4}^{09Q1}
AR(4)	0.999	1.008	1.002	0.970	0.978	0.975
VB-VAR, $q=5$	0.994	0.986	0.994	0.967	0.969	0.976
VB-VAR, $q=14$	1.044	1.028	0.999	1.017	0.994	0.979
KK, $q=5$, $p=1$	1.040	1.053	1.034	0.976	0.975	0.982
KK, $q=14$, $p=1$	1.065	1.085	1.056	0.972	0.972	0.972

(B) Forecasting GDP deflator inflation

	$h=1$			$h=4$		
Model	y_{T+1}^{T+1+1}	y_{T+1}^{T+15}	y_{T+1}^{09Q1}	y_{T+4}^{T+4+1}	y_{T+4}^{T+4+14}	y_{T+4}^{09Q1}
AR(4)	0.962	0.992	1.018	0.957	1.034	1.035
VB-VAR, $q=5$	0.988	0.987	0.993	1.034	1.074	1.070
VB-VAR, $q=14$	0.955	0.987	1.011	0.973	1.058	1.056
KK, $q=5$, $p=4$	0.997	1.008	1.026	1.057	1.096	1.098
KK, $q=14$, $p=4$	1.047	1.056	1.147	1.221	1.338	1.334

(C) Forecasting PCE deflator inflation

	$h=1$			$h=4$		
Model	y_{T+1}^{T+1+1}	y_{T+1}^{T+15}	y_{T+1}^{09Q1}	y_{T+4}^{T+4+1}	y_{T+4}^{T+4+14}	y_{T+4}^{09Q1}
AR(4)	0.974	0.992	0.996	0.964	0.990	0.992
VB-VAR, $q=5$	1.000	1.012	1.009	1.017	1.027	1.022
VB-VAR, $q=14$	1.068	1.063	1.071	1.054	1.056	1.052
KK, $q=5$, $p=4$	1.001	1.011	1.009	1.003	1.014	1.015
KK, $q=14$, $p=4$	1.008	1.018	1.039	1.123	1.157	1.170

Our replication results were close to Clements and Galvão (2013).

Table 31: Clements and Galvão (2013) Table V, Panel A

(A) Forecasting output growth													
		$h=0$				$h=1$				$h=4$			
					Ratio AR(1) _RTV				Ratio AR(1) _RTV				Ratio AR(1) _RTV
Model	Sample	y_{T+1}^{T+2}	y_{T+1}^{T+15}	y_{T+1}^{09Q1}	y_{T+1}^{09Q1}	y_{T+1}^{T+2}	y_{T+1}^{T+15}	y_{T+1}^{09Q1}	y_{T+1}^{09Q1}	y_{T+4}^{T+5}	y_{T+4}^{T+19}	y_{T+4}^{09Q1}	y_{T+4}^{09Q1}
	starts:	With Industrial production											
ADL(1,2)	1965:Q3	1.057	1.033	0.976	0.829	1.020	1.023	0.989	0.956	0.988	0.997	1.013	1.056
KK($q=5$, $p=(1,1)$)		1.223	1.119	1.060	0.902	1.030	1.019	1.003	0.969	0.963	0.956	0.959	0.999
ADL(1,2)	1979:Q1	0.933	0.947	0.893	0.826	0.973	0.980	0.952	0.958	0.944	0.982	1.000	1.054
KK($q=5$, $p=(1,1)$)		1.141	1.042	0.979	0.908	1.247	1.118	1.125	1.136	0.961	0.958	0.967	1.005
	starts:	With employment											
ADL(1,2)	1965:Q3	0.968	0.983	0.933	0.845	0.995	1.047	0.995	1.006	1.046	1.067	1.073	1.132
KK($q=5$, $p=(1,1)$)		1.043	1.034	0.997	0.902	1.015	1.068	0.999	1.007	0.931	0.934	0.948	1.000
ADL(1,2)	1979:Q1	0.965	0.984	0.941	0.853	0.932	0.986	0.948	1.004	0.954	0.996	1.007	1.052
KK($q=5$, $p=(1,1)$)		1.684	1.299	1.385	1.256	1.005	1.040	0.956	1.012	1.123	1.026	1.018	1.064

Table 32: Replication of Clements and Galvão (2013) Table V, Panel A

(A) Forecasting output growth													
		$h=0$				$h=1$				$h=4$			
					Ratio AR(1) _RTV				Ratio AR(1) _RTV				Ratio AR(1) _RTV
Model	Sample	y_{T+1}^{T+2}	y_{T+1}^{T+15}	y_{T+1}^{09Q1}	y_{T+1}^{09Q1}	y_{T+1}^{T+2}	y_{T+1}^{T+15}	y_{T+1}^{09Q1}	y_{T+1}^{09Q1}	y_{T+4}^{T+5}	y_{T+4}^{T+19}	y_{T+4}^{09Q1}	y_{T+4}^{09Q1}
	starts:	With Industrial production											
ADL(1,2)	1965:Q3	1.056	1.032	0.976	0.829	1.020	1.020	0.992	0.956	0.987	0.996	1.012	1.055
KK($q=5$, $p=(1,1)$)		1.220	1.117	1.060	0.900	1.028	1.020	1.006	0.970	0.963	0.956	0.958	0.999
ADL(1,2)	1979:Q1	0.937	0.952	0.899	0.829	0.973	0.980	0.954	0.963	0.946	0.975	0.999	1.039
KK($q=5$, $p=(1,1)$)		2.434	1.077	1.008	0.929	253.354	214.973	204.289	206.136	0.957	0.955	0.967	1.006
	starts:	With employment											
ADL(1,2)	1965:Q3	0.969	0.984	0.935	0.845	0.997	1.045	1.001	1.004	1.038	1.060	1.064	1.124
KK($q=5$, $p=(1,1)$)		1.039	1.031	0.996	0.900	1.019	1.073	1.005	1.008	0.930	0.933	0.946	0.999
ADL(1,2)	1979:Q1	0.947	0.969	0.929	0.845	0.920	0.976	0.938	0.996	0.980	1.016	1.019	1.100
KK($q=5$, $p=(1,1)$)		1.362	1.086	1.157	1.052	0.999	1.093	1.031	1.094	0.910	0.936	0.932	1.007

We were able to closely match most of the results from Clements and Galvão (2013), except for some of the KK model results for $h=1$ and $h=4$ horizons and 1979:Q1 sample.

Table 33: Clements and Galvão (2013) Table V, Panel B

(B) Forecasting GDP deflator inflation													
		$h = 0$				$h = 1$				$h = 4$			
					Ratio AR(4) _RTV				Ratio AR(4) _RTV				Ratio AR(4) _RTV
Model	Sample	y_{T+1}^{T+2}	y_{T+1}^{T+15}	y_{T+1}^{09Q1}	y_{T+1}^{09Q1}	y_{T+1}^{T+2}	y_{T+1}^{T+15}	y_{T+1}^{09Q1}	y_{T+1}^{09Q1}	y_{T+4}^{T+5}	y_{T+4}^{T+19}	y_{T+4}^{09Q1}	y_{T+4}^{09Q1}
	starts:	With Industrial production											
ADL(4,2)	1965:Q3	0.970	0.980	1.007	0.967	0.972	0.969	0.978	0.970	1.064	1.094	1.094	1.074
KK($q = 5$, $p = (4, 1)$)		1.099	1.023	1.059	1.016	1.083	0.994	1.030	1.022	1.078	1.079	1.100	1.080
ADL(4,2)	1979:Q1	0.940	0.911	0.932	0.997	0.935	0.906	0.920	0.996	0.984	0.999	0.977	1.036
KK($q = 5$, $p = (4, 1)$)		1.963	1.281	1.406	1.496	1.096	0.948	0.969	1.043	0.982	0.931	0.927	1.002
	starts:	With employment											
ADL(4,2)	1965:Q3	0.968	0.971	0.981	0.953	0.972	0.974	0.977	0.978	1.047	1.108	1.097	1.154
KK($q = 5$, $p = (4, 1)$)		1.025	0.960	0.996	0.968	1.048	0.986	0.999	1.005	1.049	0.994	1.034	1.088
ADL(4,2)	1979:Q1	0.939	0.896	0.920	0.991	0.939	0.906	0.916	0.991	0.981	0.981	0.964	1.057
KK($q = 5$, $p = (4, 1)$)		1.321	1.042	1.018	1.110	2.015	1.138	1.231	1.342	1.106	1.048	1.119	1.222

Table 34: Replication of Clements and Galvão (2013) Table V, Panel B

(B) Forecasting GDP deflator inflation													
		$h = 0$				$h = 1$				$h = 4$			
					Ratio AR(4) _RTV				Ratio AR(4) _RTV				Ratio AR(4) _RTV
Model	Sample	y_{T+1}^{T+2}	y_{T+1}^{T+15}	y_{T+1}^{09Q1}	y_{T+1}^{09Q1}	y_{T+1}^{T+2}	y_{T+1}^{T+15}	y_{T+1}^{09Q1}	y_{T+1}^{09Q1}	y_{T+4}^{T+5}	y_{T+4}^{T+19}	y_{T+4}^{09Q1}	y_{T+4}^{09Q1}
	starts:	With Industrial production											
ADL(4,2)	1965:Q3	0.970	0.980	1.007	0.967	0.972	0.970	0.978	0.970	1.064	1.094	1.094	1.074
KK($q = 5$, $p = (4, 1)$)		1.099	1.023	1.059	1.016	1.083	0.994	1.030	1.022	1.078	1.079	1.100	1.080
ADL(4,2)	1979:Q1	0.945	0.921	0.943	1.007	0.937	0.904	0.920	0.998	N/A	N/A	N/A	N/A
KK($q = 5$, $p = (4, 1)$)		2.558	1.425	1.569	1.680	1.073	0.963	0.993	1.075	N/A	N/A	N/A	N/A
	starts:	With employment											
ADL(4,2)	1965:Q3	0.968	0.971	0.981	0.953	0.972	0.974	0.977	0.978	1.047	1.050	1.060	1.061
KK($q = 5$, $p = (4, 1)$)		1.025	0.960	0.996	0.968	1.048	0.986	1.000	1.001	1.050	0.994	1.034	1.036
ADL(4,2)	1979:Q1	0.939	0.898	0.921	0.997	0.939	0.904	0.915	0.996	0.966	0.977	0.953	1.062
KK($q = 5$, $p = (4, 1)$)		1.265	0.982	0.978	1.059	7.894	2.656	2.878	3.133	1.084	1.009	1.017	1.134

We were able to closely match most of the results from Clements and Galvão (2013), except we could not obtain any estimates under the ADL and KK models for $h = 4$, the 1979:Q1 sample, and with industrial production.

Table 35: Clements and Galvão (2013) Table V, Panel C

(C) Forecasting PCE deflator inflation													
		$h=0$				$h=1$				$h=4$			
					Ratio AR(4) _RTV				Ratio AR(4) _RTV				Ratio AR(4) _RTV
Model	Sample	y_{T+1}^{T+2}	y_{T+1}^{T+15}	y_{T+1}^{09Q1}	y_{T+1}^{09Q1}	y_{T+1}^{T+2}	y_{T+1}^{T+15}	y_{T+1}^{09Q1}	y_{T+1}^{09Q1}	y_{T+4}^{T+5}	y_{T+4}^{T+19}	y_{T+4}^{09Q1}	y_{T+4}^{09Q1}
	starts:	With Industrial production											
ADL(4,2)	1965:Q3	0.985	0.994	1.004	0.951	0.974	0.988	0.998	0.955	1.028	1.062	1.070	1.034
KK($q=5$, $p=(4,1)$)		1.058	1.056	1.074	1.017	1.032	1.031	1.043	1.016	1.042	1.042	1.060	1.053
ADL(4,2)	1979:Q1	0.977	0.983	0.981	0.986	0.972	0.981	0.969	0.993	0.978	0.979	0.968	1.029
KK($q=5$, $p=(4,1)$)		1.414	1.071	1.086	1.091	0.971	0.977	0.982	1.006	0.938	0.934	0.939	0.997
	starts:	With employment											
ADL(4,2)	1965:Q3	0.978	0.980	0.982	0.934	0.970	0.978	0.975	0.972	1.028	1.036	1.039	0.980
KK($q=5$, $p=(4,1)$)		0.996	0.989	1.001	0.952	1.003	0.998	1.008	1.001	1.046	1.039	1.058	0.998
ADL(4,2)	1979:Q1	0.997	0.997	1.005	0.999	0.981	0.988	0.978	1.006	0.975	0.987	0.970	1.007
KK($q=5$, $p=(4,1)$)		0.970	0.953	0.947	0.941	1.064	1.002	1.002	1.046	0.944	0.956	0.953	0.998

Table 36: Replication of Clements and Galvão (2013) Table V, Panel C

(C) Forecasting PCE deflator inflation													
		$h = 0$				$h = 1$				$h = 4$			
					Ratio AR(4) _RTV				Ratio AR(4) _RTV				Ratio AR(4) _RTV
Model	Sample	y_{T+1}^{T+2}	y_{T+1}^{T+15}	y_{T+1}^{09Q1}	y_{T+1}^{09Q1}	y_{T+1}^{T+2}	y_{T+1}^{T+15}	y_{T+1}^{09Q1}	y_{T+1}^{09Q1}	y_{T+4}^{T+5}	y_{T+4}^{T+19}	y_{T+4}^{09Q1}	y_{T+4}^{09Q1}
	starts:	With Industrial production											
ADL(4,2)	1965:Q3	0.985	0.994	1.004	0.951	0.974	0.979	0.981	0.955	1.028	1.038	1.041	1.034
KK($q = 5$, $p = (4, 1)$)		1.058	1.056	1.074	1.017	1.032	1.031	1.043	1.016	1.042	1.042	1.060	1.053
ADL(4,2)	1979:Q1	0.994	1.002	1.005	0.989	0.993	1.003	0.994	0.998	0.990	0.997	0.985	1.043
KK($q = 5$, $p = (4, 1)$)		1.414	1.071	1.086	1.091	0.971	0.977	0.982	1.006	0.933	0.924	0.932	0.987
	starts:	With employment											
ADL(4,2)	1965:Q3	0.978	0.980	0.982	0.934	0.970	0.976	0.974	0.972	1.028	1.036	1.039	0.980
KK($q = 5$, $p = (4, 1)$)		0.996	0.990	1.001	0.952	1.003	0.998	1.008	1.006	1.046	1.039	1.058	0.998
ADL(4,2)	1979:Q1	0.997	0.997	1.005	0.999	0.981	0.988	0.978	1.006	0.975	0.987	0.970	1.007
KK($q = 5$, $p = (4, 1)$)		0.979	0.953	0.947	0.941	1.258	1.200	1.188	1.222	0.982	0.984	0.989	1.026

We were able to closely match most of the results from Clements and Galvão (2013).

7.22 Corsetti, Meier, and Müller (2012), "Fiscal Stimulus with Spending Reversals"

Dependent variables: Seven variable VAR that rotates the seventh variable between three different series. GDP as a scaling factor and a direct input as ln(per capita real GDP).

Key Results: "an exogenous increase in government spending causes a substantial rise in aggregate output," and "a positive spending shock triggers a sizable buildup of public debt, followed over time by a decline of government spending below trend" (introduction, pg. 878)

We took the key result to be the panels for output and debt in their Figures 1 and 2, as debt was scaled by GDP. We were able to replicate these panels exactly with assistance from the authors using code from *Review of Economics and Statistics*' website in Matlab R2008a (Windows). We classified this replication as successful.

7.23 D'Agostino and Surico (2012), "A Century of Inflation Forecasts"

Dependent variables: Log difference of GDP deflator.

Key Results: "output growth had marginal predictive power for inflation only during times in which... the monetary authorities did not succeed in establishing a clear nominal anchor or an inflation fighter reputation." (introduction, pg. 1097)

We took 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 *Review of Economics and Statistics*' website and were able to replicate these figures exactly without assistance using Matlab R2010a (Linux). We classified this replication as successful.

7.24 Del Negro and Schorfheide (2009), "Monetary Policy Analysis With Potentially Misspecified Models"

Dependent variables: log difference of real GDP and the CBO's real potential output, log difference of GDP deflator, Federal Funds Rate.

Key Results: No single key result, so we tried to replicate everything except Figure 1.

We obtained the data and code from the *American Economic Review's* website. Even with assistance from Frank Schorfheide, we could not get any of the Matlab code to run without crashing. The main problem was due to missing Matlab dependencies, either in terms of functions or matrix objects. Frank Schorfheide advised us to first run some code files multiple times to generate required inputs, but despite having a manifest of the data and code files, the order of code execution was unclear to produce the desired results. For these reasons, we classified this paper as "failed due to incorrect public data or code."

7.25 Den Haan and Sterk (2011), "The Myth of Financial Innovation and the Great Moderation"

Dependent variables: Federal Funds Rate, HP-Filtered or natural log of: GDP deflator, Residential Investment, Durable Expenditures, GDP, Home Mortgages, Consumer Credit.

Key Results: "the responses of real activity and consumer loans to several shocks have remained remarkably stable over time. The drop in the co-movement is due to changes in the IRFs of the monetary policy shock and the real activity shocks," and "responses of both the real activity and the loan variables switch sign following a real activity shock" (introduction, pg. 708)

We took the key figures as Figures 3, 4, 6, and 7. 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. Our replication results were close to the published version using Matlab R2010a (Linux). Note that our versions of Figures 4 and 6 from Den Haan and Sterk (2011), produced with the authors' original code, have an extra column of figures corresponding to the time period 1954Q3-2008Q1 relative to the published paper. We classified this replication as successful.

Figure 18: Den Haan and Sterk (2011) Figure 3

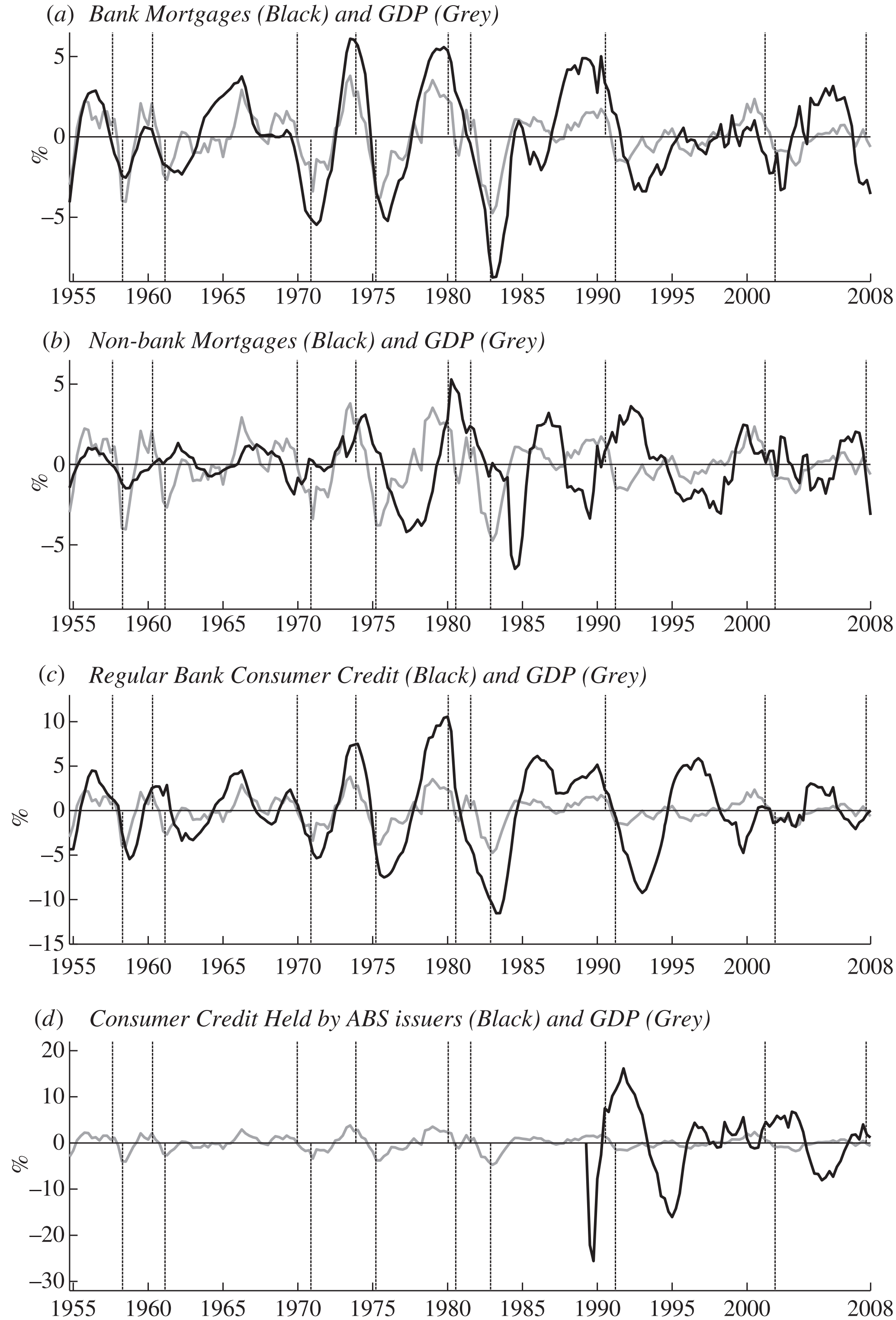
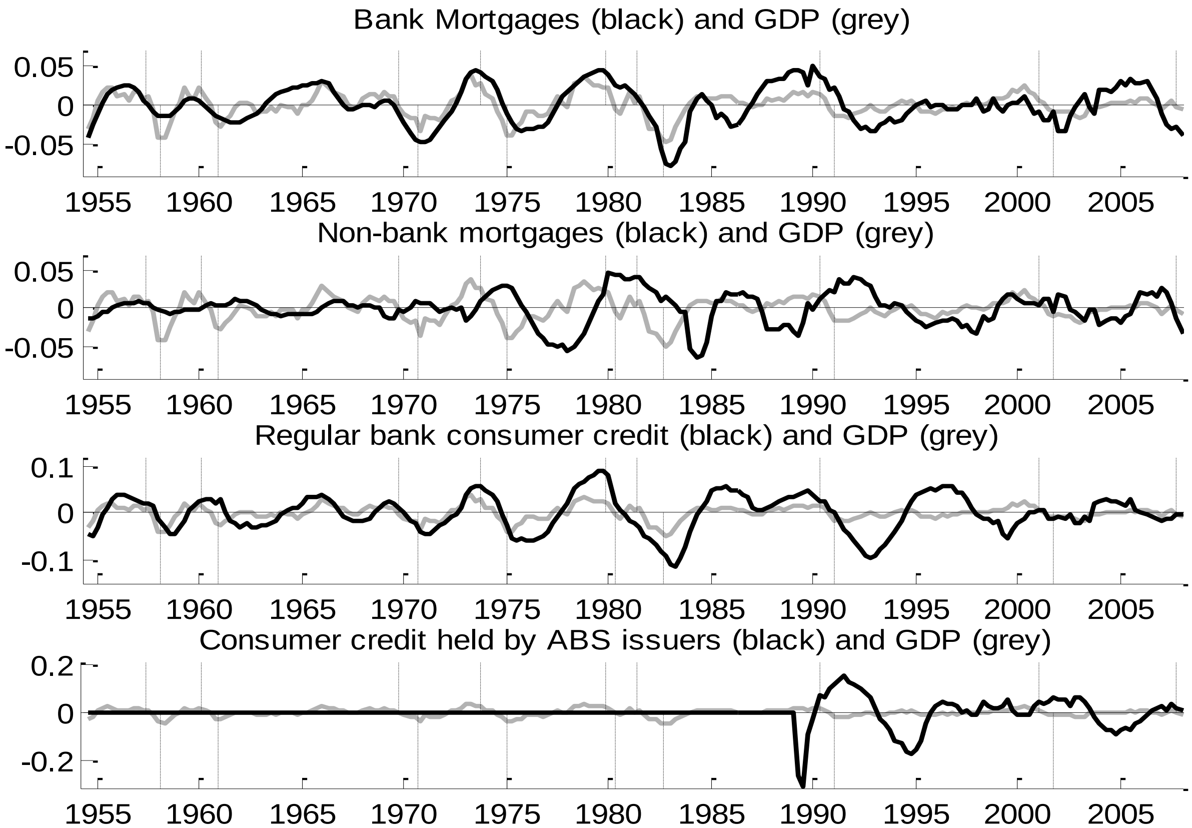


Figure 19: Replication of Den Haan and Sterk (2011) Figure 3



Our replication matches the published figure quite closely.

Figure 20: Den Haan and Sterk (2011) Figure 4

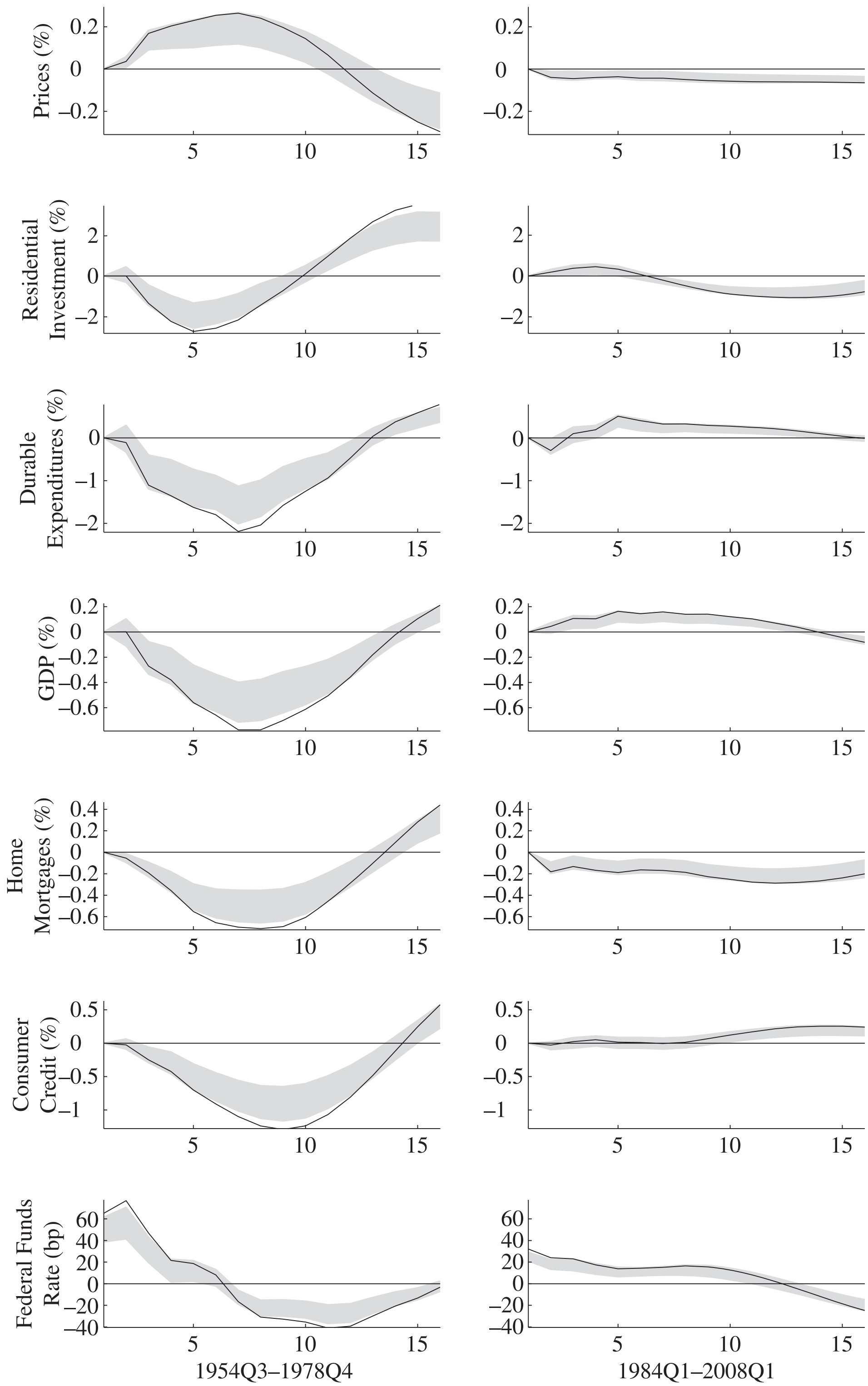
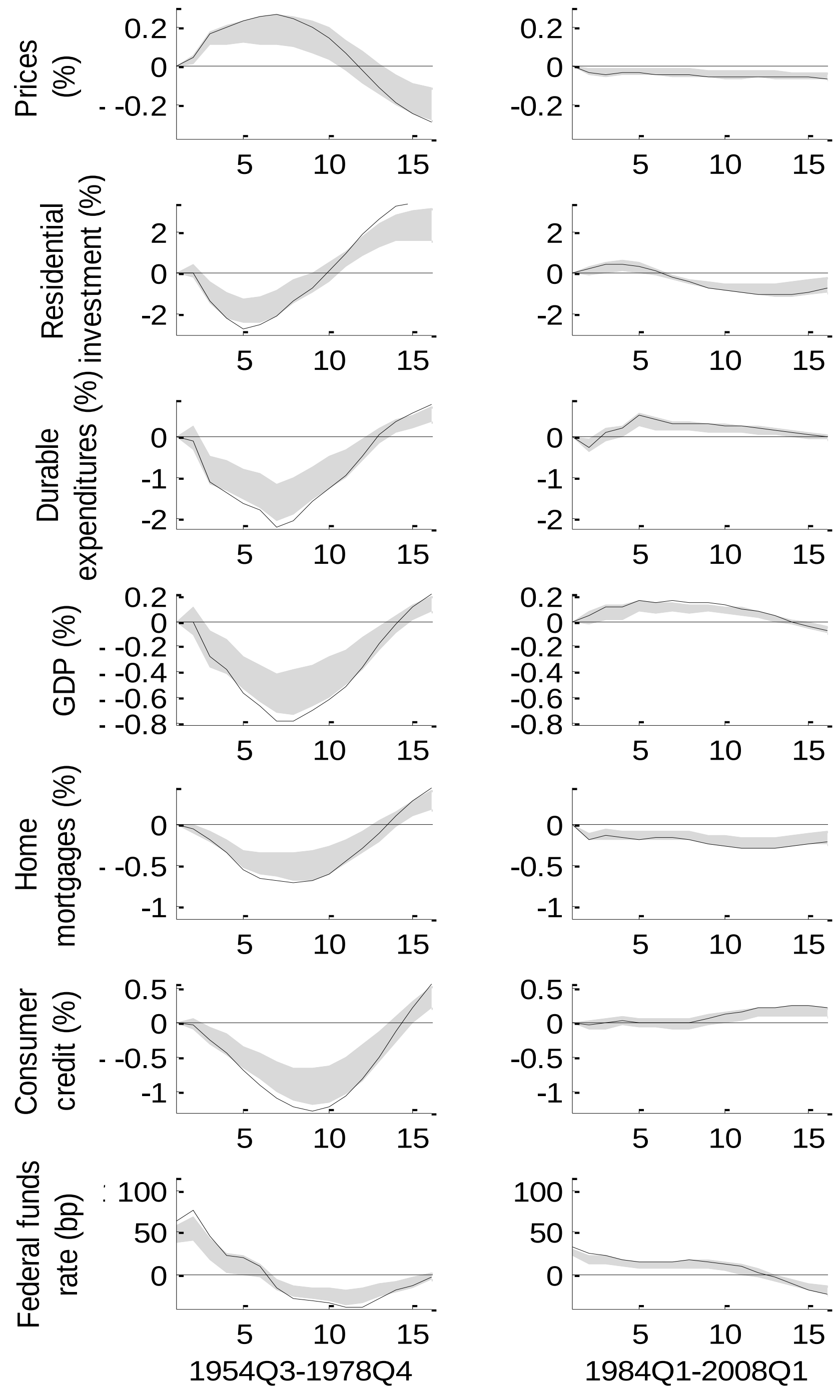


Figure 21: Replication of Den Haan and Sterk (2011) Figure 4



Our replication was very close to the published figure.

Figure 22: Den Haan and Sterk (2011) Figure 6

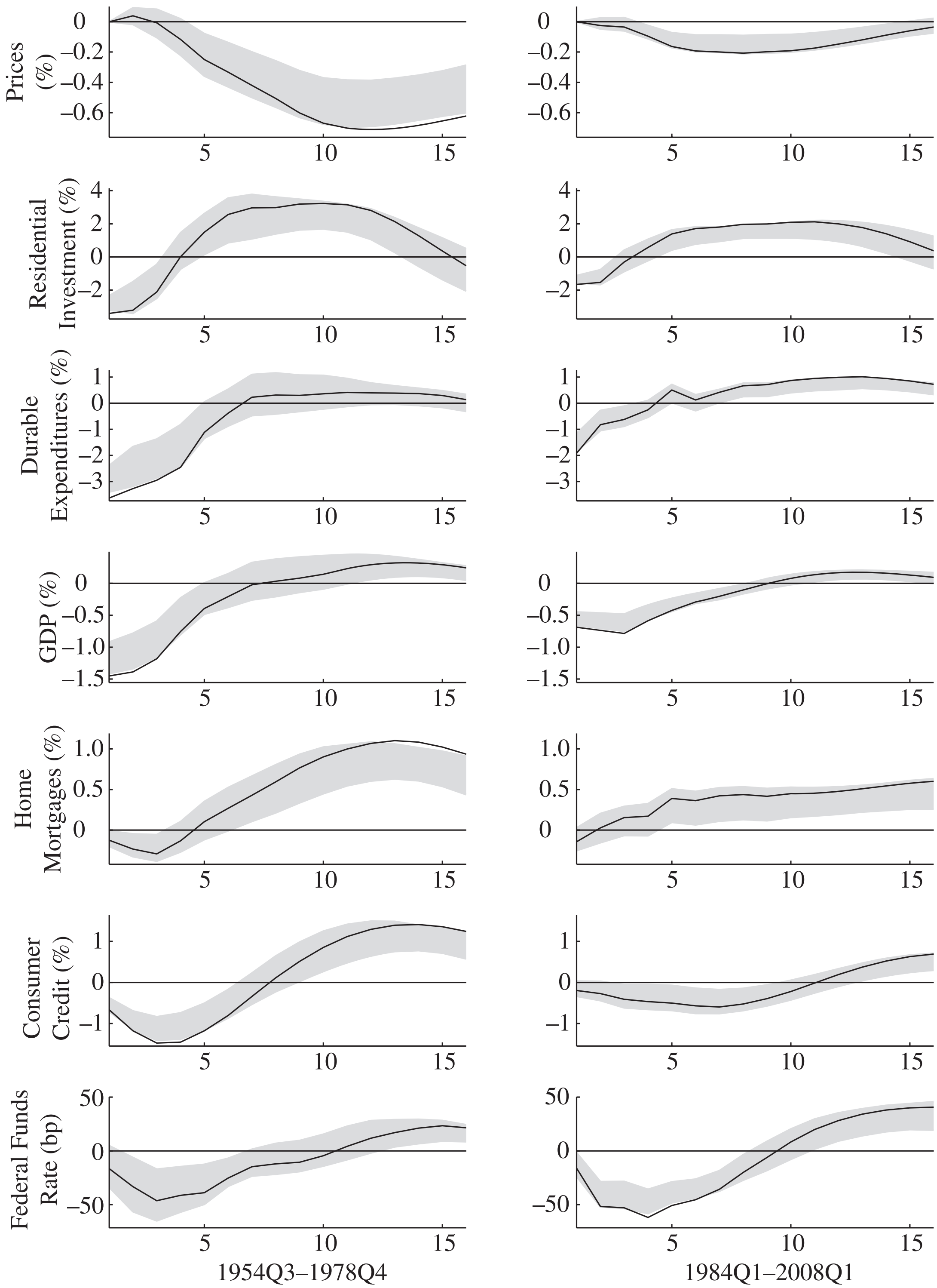
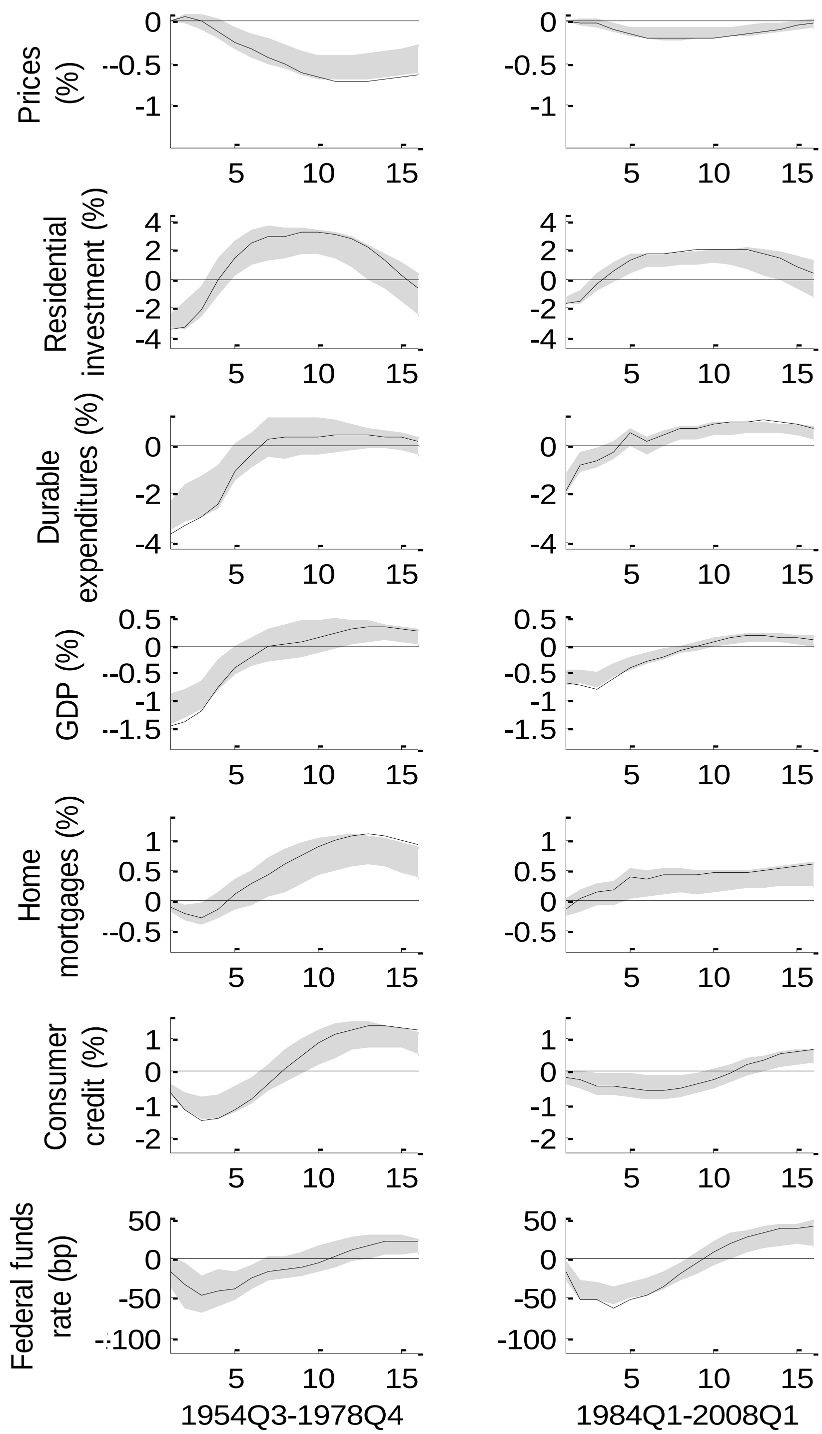


Figure 23: Replication of Den Haan and Sterk (2011) Figure 6



Our replication was very close to the published figure.

Figure 24: Den Haan and Sterk (2011) Figure 7

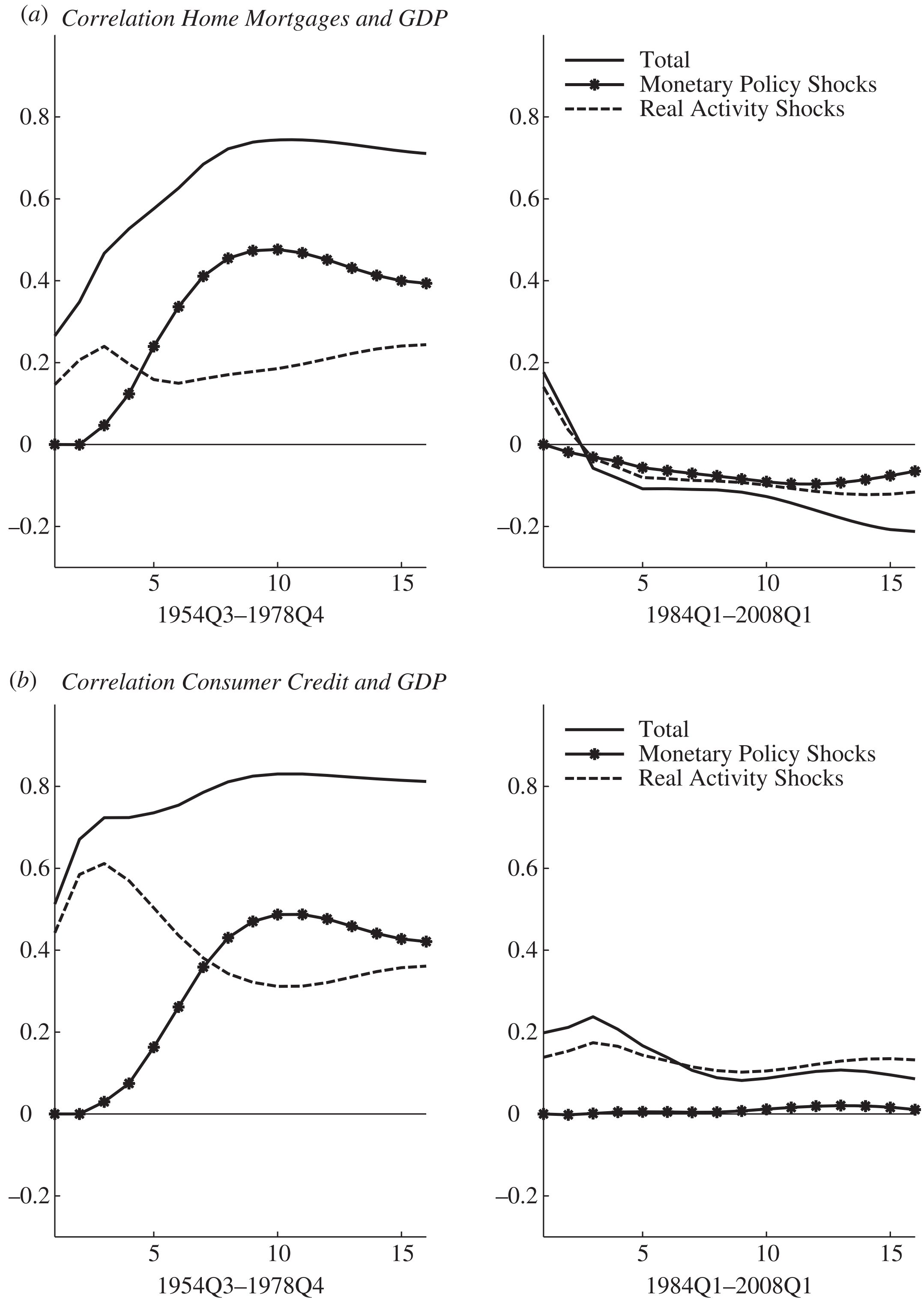
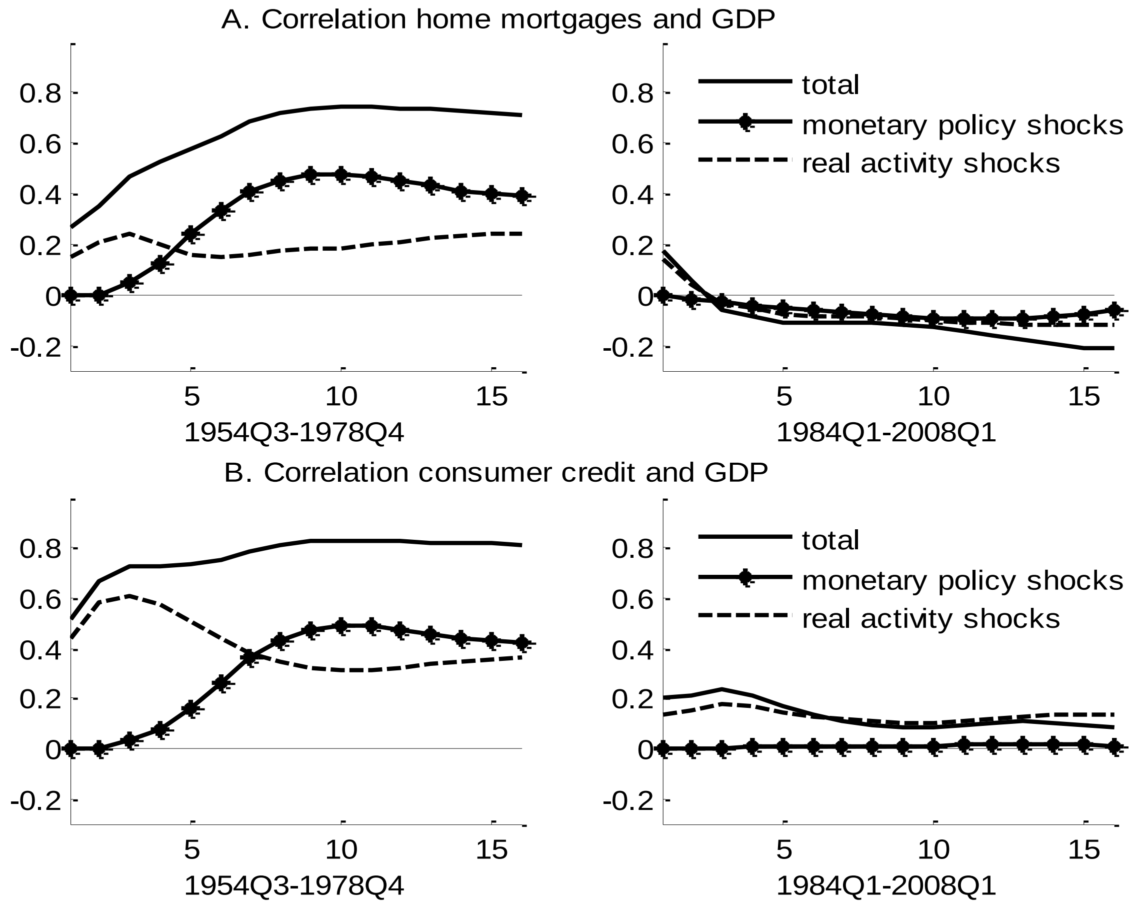


Figure 25: Replication of Den Haan and Sterk (2011) Figure 7



Our replicated correlations were very similar to the published results.

7.26 Favero and Giavazzi (2012), "Measuring Tax Multipliers: The Narrative Method in Fiscal VARs"

Dependent variables: ln(real GDP per capita), differenced ln(GDP Deflator), ln(per capita Real Primary Government Expenditure), Average Nominal Cost of Public Debt, ln(Government Receipts at Annual Rates).

Key Results: "If the effects of shocks identified by the [Romer & Romer] narrative method are analyzed in the context of a multivariate VAR... then the multiplier is not different from that obtained in the traditional fiscal VAR approach." (introduction, pg. 70)

We interpreted the key results as Figures 5 to 6, with Figures 3 to 4 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). The data and code were downloaded from the *American Economic Journal: Economic Policy*'s website. We classified this replication as successful.

7.27 Fève and Guay (2010), "Identification of Technology Shocks in Structural VARs"

Dependent variables: Hours worked, growth rate of GDP deflator, Federal Funds Rate.

Key Results: "a significant short-run decrease of hours after a technology improvement followed by a hump-shaped positive response. Additionally, the rate of inflation and the nominal interest rate displays a significant decrease after this shock." (abstract)

We had neither data or code for this paper, so we were unable to replicate any of the key results. We classified this paper as "failed due to missing public data or code."

7.28 Fisher and Peters (2010), "Using Stock Returns to Identify Government Spending Shocks"

Dependent variables: ln(real GDP), ln(per capita hours worked), ln(real consumption of nondurables and services), ln(real wages).

Key Results: "We estimate the government spending multiplier associated with increases in military spending to be about 1.5 over a horizon of 5 years."(abstract)

This paper used proprietary data, so we were unable to replicate any of the key results. We classified this paper as "failed due to proprietary data."

7.29 Gabaix (2011), "The Granular Origins of Aggregate Fluctuations"

Dependent variables: Per capita GDP growth.

Key Results: "The idiosyncratic movements of the largest 100 firms in the United States appear to explain about one-third of variations in output growth." (abstract)

We took the key results as Figure 2 and Tables 1 to 5. We were able to replicate exactly using the data and code from *Econometrica*'s website without assistance from the author using R 3.1.0 (Linux). We classified this replication as successful.

7.30 Gordon (2010), "Okun's Law and Productivity Innovations"

Dependent variables: Output per hour, Hours per employee, Employment Rate, Labor Force Participation Rate.

Key Results: "Okun's Law responses, suggested in 1965 correspond remarkably well to the actual responses over... 1962-86," "Okun's Law is obsolete for the 1986-2009 interval" and "there is no longer any procyclical responsiveness of output per hour" (pg. 13)

We had neither data or code for this paper, so we were unable to replicate any of the key results. We classified this paper as "failed due to missing public data or code."

7.31 Gilchrist and Zakrajšek (2012), "Credit Spreads and Business Cycle Fluctuations"

Dependent variables: Average change in GDP over several quarters and similar changes in various components of GDP, nominal Federal Funds Rate, excess bond premium, cumulative excess market return, log difference of GDP deflator, 10-year nominal treasury yield.

Key Results: "shocks to the excess bond premium that are orthogonal to the current state of the economy lead to economically and statistically significant declines in consumption, investment, and output, as well as to appreciable disinflation." (introduction, pg. 1693)

This paper used proprietary data, so we were unable to replicate any of the key results. We classified this paper as "failed due to proprietary data."

7.32 Hall and Sargent (2011), "Interest Rate Risk and Other Determinants of Post-WWII US Government Debt/GDP Dynamics"

Dependent variables: Federal Debt/GDP.

Key Results: "shows contributions that nominal returns, the maturity composition of the debt, inflation, and growth in real gross domestic product (GDP) have made to the evolution of the US debt-GDP ratio since World War II." (introduction, pg. 192)

This paper used proprietary data, so we were unable to replicate any of the key results. We classified this paper as "failed due to proprietary data."

7.33 Hansen, Lunde, and Nason (2011), "The Model Confidence Set"

Dependent variables: Federal Funds Rate, although the main interest of the paper is model selection.

Key Results: "a MCS for Taylor rule regressions" (introduction, pg. 454)

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

7.34 Heutel (2012), “How Should Environmental Policy Respond to Business Cycles? Optimal Policy Under Persistent Productivity Shocks”

Dependent variables: Carbon Dioxide Emissions, Various Transformations.

Key Results: “[carbon dioxide] emissions are significantly procyclical with an elasticity between 0.5 and 0.9.” (introduction, pg. 245)

We took the key result as Table 1. Using the data and code files from the *Review of Economic Dynamics*’ website, we were able to replicate the table exactly without the author’s assistance using Stata 13.1 SE (Windows). We classified this replication as successful.

7.35 Ilbas (2012), "Revealing the Preferences of the US Federal Reserve"

Dependent variables: Various, including log difference of real GDP.

Key Results: "monetary policy in the Volcker-Greenspan period performed optimally under commitment." and "the Great Moderation of output growth is explained by a combination of two factors: the decrease in the volatility of the structural shocks and the improved monetary policy conduct." (abstract)

We only had the data but not the code for this paper, so we were unable to replicate any of the key results. We classified this paper as "failed due to missing public data or code."

7.36 Inoue and Rossi (2011), "Identifying the Sources of Instabilities in Macroeconomic Fluctuations"

Dependent variables: Volatilities of: real GDP per capita, Three-month T-bill rate, first difference of GDP deflator.

Key Results: "the Great Moderation was due not only to changes in shock volatilities but also to changes in monetary policy parameters, as well as in the private sector's parameters." and "alternative sources of instabilities appear to have offsetting effects on output volatility..." (abstract)

We interpreted 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) concluded that most of the parameters in their New Keynesian model were unstable (hypothesis tests generally rejected 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 showed that the contribution to the standard deviations of inflation and output from the model parameters were often of opposite signs, so the net effect on the standard deviation of output and inflation was mitigated. Using the replication files from *Review of Economics and Statistics*' website, we were able to qualitatively match the results from their Tables 1 and 3, but there were some minor differences using Matlab R2010a (Linux). In particular, we were able to exactly match a vast majority of the p-values from their Table 1, but we found minor differences in many of the estimates from Table 2. We classified this replication as successful.

Table 37: 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
ρ_{gy}	0	0
σ_e	0	0
ρ_υ	0	0
ρ_π	0	0
σ_z	0.76	0.76

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 38: 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." We continued to find that only σ_z was a stable parameter, although we found a different p-value than Inoue and Rossi (2011).

Table 39: Inoue and Rossi (2011) Table 3

Parameter:	Output	Inflation	Interest Rate
No change: (actual S.D.)	0.88	0.44	0.29
Unstable Parameters	% Contribution to Change		
ρ_e	7%	10%	0%
σ_ν	62%	33%	83%
α	-3%	16%	7%
σ_a	-15%	-3%	-84%
σ_π	3%	12%	26%
ρ_a	17%	2%	76%
γ	32%	0%	-9%
ψ	0%	0%	0%
ρ_{gy}	-43%	2%	17%
σ_e	-2%	-5%	0%
ρ_υ	5%	4%	-8%
ρ_π	-13%	-23%	5%
Stable Parameters:			
σ_z	49%	52%	3%
All change: (actual S.D.)	1.44	0.91	0.38

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).

Table 40: 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). Our replication of this table had some minor differences to the published results, but the contributions of our replicated parameters to the volatilities of output, inflation, and the interest rate were similar in magnitude to Inoue and Rossi (2011).

7.37 Ireland (2009), "On the Welfare Cost of Inflation and the Recent Behavior of Money Demand"

Dependent variables: Various measures for: ln(nominal money balances/ nominal income), ln(short-term nominal interest rate).

Key Results: "the semi-log specification (2) with its finite satiation point may now provide a more accurate description of money demand [compared to a log-log specification]" (pg. 1043)

This key result corresponded to Table 2, which we were able to replicate exactly using code and data from *American Economic Review*'s website without help from Peter Ireland using Matlab R2010a (Linux). We classified this replication as successful.

7.38 Jermann and Quadrini (2012), "Macroeconomic Effects of Financial Shocks"

Dependent variables: Various, including real GDP growth and real GDP level.

Key Results: "financial shocks are important not only for capturing the dynamics of financial flows but also for the dynamics of the real business cycle quantities, especially labor." and "financial shocks contribute to almost half of the volatility of output..." (introduction, pg. 239)

We did not have the correct software to run the code for this paper. We classified this paper as "failed due to missing software."

7.39 Jore, Mitchell, and Vahey (2010), "Combining Forecast Densities from VARs with Uncertain Instabilities"

Dependent variables: Quarterly logarithmic changes of GDP growth and GDP deflator. Various short-term interest rates.

Key Results: "Our proposed recursive-weight density combination strategy... produces well-calibrated predictive densities for US real-time data by giving substantial weight to models that allow for structural breaks. In contrast, equal-weight combinations produce poorly calibrated forecast densities..." (abstract)

We only had the data but not the code for this paper, so we were unable to replicate any of the key results. We classified this paper as "failed due to missing public data or code."

7.40 Justiniano, Primiceri, and Tambalotti (2011), "Investment Shocks and the Relative Price of Investment"

Dependent variables: Various, including first-differenced real per capita GDP.

Key Results: "[marginal efficiency of investment] shocks explain between 60 and 85 percent of the variance of output, hours and investment at business cycle frequencies..." and "the contribution of the [investment-specific technology] shocks is minimal" and "[marginal efficiency of investment] shocks implied by our estimates is highly correlated with credit spreads and that it accounts for most of the fall in output and hours in 2007 and 2008." (introduction, pg. 103)

We attempted to replicate Tables 2-4 and Figures 2-5. From the RED's website, we downloaded an archive that had data and code, but we were only able to use the archive to qualitatively replicate Table 1, and Table 1 was not one of the key results that we identified in our preanalysis plan. We did not find the code to replicate Tables 2-4 and Figures 2-5. Therefore, we marked this paper as a "failed due to missing public data or code."

7.41 Keen (2009) "Output, Inflation, and Interest Rates in an Estimated Optimizing Model of Monetary Policy"

Dependent variables: real GDP per capita, percent change in GDP implicit price deflator, Federal Funds Rate.

Key Results: "a sticky price and limited participation model fits the data [best]," "[this] model is able to produce simultaneously the output, inflation, and liquidity effects after a monetary disturbance." and "monetary policy shocks account for a modest portion of the variability in output, inflation, and the nominal interest rate." (introduction, pg. 328).

We attributed these results to Figures 1-2 and Tables 4-6. Using data and code from *Review of Economic Dynamics*' website, we were able to closely replicate Figures 1-2, and exactly replicate Tables 5-6, but were unable to obtain Table 4. The programs were run twice for over two months each time, but they did not terminate, either because the code did not finish or the code was stuck somewhere. Because the code did not terminate and we were unable to obtain a majority of the key results for each table and figure, we classified this paper as "failed due to incorrect public data or code."

Figure 26: Keen (2009) Figure 1

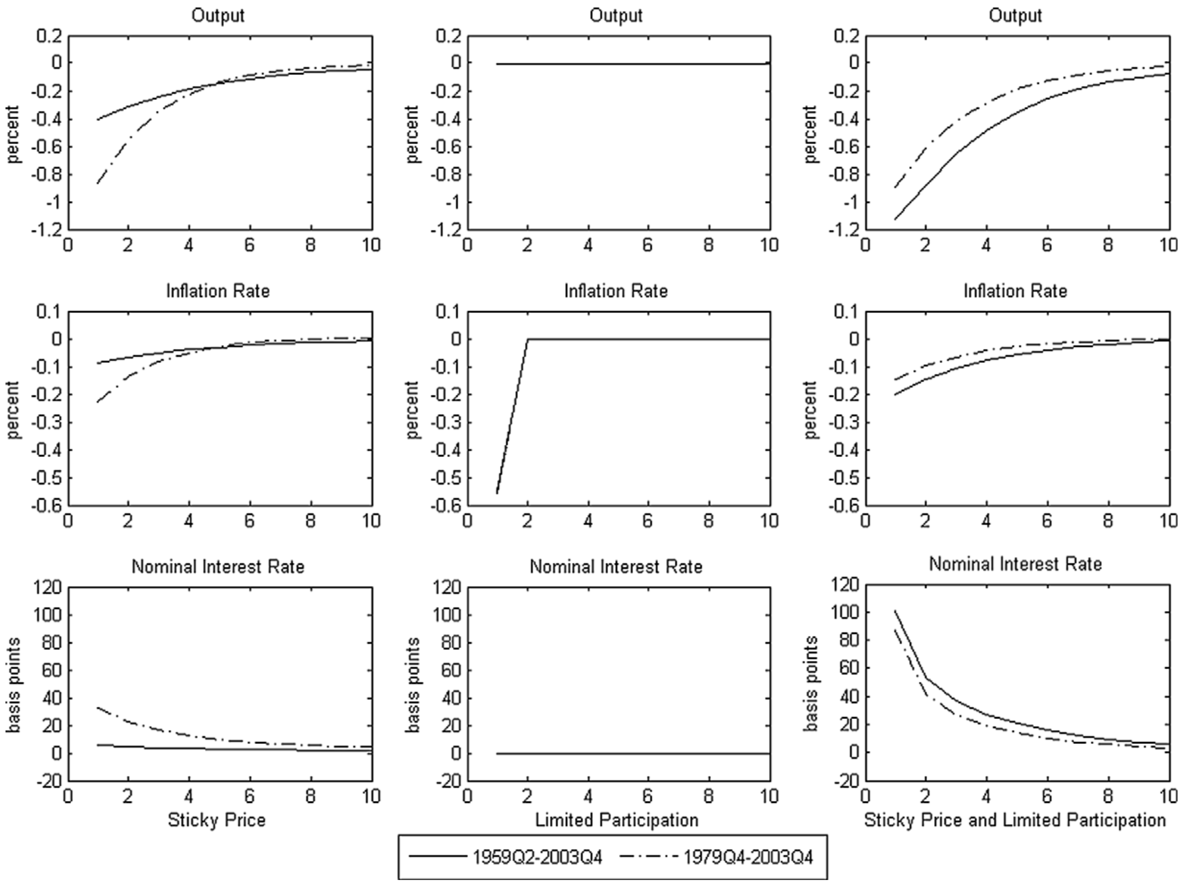
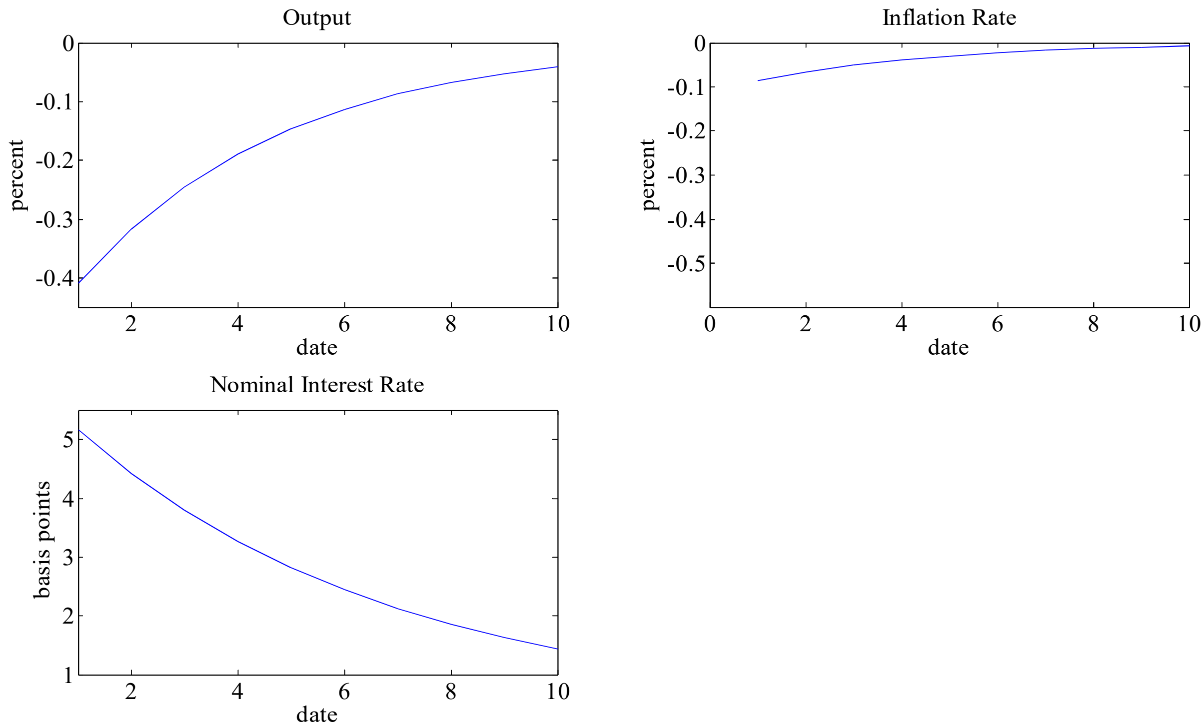
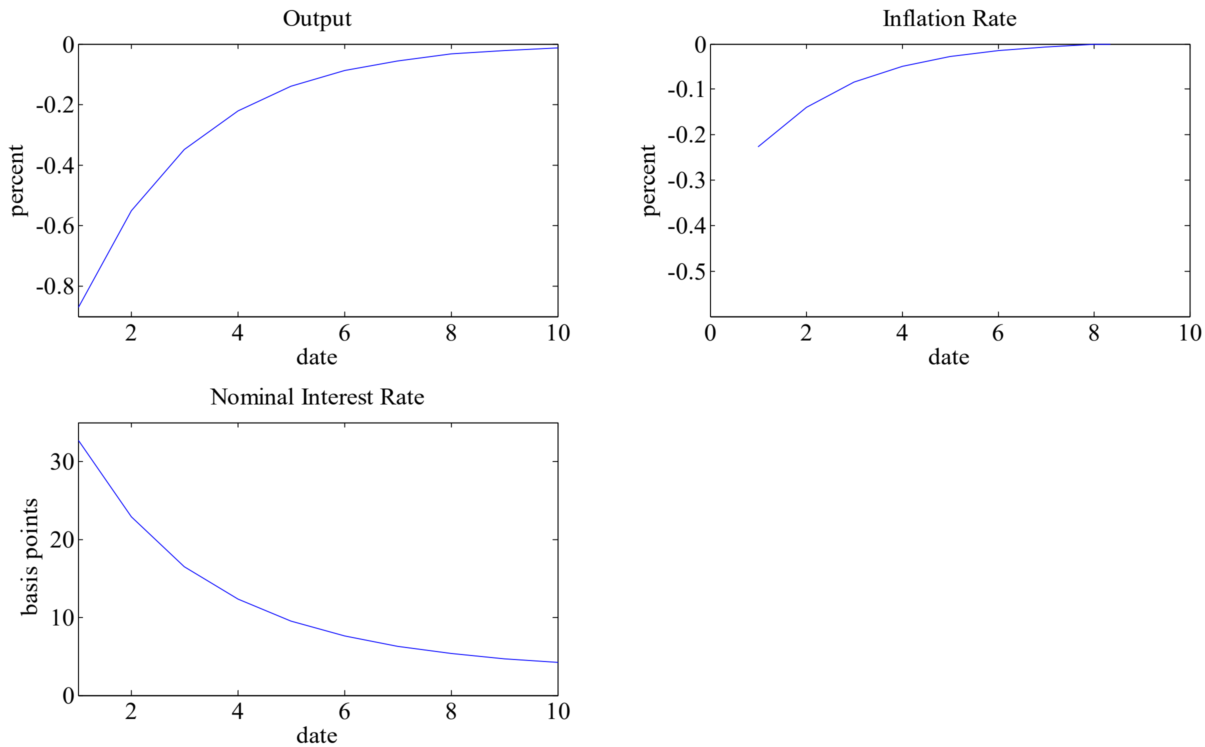


Figure 27: Replication of Keen (2009) Figure 1, Sticky Price plots for 1959Q2-2003Q4



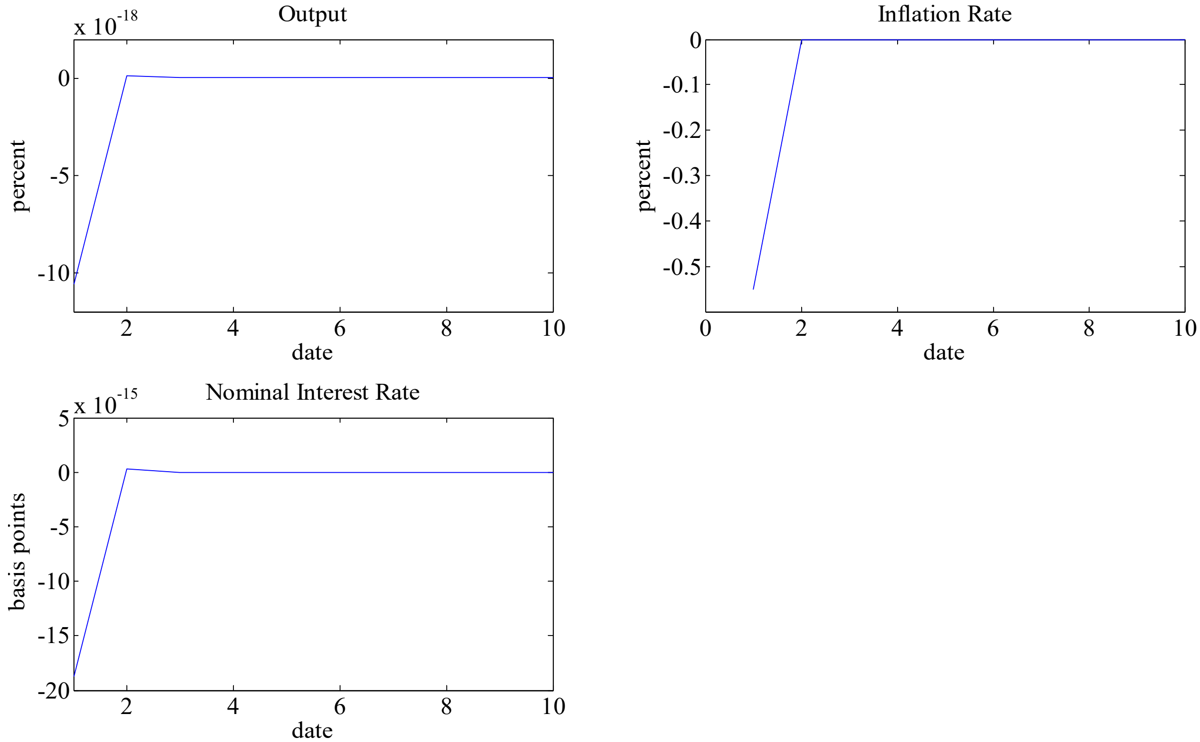
Our replicated plots were very similar to the published results.

Figure 28: Replication of Keen (2009) Figure 1, Sticky Price plots for 1979Q4-2003Q4



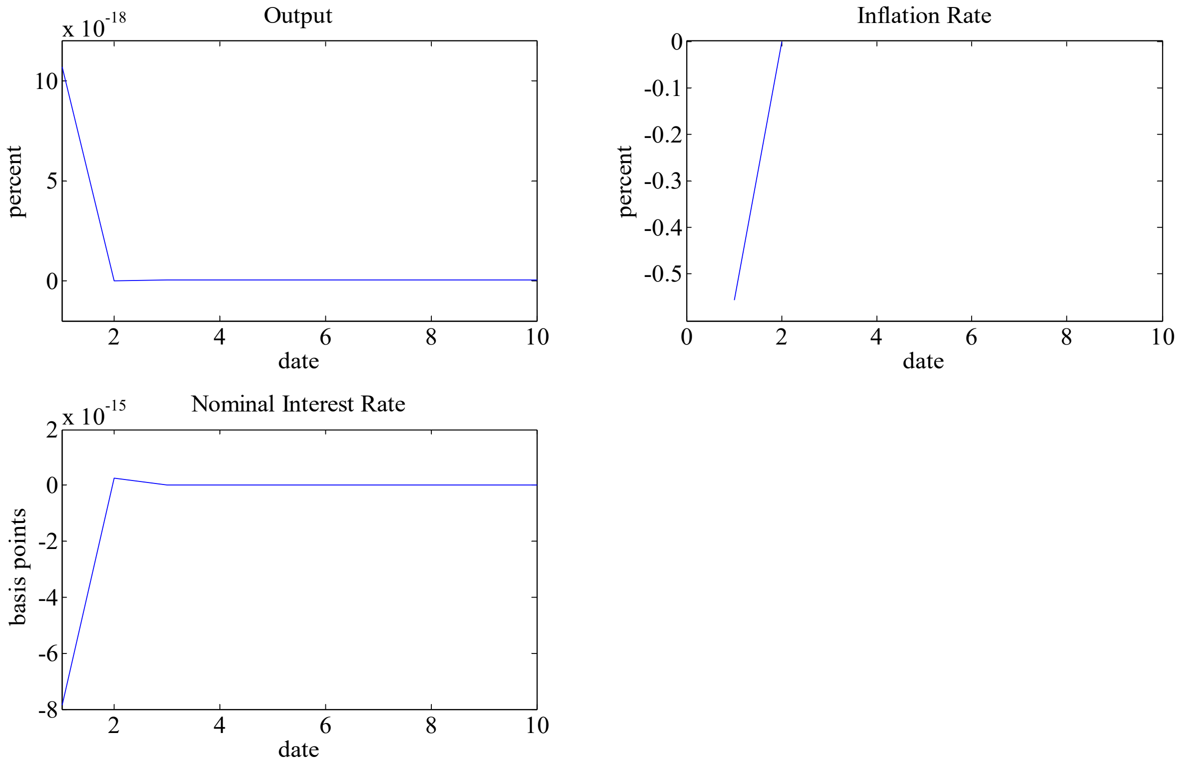
Our replicated plots were very similar to the published results.

Figure 29: Replication of Keen (2009) Figure 1, Limited Participation plots for 1959Q2-2003Q4



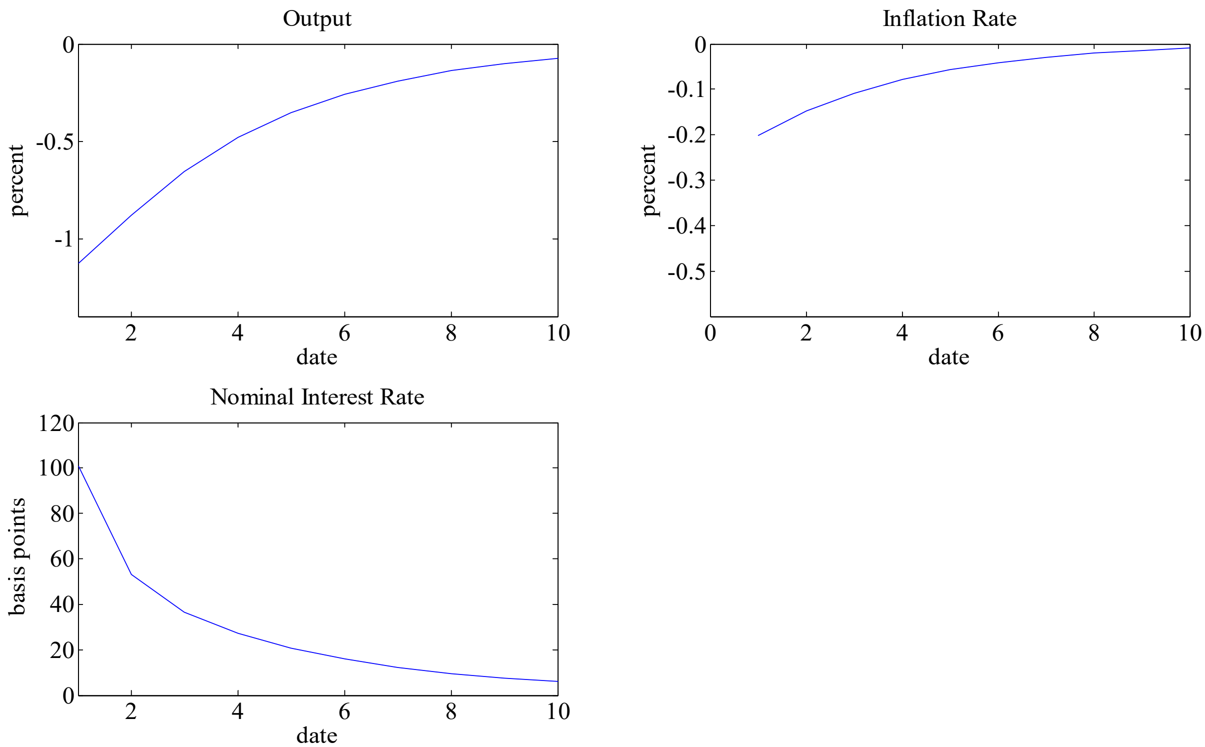
Our replicated plots were very similar to the published results.

Figure 30: Replication of Keen (2009) Figure 1, Limited Participation plots for 1979Q4-2003Q4



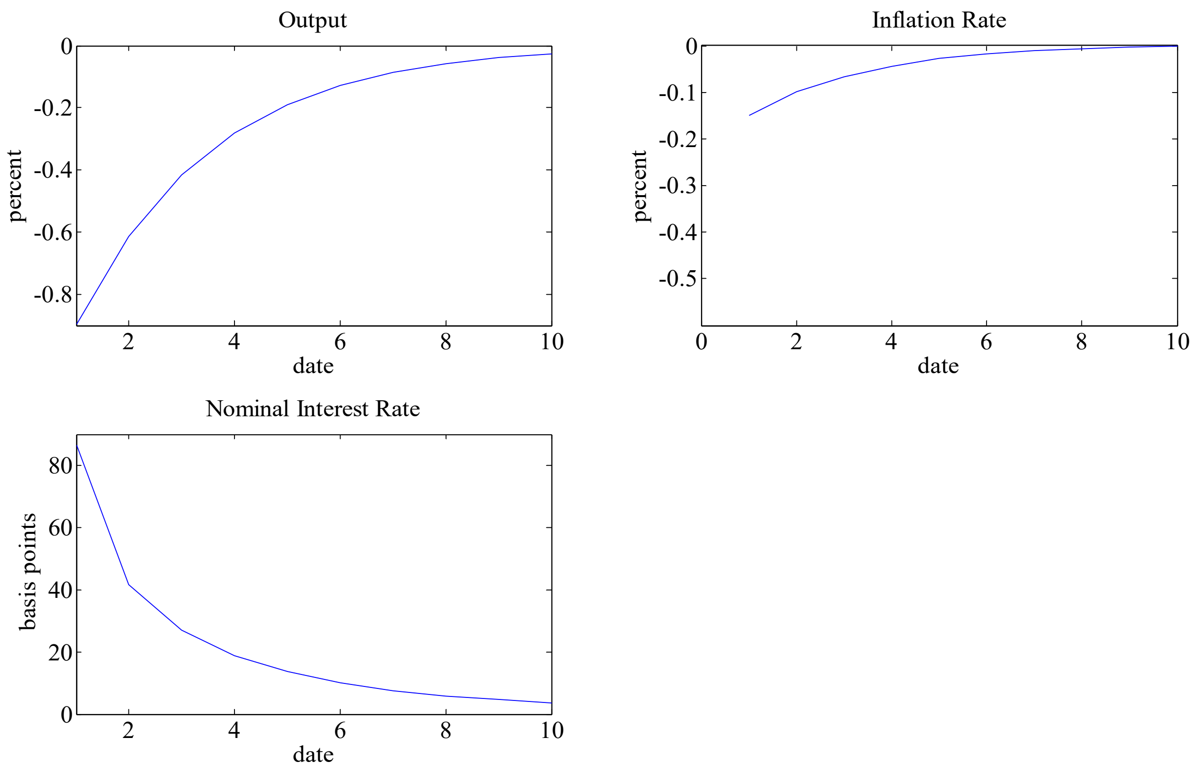
Our replicated plots were very similar to the published results.

Figure 31: Replication of Keen (2009) Figure 1, Sticky Price and Limited Participation plots for 1959Q2-2003Q4



Our replicated plots were very similar to the published results.

Figure 32: Replication of Keen (2009) Figure 1, Sticky Price and Limited Participation plots for 1979Q4-2003Q4



Our replicated plots were very similar to the published results.

Figure 33: Keen (2009) Figure 2

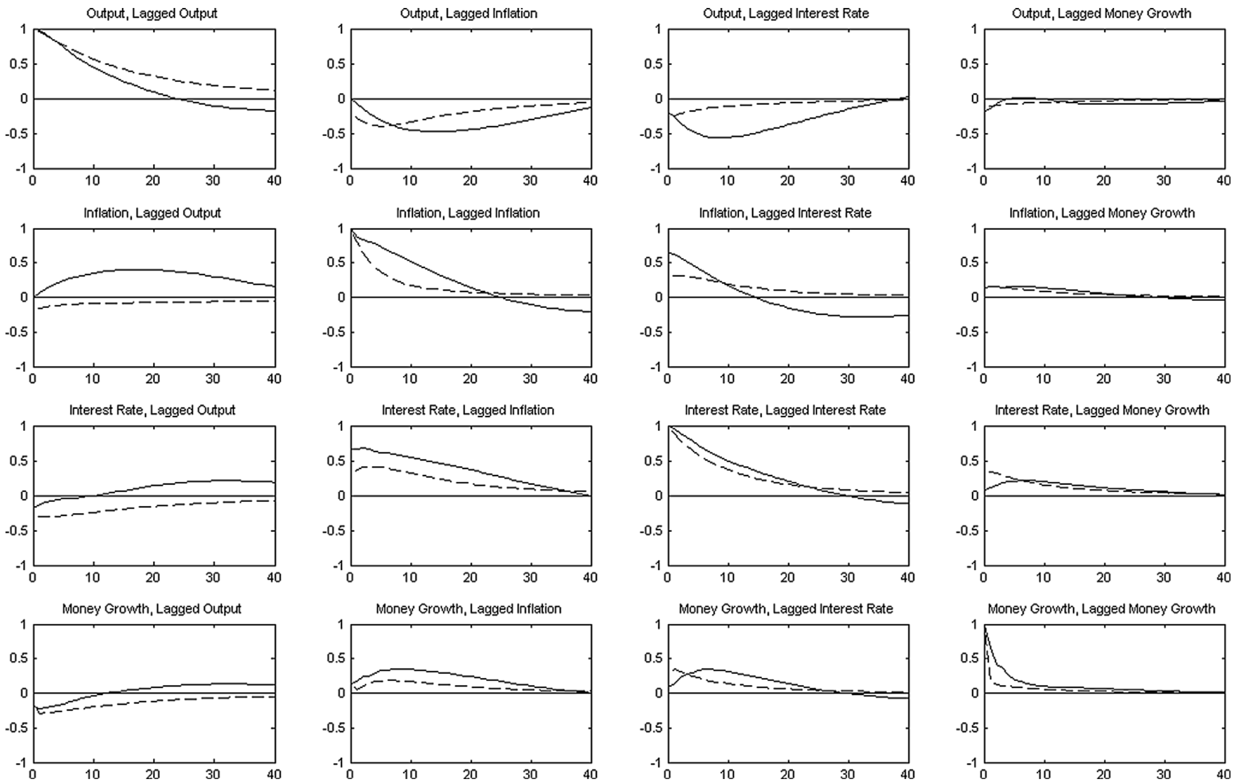
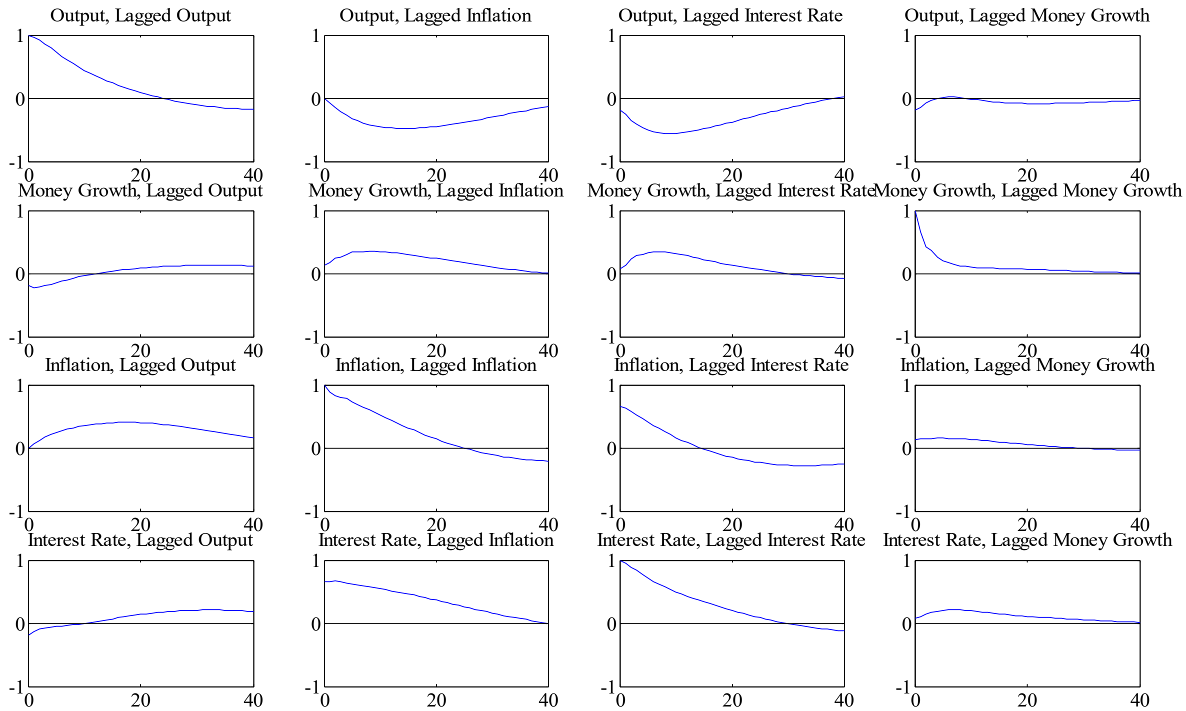
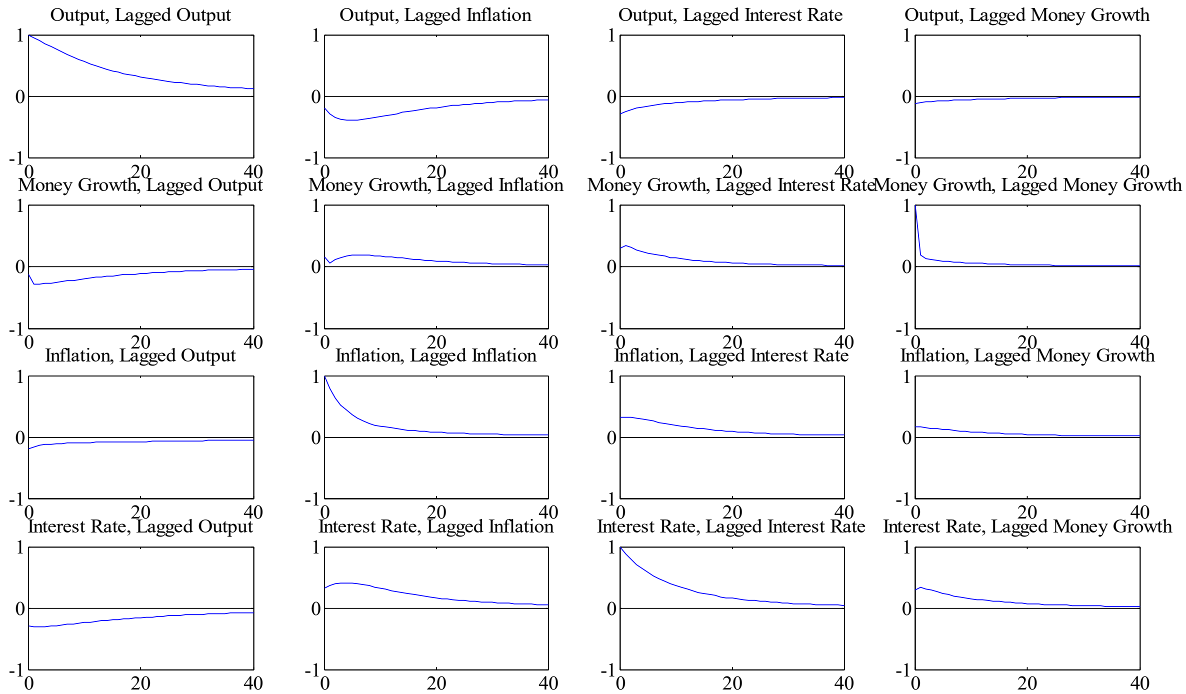


Figure 34: Replication of Keen (2009) Figure 2, Data plots



Our replicated plots were very similar to the published results.

Figure 35: Replication of Keen (2009) Figure 2, Sticky Price and Limited Participation model plots



Our replicated plots were very similar to the published results.

7.42 Kilian (2009), "Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market"

Dependent variables: CPI, real GDP growth.

Key Results: "oil price increases my have very different effects on the real price of oil, depending on the underlying cause of the price increase." (followed by list of examples, introduction, pg. 1053)

We took the key figure as Figure 5, namely the responses of GDP to oil supply shocks. We were able to replicate exactly this figure using data and code from the *American Economic Review*'s website without assistance from the author in Matlab R2013a (Windows). We classified this replication as successful.

7.43 Kormilitsina (2011), “Oil Price Shocks and the Optimality of Monetary Policy”

Dependent variables: Various, including seasonally adjusted GDP per capita.

Key Results: “I find that monetary policy amplified the negative effect of the oil price shock. The optimal response to the shock would have been to raise inflation and interest rates above what had been seen in the past.” (abstract)

We interpreted the key result of this paper as Figure 2, which contrasted 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 *Review of Economic Dynamics*’ website. We used Fortran f90 (Linux) and Matlab R2012a (Windows). We classified this replication as successful.

7.44 Krishnamurthy and Vissing-Jorgensen (2012), "The Aggregate Demand for Treasury Debt"

Dependent variables: Various Measures of Bond Spreads.

Key Results: "When the supply of Treasuries is low, the value that investors assign to the liquidity and safety attributes offered by Treasuries... is high."(introduction, pg. 235)

We took 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) was proprietary, so we replicated the columns in Tables 1 to 2 that did 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 *Journal of Political Economy* in Matlab R2013a (Windows). We classified this replication as successful.

7.45 Levine, Pearlman, Perendia, and Yang (2012), "Endogenous Persistence in an Estimated DSGE Model Under Imperfect Information"

Dependent variables: N/A

Key Results: "empirical support for [imperfect information] as an endogenous persistence mechanism, but this is dominated by that from habit and adaptive learning." (abstract)

We only had the data but not the code for this paper. Since we were unable to replicate any of the key results, we classified this paper as "failed due to missing public data or code."

7.46 Maheu and Gordon (2008), "Learning, Forecasting, and Structural Breaks"

Dependent variables: Real GDP growth.

Key Results: "We consider predictions of real US GDP and document... a gradual reduction in volatility over time with evidence of three separate regimes. The model is particularly useful in forecasting the probability of positive growth." (introduction, pg. 556)

We only had the data but not the code for this paper. Since we were unable to replicate any of the key results, we classified this paper as "failed due to missing public data or code."

7.47 Mavroeidis (2010), "Monetary Policy Rules and Macroeconomic Stability: Some New Evidence"

Dependent variables: Federal Funds Rate, although it is not the main focus of the paper.

Key Results: "[Taylor] policy rule parameters appear to be well identified in the pre-Volcker sample." and "in subsequent periods... the policy reaction function is not well identified." (introduction, pg. 491)

We were able to replicate these figures exactly without assistance from the author using data and code from the *American Economic Review* website in OX 6.30 (Linux). We classified this replication as successful.

7.48 Mertens and Ravn (2010), "Measuring the Impact of Fiscal Policy in the Face of Anticipation: A Structural VAR Approach"

Dependent variables: per capita real: GDP, private consumption of nondurables and services, government consumption expenditures.

Key Results: "output and consumption rise in response to an unanticipated permanent increase in government spending..." (introduction, pg. 395)

The key result corresponded to Figures 4-8. We obtained the data and code archive from Morten O. Ravn but were unable to produce any of the desired figures. The archive did not contain any replication instructions, and we could not identify the appropriate order of code execution to produce the results. For these reasons, we classified this paper as "failed due to incorrect public data or code."

7.49 Mertens and Ravn (2011), “Understanding the Aggregate Effects of Anticipated and Unanticipated Tax Policy Shocks”

Dependent variables: ln(per capita real GDP), ln(real nondurables consumption per capita), ln(real durables consumption per capita), ln(real investment per capita), ln(average hours worked).

Key Results: “unanticipated tax cuts have persistent expansionary effects on output, consumption, investment, and hours worked. Anticipated tax cuts give rise to contractions in output, investment and hours worked prior to their implementation, while stimulating the economy when implemented.” (abstract)

We took 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 *Review of Economic Dynamics*’ website using Matlab R2013a (Windows). We classified this replication as successful.

7.50 Mertens and Ravn (2012), "Empirical Evidence on the Aggregate Effects of Anticipated and Unanticipated US Tax Policy Shocks"

Dependent variables: ln(real GDP per adult), ln(real per capita consumption), ln(real per capita gross investment), ln(per capita hours worked), real wage.

Key Results: "Preannounced but not yet implemented tax cuts give rise to contractions in output, investment, and hours worked while real wages increase." and "Implemented tax cuts... have expansionary effects, on output, consumption, investment, hours worked, and real wages." (abstract)

We only had the data but not the code for this paper, so we were unable to replicate any of the key results. We classified this paper as "failed due to missing public data or code."

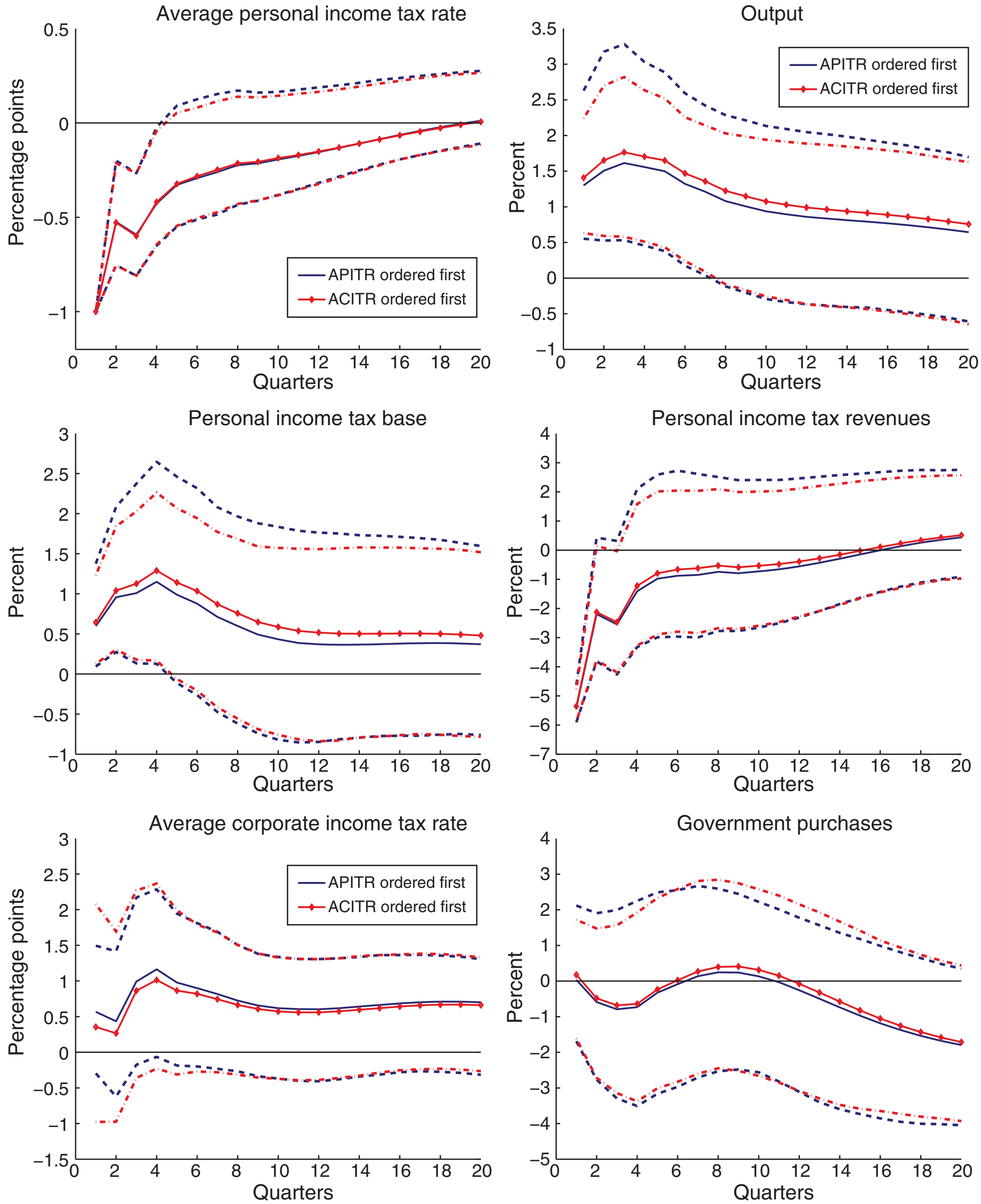
7.51 Mertens and Ravn (2013), "The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States"

Dependent variables: Average personal income tax rate, Average corporate income tax rate, ln(personal income tax base), ln(corporate income tax base), ln(real government purchases per capita), ln(real GDP per capita), ln(real federal government debt per capita)

Key Results: "short run output effects of tax shocks are large" (abstract, emphasized on pg. 1228)

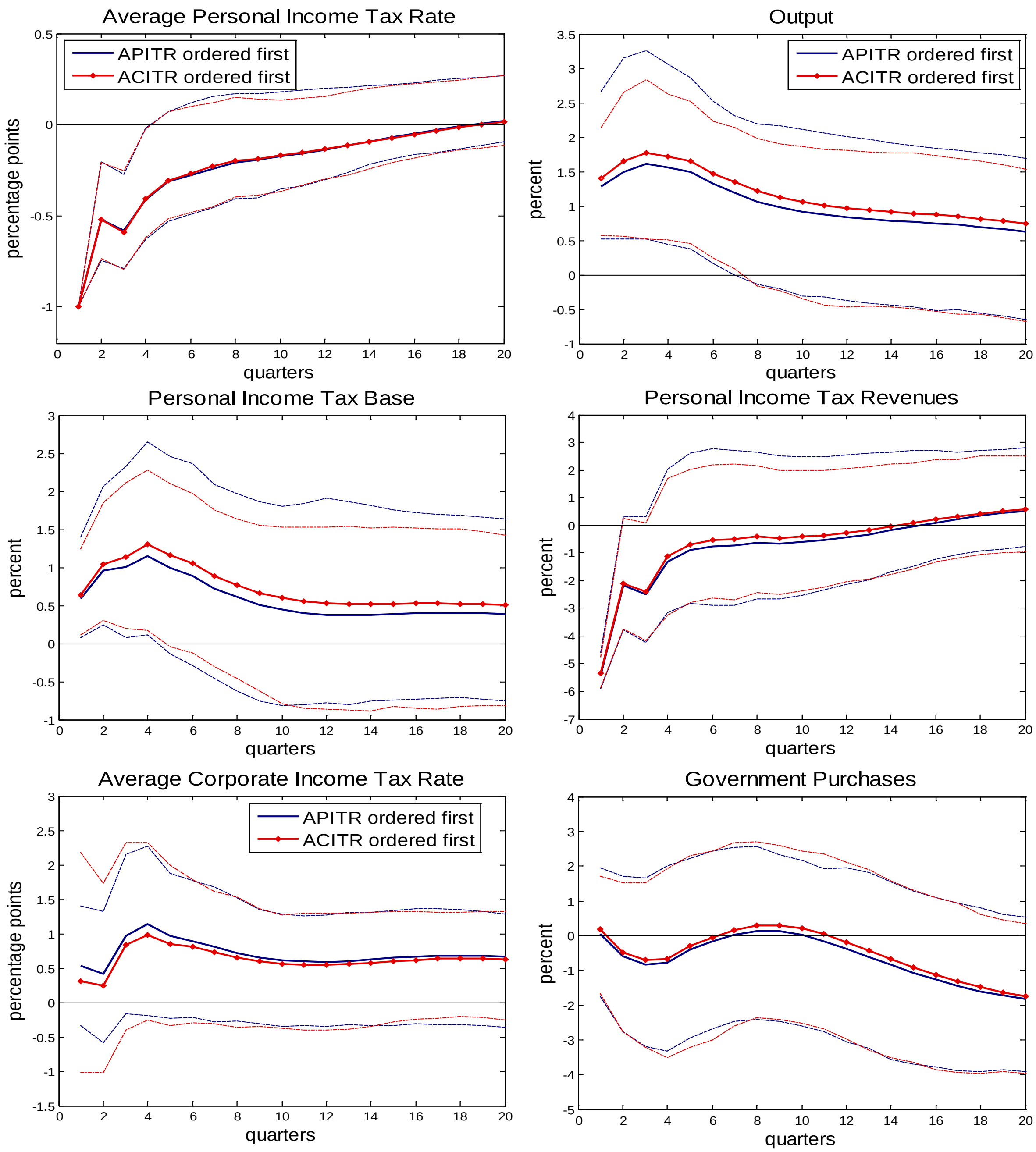
This result corresponded to Figures 2 and 3. We were able to use data and code from the *American Economic Review*'s website and replicate these figures quite closely without help from the authors using Matlab R2013a (Windows). We classified this replication as successful.

Figure 36: Mertens and Ravn (2013) Figure 2



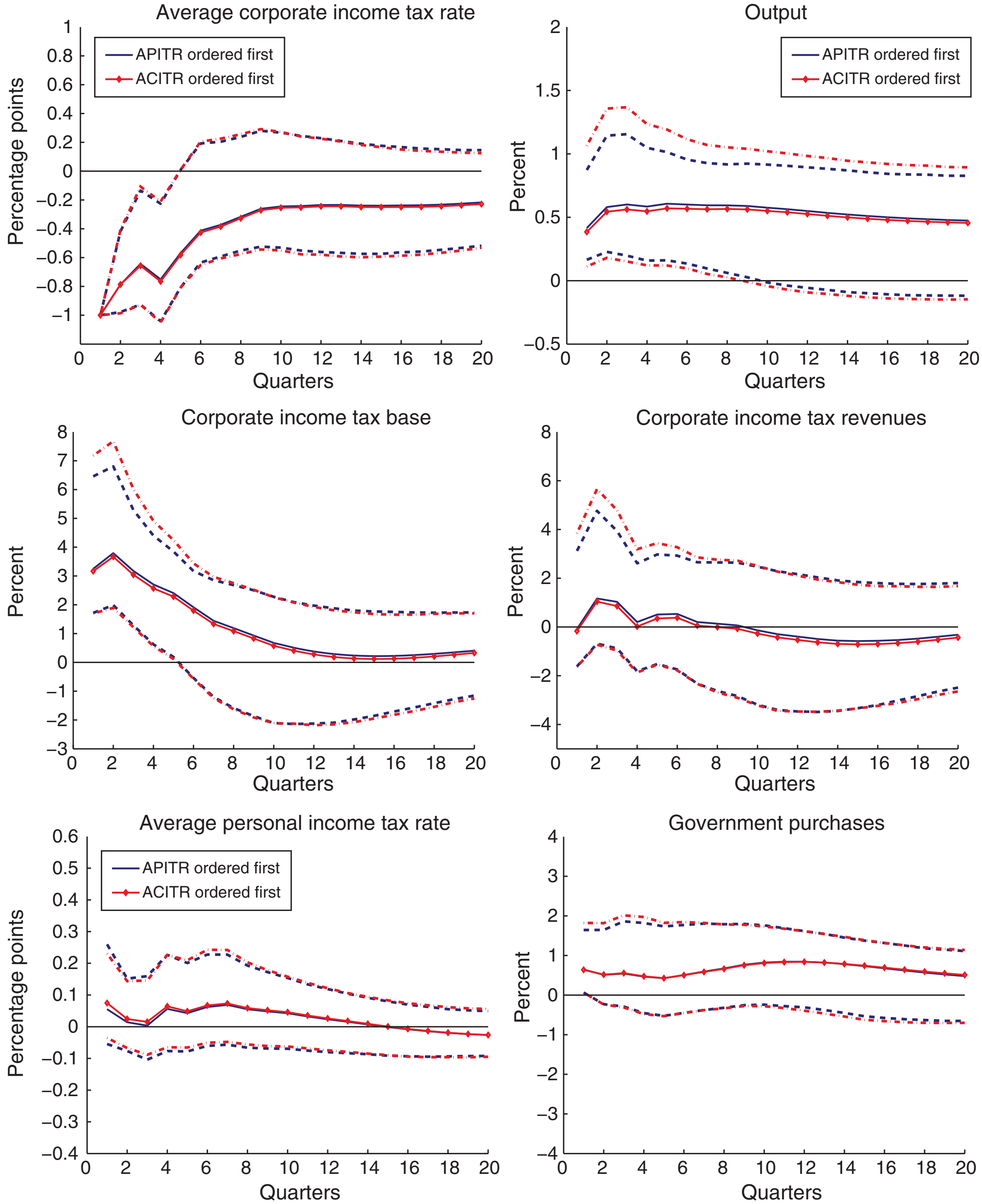
Impulse responses from a one percentage point cut in the average personal income tax rate. Dashed lines are 95% confidence intervals.

Figure 37: Replication of Mertens and Ravn (2013) Figure 2



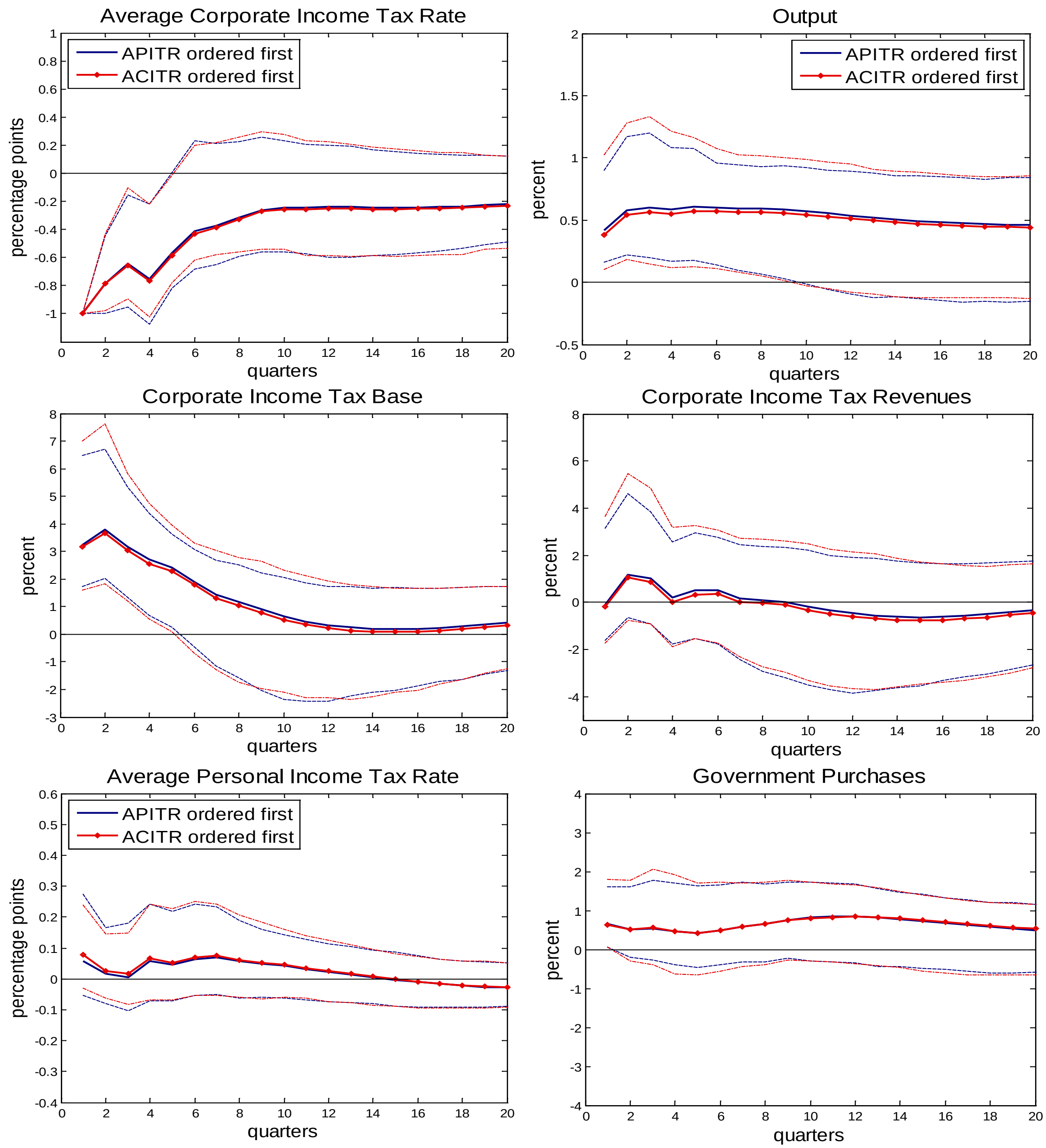
Impulse responses from a one percentage point cut in the average personal income tax rate. Dashed lines are 95% confidence intervals. Our replication was very close to the published figure.

Figure 38: Mertens and Ravn (2013) Figure 3



Impulse responses from a one percentage point cut in the average corporate income tax rate. Dashed lines are 95% confidence intervals.

Figure 39: Replication of Mertens and Ravn (2013) Figure 3



Impulse responses from a one percentage point cut in the average corporate income tax rate. Dashed lines are 95% confidence intervals. Our replication was very close to the published figure.

7.52 Milani (2011), "Expectation Shocks and Learning as Drivers of the Business Cycle"

Dependent variables: Detrended Real GDP, GDP deflator, 3-month Treasury Bill Rate, various expectations of these variables.

Key Results: "Expectation shocks explain roughly half of business cycle movements." (introduction,pg. 381)

We had neither data or code for this paper, so we were unable to replicate any of the key results. We classified this paper as "failed due to missing public data or code."

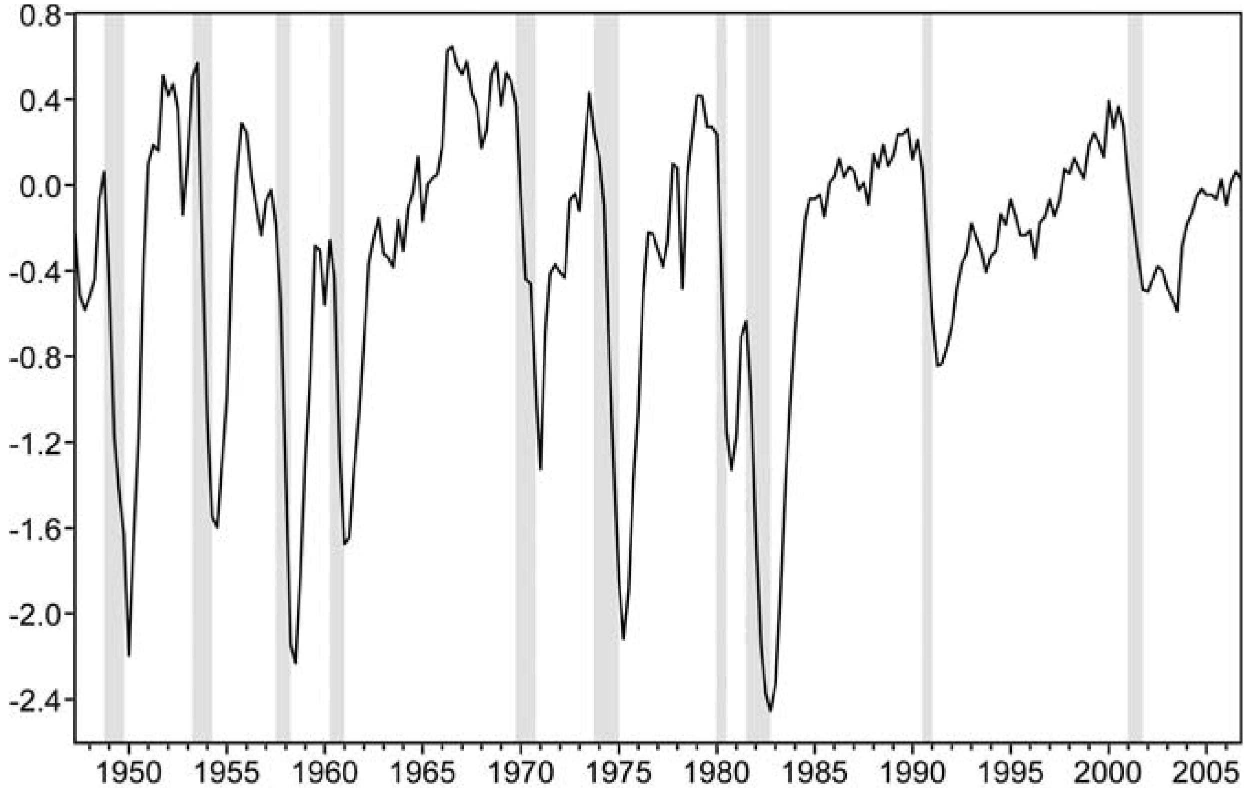
7.53 Morley and Piger (2012), "The Asymmetric Business Cycle"

Dependent variables: Various transformations of real GDP.

Key Results: "[Our model-averaged] measure [of the business cycle] also displays an asymmetric shape."(abstract)

We took the key result from this paper as Figure 3, namely that their model-averaged measure of the business cycle displayed an asymmetric shape, with the variance of output in a recession being greater than the variance of output during an expansion. We obtained the replication files from *Review of Economics and Statistics*' website. The authors specified that they ran their programs in Gauss 10 on Mac OS X in their readme (one of the few papers that specified 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 replicated their Figure 3 quite closely with help from the authors. We classified this replication as successful.

Figure 40: Morley and Piger (2012) Figure 3



NBER recessions are shaded.

A model-averaged measure of the business cycle.

Figure 41: Replication of Morley and Piger (2012) Figure 3

NBER recessions are shaded.

A model-averaged measure of the business cycle. Our replication of this figure was almost exact.

7.54 Mountford and Uhlig (2009), "What Are the Effects of Fiscal Policy Shocks?"

Dependent variables: Ten variable VAR, including ln(real per capita GDP)

Key Results: "deficit-financed tax cuts work best... to improve GDP, with a maximal present value multiplier of five dollars of total additional GDP per each years after the shock." (abstract)

This result was attributed to Figures 10-13 and Tables 2-5. We used the replication archive from *Journal of Applied Econometrics*' website. The programs produced Figure 7 from the published paper but did not produce any of the other figures or tables. For these reasons, we classified this paper as "failed due to incorrect public data or code."

7.55 Nakov and Pescatori (2010), "Oil and the Great Moderation"

Dependent variables: Volatility of: real GDP growth, GDP deflator, Federal Funds Rate, real West Texas Intermediate oil price.

Key Results: "57% of the reduced volatility of GDP growth [during the Great Moderation] is attributed to smaller TFP shocks. Oil related effects explain around a third." (abstract)

We took the key result from Nakov and Pescatori (2010) as Table 5, although Tables 3 to 4 and 6 to 7 lent supporting evidence to Table 5. Our estimates of the posterior distribution in their Table 3 were slightly off, but the qualitative results of the paper still held. Using data and code from Anton Nakov, we were able to replicate Tables 4 to 5 almost exactly, and Tables 6 to 7 exactly. We performed the estimation in Matlab R2012a (Linux). We classified this replication as successful.

Table 41: Nakov and Pescatori (2010) Table 3, Posterior Parameters

	Mean	Standard Deviation	Mode
1970-1983			
θ	0.614	0.068	0.622
ψ	0.961	0.224	0.897
ϕ_i	0.545	0.079	0.537
ϕ_π	2.438	0.359	2.224
ϕ_y	0.545	0.108	0.531
ρ_a	0.958	0.017	0.969
ρ_b	0.890	0.035	0.896
ρ_z	0.917	0.032	0.927
ρ_ω	0.926	0.031	0.937
$100\sigma_a$	1.359	0.127	1.331
$100\sigma_b$	2.762	0.599	2.207
$100\sigma_z$	21.27	2.436	21.33
$100\sigma_\omega$	34.95	6.763	30.73
$100\sigma_r$	0.530	0.068	0.494
1984-2007			
θ	0.473	0.063	0.477
ψ	1.070	0.238	1.009
ϕ_i	0.676	0.059	0.691
ϕ_π	3.191	0.295	3.100
ϕ_y	0.535	0.098	0.539
ρ_a	0.978	0.010	0.983
ρ_b	0.951	0.015	0.951
ρ_z	0.881	0.033	0.882
ρ_ω	0.954	0.018	0.960
$100\sigma_a$	0.630	0.045	0.620
$100\sigma_b$	2.133	0.516	1.862
$100\sigma_z$	14.92	1.596	15.18
$100\sigma_\omega$	25.43	4.721	23.40
$100\sigma_r$	0.231	0.034	0.212

Table 42: 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
$100\sigma_a$	1.415	0.127	1.331
$100\sigma_b$	3.070	0.599	2.207
$100\sigma_z$	19.849	2.436	21.332
$100\sigma_\omega$	37.359	6.764	30.730
$100\sigma_r$	0.429	0.068	0.494
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
$100\sigma_a$	0.628	0.045	0.620
$100\sigma_b$	2.258	0.516	1.862
$100\sigma_z$	14.884	1.596	15.177
$100\sigma_\omega$	25.682	4.721	23.397
$100\sigma_r$	0.231	0.034	0.212

Our replication of the posterior mean estimates was slightly off, but still close to the published results. Our replication of the standard deviations and posterior modes was almost exact.

Table 43: Nakov and Pescatori (2010) Table 4

	1970-1983		1984-2007		Volatility Reduction	
	Data	Model	Data	Model	Data	Model
GDP Growth	1.20	1.64	0.52	0.72	57%	56%
Inflation	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	19.0	16.6	13.0	12.0	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 44: Replication of Nakov and Pescatori (2010) Table 4

	1970-1983		1984-2007		Volatility Reduction	
	Data	Model	Data	Model	Data	Model
GDP Growth	1.20	1.64	0.52	0.72	57%	56%
Inflation	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. Our replication of this table was almost exact.

Table 45: Nakov and Pescatori (2010) Table 5

	Oil		Monetary Policy			
	Share	Shocks	Rule	Shocks	TFP Shock	Other Factors
Inflation	32%	17%	40%	11%	2%	-2%
GDP Growth	18%	11%	0%	4%	57%	10%
Interest Rate	12%	3%	37%	4%	8%	36%
Real Oil Price	-3%	101%	0%	0%	0%	2%

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 46: 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. Our replication of this table was almost exact.

7.56 Perotti (2012), "The Effects of Tax Shocks on Output: Not so Large, but Not Small Either"

Dependent variables: Log change of GDP per capita.

Key Results: "responses to a tax shock that allow for a distinction between the discretionary and the endogenous components of tax changes are about halfway between the large effects estimated by Romer and Romer (2010) and the smaller effects estimated..." and "there is almost no... evidence of anticipation effects." (abstract)

We had neither data or code for this paper, so we were unable to replicate any of the key results. We classified this paper as "failed due to missing public data or code."

7.57 Piazzesi and Schneider (2010), "Interest Rate Risk in Credit Markets"

Dependent variables: Household assets and liabilities as a percent of GDP.

Key Results: "we obtain dollar holdings in spanning bonds as a percent of GDP" (description follows, pg. 582) and "the increase in household debt [since 1950 as a percent of GDP] happened in both short and long term debt instruments."(pg. 583)

We had neither data or code for this paper, so we were unable to replicate any of the key results. We classified this paper as "failed due to missing public data or code."

7.58 Polito and Wickens (2012), "Optimal Monetary Policy Using an Unrestricted VAR"

Dependent variables: Deviation of real GDP from HP-trend, CPI.

Key Results: "the optimal solutions for inflation and output, restricting the interest rate to be non-negative... results in a continuous fall in output and a sharp decrease in inflation [over the last three quarters of 2009]" (introduction, pg. 526-527)

We interpreted the key result as Figure 15. We obtained the data from *Journal of Applied Econometrics*' website and the Matlab code from Vito Polito. The code crashed due to a concatenation error that we could not fix and did not produce the desired figure. For these reasons, we classified this paper as "failed due to incorrect public data or code."

7.59 Ramey (2011), "Identifying Government Spending Shocks: It's all in the Timing"

Dependent variables: Six variable VAR that rotates the sixth variable between eight different series. Always includes ln(real per capita GDP).

Key Results: "government spending multipliers range from 0.6 to 1.2." (abstract)

We took the key results from this paper as Figures 10 and 12. The *Quarterly Journal of Economics* did 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). We classified this replication as successful.

7.60 Reis and Watson (2010), "Relative Goods' Prices, Pure Inflation, and The Phillips Correlation"

Dependent variables: Various measures of inflation, GDP, industrial production, unemployment rate, consumption, employment.

Key Results: "macroeconomic shocks account for almost as much as one-third of the movement in sectoral prices" and "controlling for relative goods prices, the Phillips correlation becomes quantitatively negligible." (introduction, pg. 130)

We took the key results to be their Tables 4 and 5. In their Table 4, panels A-C there existed a significant association between GDP and PCE inflation because these panels only controlled for absolute price changes. Their Table 4, panels D-E controlled for relative price changes, and the association between GDP and PCE inflation disappeared. Similarly, in their Table 5, controlling for relative prices in panels B-D removed the association between pure inflation and GDP. The data and code were downloaded from *American Economic Journal: Macroeconomics*' website. 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 had 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. We classified this replication as successful.

7.61 Romer and Romer (2010), "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks"

Dependent variables: ln(real GDP), Tax Changes.

Key Results: "The behavior of output following these more exogenous [tax] changes indicates that tax increases are highly contractionary." (abstract)

We took the key figures as their Figures 6, 7, and 9. Due to data constraints, we only replicated panel B of Figure 7, although we had the data to replicate all of their Figures 6 and 9. We were able to replicate exactly their Figures 6, 9, and panel B of their Figure 7 using code and data from the *American Economic Review*'s website without assistance from the authors using RATS 7.10 (Linux). We classified this replication as successful.

7.62 Rudebusch and Wu (2008), "A Macro-Finance Model of the Term Structure, Monetary Policy and the Economy"

Dependent variables: Various yield curve parts, industrial capacity utilization, 12-month percentage change in PCE prices.

Key Results: "(1) the latent term structure factors from no-arbitrage finance models appear to have important macroeconomic and monetary policy underpinnings, (2) there is no evidence of a slow partial adjustment of the policy interest rate by the central bank, and (3) both forward-looking and backward-looking elements play roles in macroeconomic dynamics."(abstract)

In preparing replication archives, we realized that, for estimating their Taylor Rule, Rudebusch and Wu (2008) used capacity utilization instead of GDP minus potential GDP as the measure of the output gap. Therefore, this paper should not have been included in our original sample. However, because we classified the paper as in-sample during our preanalysis plan and did not discover the classification error until after our results were finalized, we are keeping this paper in our analysis.

We interpreted the key findings to Figures 2-7 and Table 2. We acquired a replication archive from Glenn Rudebusch, but the code was missing dependencies. A supplemental file was obtained from Tao Wu, but we were still unable to run the code in the archive. For these reasons, we classified this paper as "failed due to incorrect public data or code."

7.63 Schmitt-Grohé and Uribe (2011), "Business Cycles with a Common Trend in Neutral and Investment-Specific Productivity"

Dependent variables: per capita real GDP growth, per capita real nondurables and services consumption growth, per capita investment growth, ln(nonfarm business hours).

Key Results: "the common stochastic trend in neutral and investment-specific productivity plays a sizable role in driving business cycles." (introduction, pg. 123)

We used the data and code files from the *Review of Economic Dyanmics*' 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). We classified this replication as successful.

7.64 Schmitt-Grohé and Uribe (2012), "What's News in Business Cycles"

Dependent variables: First-difference natural logs of: real GDP per capita, real consumption per capita, real investment per capita, real government expenditure per capita, hours per capita, TFP, and the relative price of investment. Variance of: real GDP per capita, real consumption per capita, real investment per capita, and hours per capita.

Key Results: "anticipated shocks account for about half of predicted aggregate fluctuations in output, consumption, investment, and employment." (abstract)

We used the replication files on *Econometrica*'s website, which produced exactly the author's results without assistance using Matlab R2008a (Windows). We classified this replication as successful.

7.65 Senyuz (2011), "Factor Analysis of Permanent and Transitory Dynamics of the US Economy and the Stock Market"

Dependent variables: Recessions, Recession Probabilities.

Key Results: "all 10 recessions in the post-war sample... are identified by the permanent and transitory components" (introduction, pg. 977)

We did not have the correct software to run the code for this paper. We classified this paper as "failed due to missing software."

7.66 Shore (2010), "For Better, For Worse: Intrahousehold Risk-Sharing Over the Business Cycle"

Dependent variables: Cross-sectional variance of income, Cross-sectional covariance of Couples' Incomes.

Key Results: "husbands' and wives' income changes are more positively correlated when the economy is growing rapidly." (abstract)

This result corresponded to Tables 1 and 2. For our Finance and Economics Discussion series version of this paper, Chang and Li (2015a), we mistakenly thought that we lacked the files to replicate Shore (2010). However, after preparing our replication datasets for the *Critical Finance Review* we realized that the archive that we had for Shore (2010) was complete, which we were able to use to exactly replicate Shore (2010) using Stata SE 11.0 (Windows). The files came directly from Stephen Shore. Our Sankey diagrams in Figures 1 and 2 classify Shore (2010) as "failed due to incorrect public data or code," as that was what we believed in our working paper. The heterogeneity analysis in subsections 5.2 to 5.4 classify Shore (2010) as successful.

7.67 Šustek (2011), "Monetary Business Cycle Accounting"

Dependent variables: 3-month Treasury Yield, GDP Deflator.

Key Results: "the [total factor productivity] efficiency and asset market wedges are both necessary, and to some extent also sufficient, for generating the observed lag-lead pattern of [nominal inflation and the interest rate]. The other four wedges are less important." (introduction, pg. 593)

The key results were Figures 2-4. Using data and code from *Review of Economic Dynamics*' website, we were unable to fully replicate the results even with assistance from the author. Our initial attempts led to nonsensical parameter estimates. For example, one of the relative standard deviations was estimated to be 459987561354116. We believed that this was primary due to the simulated annealing algorithm not converging to a global maximum, and we ultimately could not get the simulated annealing algorithm to work. However, using the model parameter estimates in the published paper without running the simulated annealing estimation procedure, we were able to replicate the "no efficiency wedge" panels in Figures 2 and 3. For these reasons, we classified this paper as "failed due to incorrect public data or code."