Internet Appendix to
"Impulse Response Functions for Self-Exciting Nonlinear Models"
by Neville Francis, Michael T. Owyang, and Daniel Soques

A Additional Simulation Results

A.1 Example Simulation IRFs

To illustrate the dynamics of our simulation exercise, this section presents a few examples of different data generating processes for a single parameterization. Figure 1 shows the posterior median IRF for each estimation method when the true data generating process is a TVAR. Similarly, Figure 2 shows the estimated responses when the truth is a STVAR. Note that these example plots are for a single parameterization of the model.

A.2 Horizon-Specific CRPS

In the main text of the paper, we focused on the average CRPS across response horizons h to diagnose performance of each estimation method. However, it may be that some methods perform better at certain horizons than others. Figure 3 shows the CRPS at each horizon for each estimation method under the various forms of model misspecification. These results largely mimic the average CRPS results, though there are some instances where a method has relatively better performance at shorter or longer horizons. In almost all instances, the MA-GIRF outperforms all but the true data generating process. The only exception with regard to LP is when the VAR lag is truncated.

A.3 State Asymmetry

In the class of models we study, the IRFs will depend on state of the economy at the time of the shock. State Asymmetry (SA) measures the difference in the responses across the initial regime at each horizon, h:

$$SA(h, \delta, \Theta) = \Phi^{0}_{GIRF}\left(h, \delta, \Theta\right) - \Phi^{1}_{GIRF}\left(h, \delta, \Theta\right),$$

where the superscripts indicate the state of the economy on impact. Evaluated at $\widetilde{\Theta}$, $SA(h, \delta, \widetilde{\Theta})$ is the true degree of state asymmetry, while $SA(h, \delta, \Theta))$ is computed by averaging across Gibbs draws. We define $DSA\left(\delta, \Theta\right)$ as the sum of the squared difference between the true response and the average SA across horizons [see also Karamé (2015)]:

$$DSA\left(\delta, \Theta\right) = \sum_{h} \left[SA(h, \delta, \Theta) - SA(h, \delta, \widetilde{\Theta})\right]^{2}.$$

A lower $DSA\left(\delta, \Theta\right)$ suggests the estimated state asymmetry for a given model is closer to the true degree of state asymmetry. Table 1 shows the DSA results for our simulation exercise under various forms of model misspecification. The models which tend to have lower bias (i.e., lower CRPS) also tend to have lower DSA.

Figure 1: **Example IRFs From TVAR Simulations** – This figure shows an example IRF for a single simulation. For each model, we plot the posterior median response at each horizon. The true model is a TVAR.

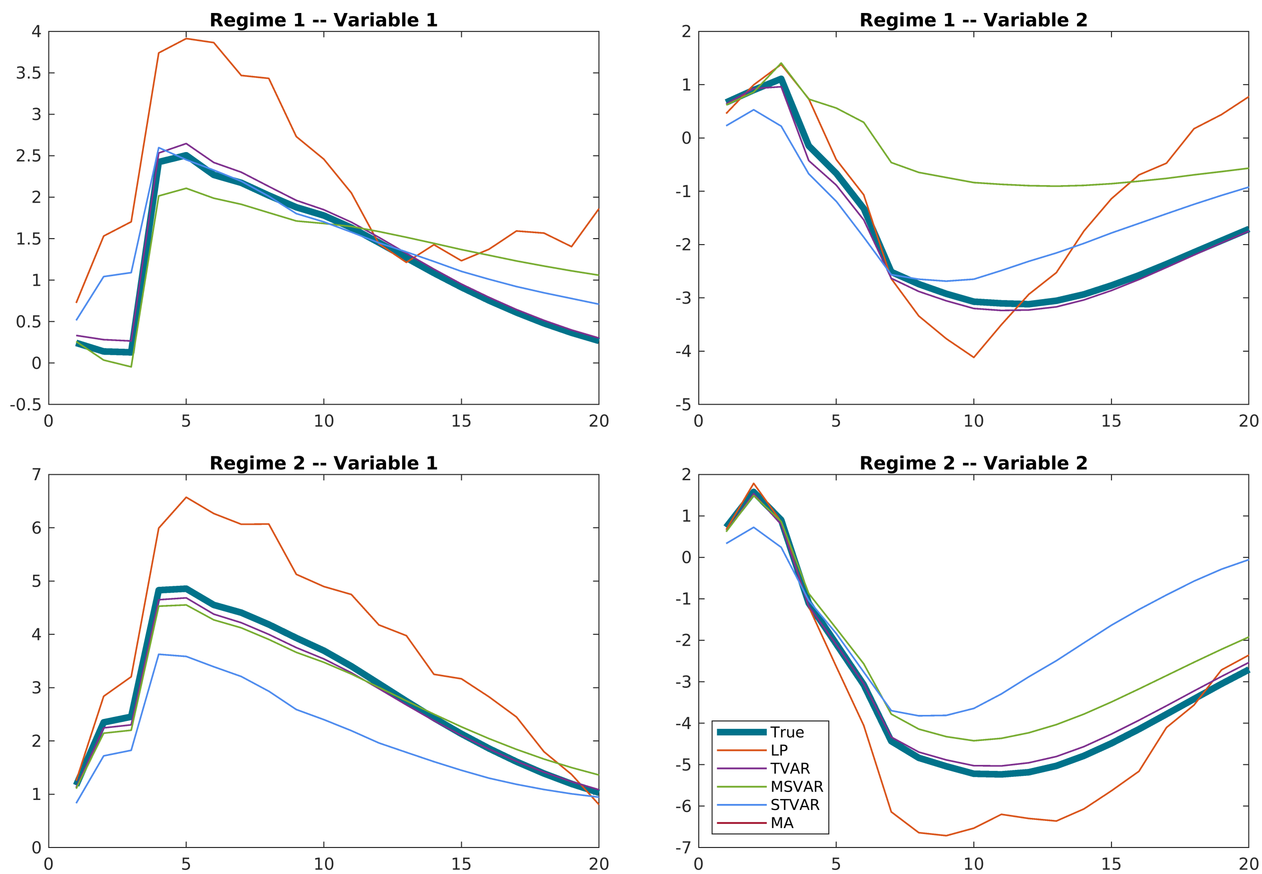


Figure 2: **Example IRFs From STVAR Simulations** – This figure shows an example IRF for a single simulation. For each model, we plot the posterior median response at each horizon. The true model is a STVAR.

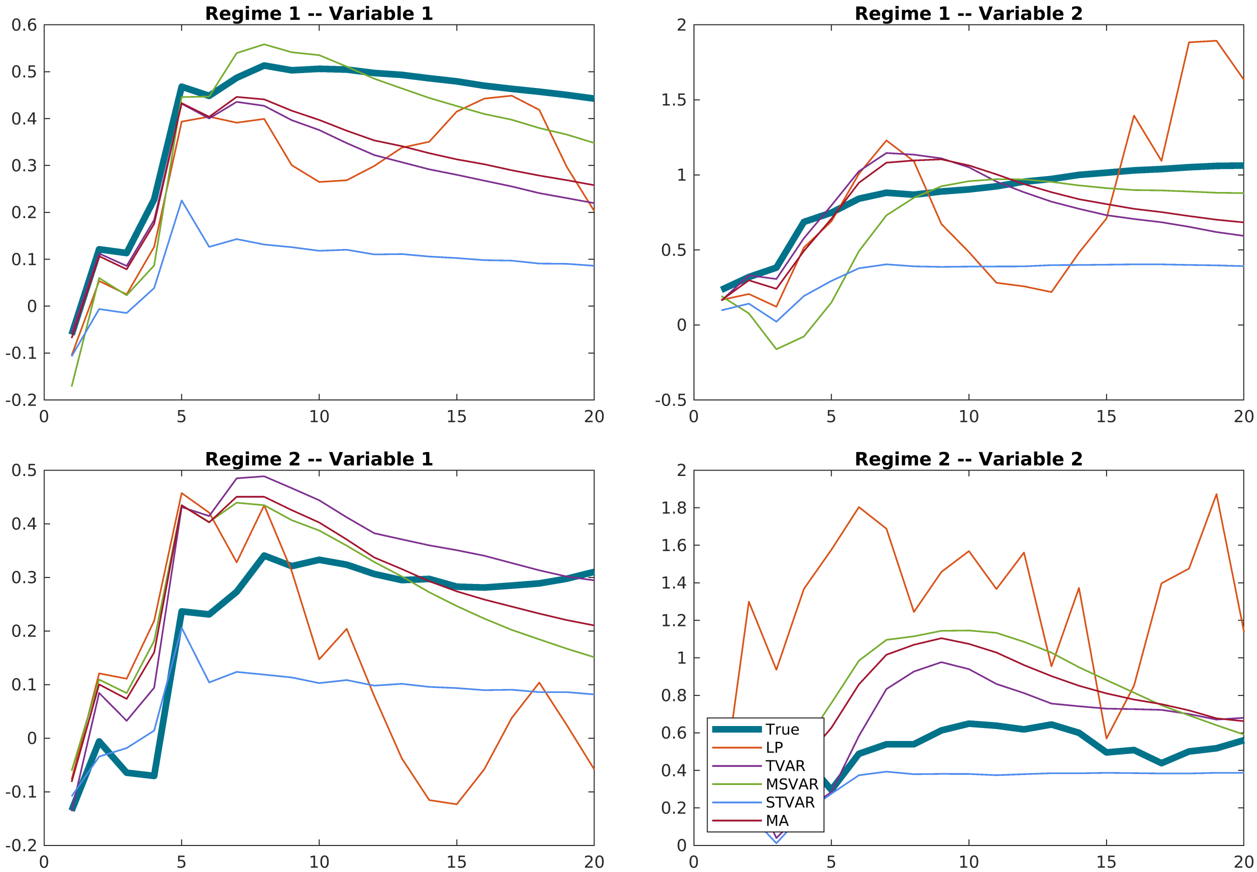
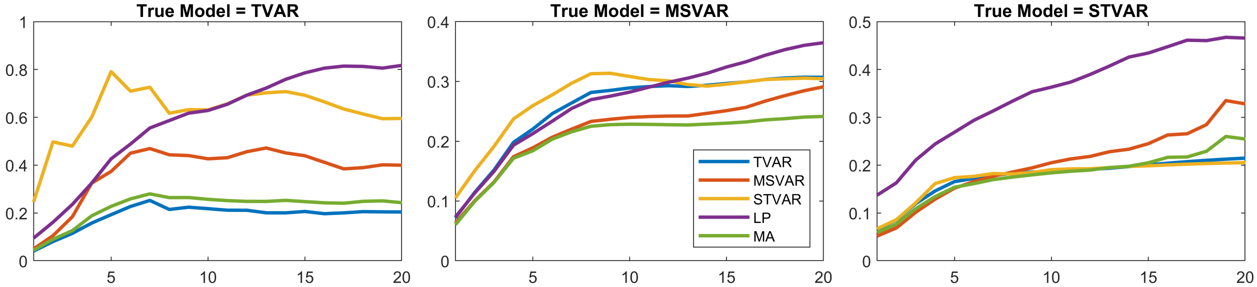
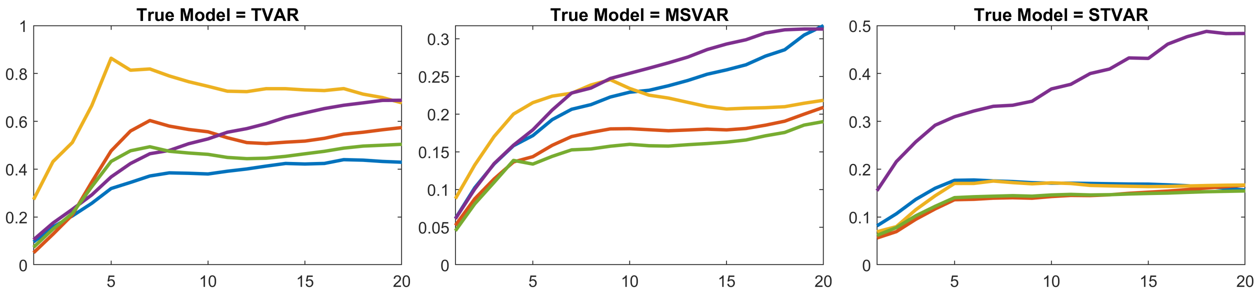


Figure 3: **CRPS By Horizon For Various Misspecifications** – This figure shows the CRPS across impulse response horizons for various model misspecifications.

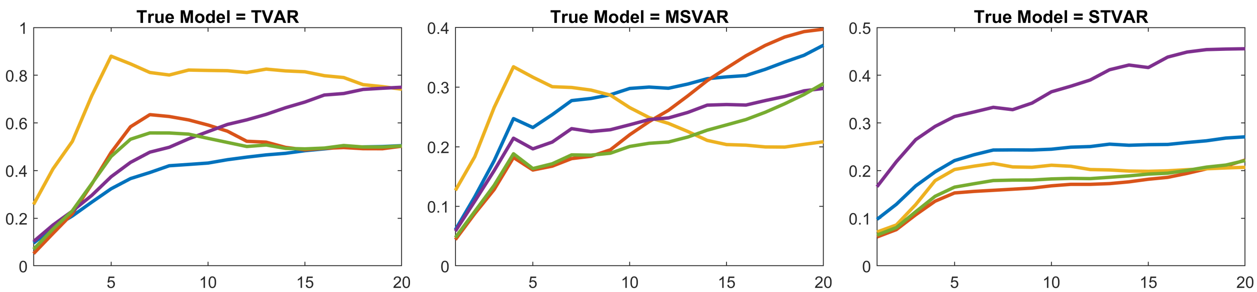
(a) Transition Process Only



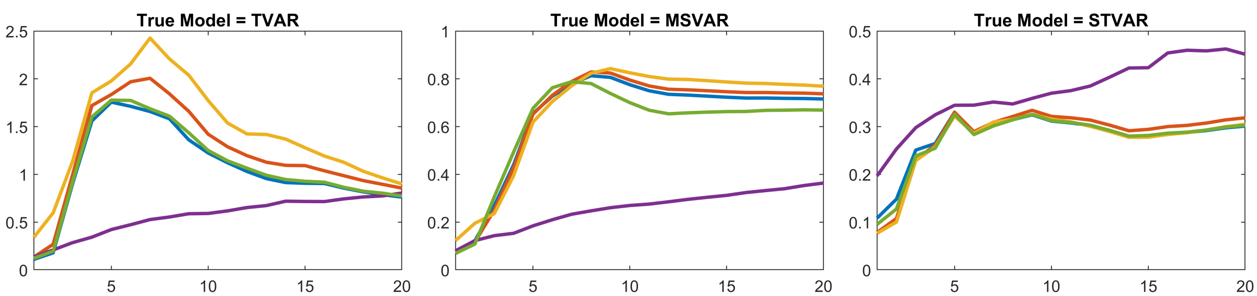
(b) Incorrect Lag on Transition Variable



(c) Incorrect Transition Variable



(d) Incorrect VAR Lag (1)



(e) Incorrect VAR Lag (6)

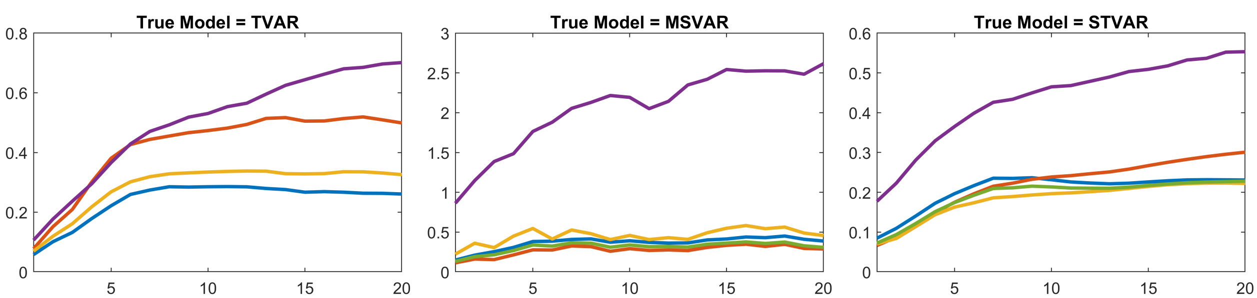


Table 1: **DSA Across Simulation Specifications** – This table shows the median Deviation in State Asymmetry (DSA) for each estimation method for a given true data generating process and specification error. The 68% posterior interval is shown in brackets beneath the median.

Specification Error:	**True Model:**	**Estimated Model:**				
		TVAR	*MSVAR*	*STVAR*	*LP*	*MA*
Transition	TVAR	5.86 [1.32,25.98]	29.79 [5.54,251.30]	18.54 [4.92,58.81]	70.65 [26.74,206.05]	9.79 [1.93,46.62]
Process	MSVAR	10.37 [2.35,48.83]	15.21 [1.06,74.41]	13.77 [3.49,48.75]	15.97 [5.64,57.16]	9.65 [2.12,46.25]
Only	STVAR	3.84 [0.88,16.51]	3.60 [1.17,20.47]	1.76 [0.84,5.72]	33.44 [8.13,129.71]	3.00 [0.75,16.07]
Incorrect Lag	TVAR	22.53 [6.21,92.12]	42.76 [5.48,356.48]	42.51 [9.28,143.18]	68.57 [23.11,222.21]	30.15 [5.51,150.86]
on Transition	MSVAR	15.08 [3.11,46.02]	6.26 [0.93,34.70]	13.69 [2.61,29.10]	14.59 [5.76,33.21]	9.43 [1.61,31.51]
Variable	STVAR	3.14 [0.98,10.08]	2.43 [0.67,7.50]	1.58 [0.41,2.93]	48.13 [6.81,169.90]	1.91 [0.64,5.83]
Incorrect	TVAR	39.42 [10.25,179.06]	38.58 [7.09,306.74]	46.66 [10.33,158.37]	88.91 [30.06,266.75]	40.53 [7.30,226.92]
Transition	MSVAR	9.05 [2.65,33.13]	6.24 [1.60,39.48]	8.18 [2.00,19.77]	13.75 [5.77,29.00]	7.22 [1.76,35.84]
Variable	STVAR	2.73 [0.80,10.78]	2.63 [0.82,9.37]	1.59 [0.39,3.58]	31.78 [8.09,126.67]	2.21 [0.73,7.58]
Incorrect	TVAR	31.93 [10.87,106.65]	57.81 [13.38,192.98]	49.46 [13.90,145.34]	57.79 [22.22,221.25]	33.06 [10.87,99.84]
VAR	MSVAR	11.56 [2.79,53.47]	11.34 [2.78,54.35]	13.27 [4.50,43.02]	17.43 [5.90,58.29]	9.85 [2.81,44.41]
Lag (1)	STVAR	5.26 [1.22,23.34]	4.01 [0.94,25.09]	2.69 [0.72,19.37]	47.86 [13.06,176.90]	3.90 [0.90,22.09]
Incorrect	TVAR	6.49 [1.42,42.02]	49.78 [4.26,297.77]	32.98 [7.49,120.41]	68.40 [23.94,250.21]	16.91 [2.50,81.76]
VAR	MSVAR	47.09 [24.25,90.06]	38.86 [12.94,76.30]	30.18 [17.97,64.08]	341.89 [47.41,3428.27]	43.47 [24.27,75.37]
Lag (6)	STVAR	5.11 [1.49,25.03]	4.51 [1.20,30.76]	2.16 [0.70,14.55]	52.23 [14.31,204.24]	3.75 [0.96,21.16]
True Model is Linear	VAR	24.07 [3.21,196.96]	9.76 [2.15,32.85]	26.49 [10.38,49.20]	22.39 [7.84,60.01]	13.87 [2.17,117.20]

B Additional Application Results

B.1 Model Averaging Multiplier Estimates

In the baseline results, the model average multiplier is computed as the weighted GIRF of output divided by the weighted GIRF of government spending. Alternatively, we could calculate the model average multiplier by applying the weights directly to the multiplier from each respective model. Tables 2 and 3 show that results are qualitatively similar using this calculation.

Table 2: **Fiscal Spending Multipliers: Full Sample (Model Average Over Multipliers)** – This table shows the median posterior five-year fiscal multiplier for each model under each business cycle regime. The 68% highest posterior density interval is shown below in brackets. The multiplier is computed as the cumulative response of output divided by the cumulative response of government spending. The first and second columns show the multiplier based on the Ramey news shock series under slack and nonslack regimes, respectively. The third and fourth columns show similar estimates when using Blanchard-Perotti shock identification. In each case the shock size is a one-percent of GDP increase in government spending. Four lags of Y_t are used in both the VAR and LP specifications. The last row shows the multiplier implied by the model average impulse response when averaging over all three model specifications (TVAR, MSVAR, and STVAR) as well as both shocks (Ramey News and Blanchard-Perotti). The model average multipliers are computed using the individual model's multiplier estimates rather than the underlying GIRFs.

Sample:	1890Q1-2015Q4			
Shock ID:	Ramey News		Blanchard-Perotti	
	Slack	Nonslack	Slack	Nonslack
TVAR	0.49 [0.15,0.87]	0.59 [0.37,0.83]	0.42 [0.13,0.68]	0.34 [0.18,0.51]
MSVAR	0.93 [0.57,1.34]	0.67 [0.33,1.00]	0.59 [0.45,0.73]	0.32 [0.13,0.50]
STVAR	0.98 [0.45,1.54]	0.90 [0.35,1.41]	0.49 [0.35,0.66]	0.29 [0.20,0.40]
Model Avg	0.53 [0.18,0.92]	0.59 [0.38,0.83]	0.46 [0.22,0.65]	0.31 [0.18,0.47]
Local Proj	0.53 [0.46,0.61]	0.68 [0.55,0.81]	0.79 [0.72,0.87]	0.31 [0.24,0.37]
		Slack	Nonslack	
Model Avg (Combined)		0.46 [0.23,0.67]	0.35 [0.20,0.53]	

Table 3: **Fiscal Spending Multipliers: Short Sample (Model Average Over Multipliers)** – This table shows the median posterior five-year fiscal multiplier for each model under each business cycle regime. The 68% highest posterior density interval is shown below in brackets. The multiplier is computed as the cumulative response of output divided by the cumulative response of government spending. The first and second columns show the multiplier based on the Ramey news shock series under slack and nonslack regimes, respectively. The third and fourth columns show corresponding estimates when using Blanchard-Perotti shock identification. In each case the shock size is a one-percent of GDP increase in government spending. Four lags of Y_t are used in both the VAR and LP specifications. The last row shows the multiplier implied by the model average impulse response when averaging over all three model specifications (TVAR, MSVAR, and STVAR) as well as both shocks (Ramey News and Blanchard-Perotti). The model average multipliers are computed using the individual model's multiplier estimates rather than the underlying GIRFs.

Sample:	1969Q1-2015Q4			
Shock ID:	Ramey News		Blanchard-Perotti	
	Slack	Nonslack	Slack	Nonslack
TVAR	1.49 [−3.68,7.44]	1.88 [−7.94,11.37]	0.83 [−0.05,1.71]	0.02 [−1.19,1.12]
MSVAR	0.67 [−8.33,11.00]	0.97 [−5.62,7.74]	0.95 [−0.11,2.07]	1.06 [−0.01,2.20]
STVAR	0.97 [−3.93,5.72]	2.26 [−11.25,14.68]	1.53 [0.68,2.45]	0.33 [−0.55,1.11]
Model Avg	1.27 [−4.24,7.36]	1.64 [−8.34,11.57]	0.84 [0.02,1.71]	0.28 [−0.86,1.35]
Local Proj	−0.24 [−0.95,0.48]	−2.00 [−8.15,2.68]	0.56 [0.34,0.78]	−2.59 [−4.04,−1.49]
		Slack	Nonslack	
Model Avg (Combined)		0.85 [−0.17,1.94]	0.30 [−1.23,1.75]	

B.2 Additional Robustness Results

In the main text, we check the robustness of fiscal spending multiplier estimates when changing the number of VAR lags or the transition variable. Those results used Blanchard-Perotti shock identification due to the poor identification of the Ramey news shock series post-Korean War. Table 4 shows the fiscal multiplier estimates when using different VAR lags both under the full and short sample. Table 5 shows the sensitivity of estimates when using the unemployment rate rather than the output gap as the transition variable in the full sample. When using the full sample, the results mostly match those when using BP identification. However, the short sample estimates give unreliable point estimates with large error bands, again indicating the Ramey shocks are a poor instrument during this subperiod.

Table 6 shows the comprehensive multiplier estimate when calculating the model average over all transition functions, both shocks, all VAR lag lengths, and the two transition variables. In both samples, the point estimates are larger in times of slack compared to normal times, however the error bands overlap. As in the main text, this provides further evidence that the multiplier is less than one and not substantially different in times of slack and nonslack.

Table 4: **Fiscal Spending Multipliers: Different VAR Lags (Ramey Shocks)** – This table shows the median posterior five-year fiscal multiplier for each model under each business cycle regime for various lag lengths of Y_t in both the VAR and LP specifications. The 68% highest posterior density interval is shown below in brackets. The multiplier is computed as the cumulative response of output divided by the cumulative response of government spending. In each case the shock size is a one-percent of GDP increase in government spending. The last row shows the multiplier implied by the model average impulse response when averaging over all three model specifications (TVAR, MSVAR, and STVAR) as well as all three lag lengths (1, 4, or 6).

(a) Full Sample

Shock ID:	Ramey News					
Sample:	1890Q1-2015Q4					
VAR Lags	1		4		6	
	Slack	Nonslack	Slack	Nonslack	Slack	Nonslack
TVAR	0.96 [0.58,1.42]	0.75 [0.60,0.95]	0.49 [0.15,0.87]	0.59 [0.37,0.83]	0.55 [0.22,0.87]	0.61 [0.42,0.81]
MSVAR	0.75 [0.60,0.90]	0.64 [0.48,0.80]	0.93 [0.57,1.34]	0.67 [0.33,1.00]	1.34 [1.09,1.61]	1.39 [1.14,1.70]
STVAR	0.91 [0.60,1.76]	0.79 [0.69,0.91]	0.98 [0.45,1.54]	0.90 [0.35,1.41]	0.75 [0.44,1.52]	0.66 [0.47,0.97]
Model Avg	0.81 [0.61,1.14]	0.70 [0.55,0.84]	0.53 [0.17,0.92]	0.59 [0.38,0.83]	0.70 [0.35,1.24]	0.69 [0.47,1.25]
Local Proj	0.68 [0.62,0.73]	0.74 [0.63,0.85]	0.53 [0.46,0.61]	0.68 [0.55,0.82]	0.53 [0.44,0.63]	0.78 [0.63,0.94]
			Slack	Nonslack		
Model Average (All VAR Lags)			0.79 [0.59,1.11]	0.69 [0.55,0.84]		

(b) Short Sample

Shock ID:	Ramey News					
Sample:	19690Q1-2015Q4					
VAR Lags	1		4		6	
	Slack	Nonslack	Slack	Nonslack	Slack	Nonslack
TVAR	0.14 [−4.46,3.58]	−2.80 [−17.89,18.93]	1.49 [−3.68,7.44]	1.88 [−7.94,11.37]	1.52 [−3.15,5.58]	0.37 [−6.41,7.31]
MSVAR	2.54 [−5.89,4.66]	2.04 [−5.76,4.16]	0.67 [−8.33,11.00]	0.97 [−5.62,7.74]	−2.60 [−3.36,−2.20]	−2.56 [−3.40,−2.17]
STVAR	1.26 [−5.32,7.81]	0.47 [−7.11,7.87]	0.97 [−3.93,5.72]	2.26 [−11.25,14.68]	1.42 [−4.48,7.53]	0.34 [−6.50,7.14]
Model Avg	−0.17 [−5.94,4.45]	−1.34 [−14.27,13.96]	1.27 [−4.24,7.37]	1.65 [−8.39,11.63]	−2.59 [−3.35,−2.17]	−2.55 [−3.39,−2.16]
Local Proj	−0.62 [−1.69,0.37]	1.08 [−0.47,3.07]	−0.24 [−0.97,0.47]	−2.01 [−8.02,3.03]	0.40 [−0.53,1.39]	−0.50 [−15.21,14.36]
			Slack	Nonslack		
Model Average (All VAR Lags)			−2.50 [−3.36,−1.94]	−2.44 [−3.33,−2.02]		

Table 5: **Fiscal Spending Multipliers: Different Transition Variables (Ramey Shocks)** – This table shows the median posterior five-year fiscal multiplier for each model under each business cycle regime for the full sample of 1890Q1 - 2015Q4. The 68% highest posterior density interval is shown below in brackets. The multiplier is computed as the cumulative response of output divided by the cumulative response of government spending. The first and second columns show the multiplier under lack and nonslack regimes, respectively, when the transition variable is the output gap. The third and fourth columns show similar estimates when using the unemployment rate as the transition variable. In each case the shock size is a one-percent of GDP increase in government spending. The top panel uses the Ramey news shock series and the bottom uses Blanchard-Perotti. Four lags of Y_t are used in both the VAR and LP specifications. The last row shows the multiplier implied by the model average impulse response when averaging over all three model specifications (TVAR, MSVAR, and STVAR) as well as both transition variables (the output gap and unemployment rate).

Sample:	1890Q1-2015Q4			
Shock ID:	Ramey News			
Transition Variable:	Output Gap		UR	
	Slack	Nonslack	Slack	Nonslack
TVAR	0.49 [0.15,0.87]	0.59 [0.37,0.83]	0.76 [0.55,0.97]	0.90 [0.53,1.46]
MSVAR	0.93 [0.57,1.34]	0.67 [0.33,1.00]	0.87 [0.55,1.17]	0.69 [0.47,0.90]
STVAR	0.98 [0.45,1.54]	0.90 [0.35,1.41]	0.72 [0.59,0.86]	0.90 [0.68,1.13]
Model Avg	0.53 [0.17,0.92]	0.59 [0.38,0.83]	0.78 [0.56,1.02]	0.80 [0.52,1.24]
Local Proj	0.53 [0.46,0.61]	0.68 [0.55,0.81]	0.64 [0.59,0.70]	0.48 [0.32,0.65]
		Slack	Nonslack	
Model Avg (Both Transition Var.)		0.60 [0.25,0.97]	0.64 [0.42,0.87]	

Table 6: **Comprehensive Model Average Mutiplier** – This table shows the median posterior five-year fiscal multiplier for the comprehensive model average under each business cycle regime. The 68% highest posterior density interval is shown below in brackets. The comprehensive multiplier is the multiplier implied by the model average impulse response when averaging over all three model specifications (TVAR, MSVAR, and STVAR), both shocks (Ramey News and Blanchard-Perotti), VAR lag lengths (1,4, and 6), and transition variables (output gap and unemployment rate). The full sample (1890:Q1-2015:Q4) and short sample (1969:Q1-2015:Q4) results are shown in the top and bottom rows, respectively.

	Slack	Nonslack
Full Sample	0.64 [0.41,0.91]	0.54 [0.32,0.73]
Short Sample	0.40 [−0.10,0.84]	0.26 [−0.18,0.70]