Cross-sectional Identification with Heterogeneous Exposure to General Equilibrium Effects

Paula Donaldson*
UC San Diego
Job Market Paper

This version: April 2026
Click here for the latest version.

Cross-sectional designs that exploit heterogeneous exposure to aggregate shocks while controlling for time fixed effects are widely used to estimate elasticities of macroeconomic importance, such as cross-sectional fiscal multipliers. I show that these designs fail to identify partial-equilibrium elasticities when exposure to the shock of interest is correlated with exposure to other aggregate variables that move in general equilibrium. I develop a test for this identification failure and propose a new decomposition method that leverages both cross-sectional and time-series variation to recover elasticities purged of general-equilibrium effects. Applying the method to estimate US cross-sectional fiscal multipliers, I find that accounting for the general equilibrium effects operating through monetary policy reduces the estimated two-year multiplier from 1.5 to 1. This result shows that monetary policy can bias cross-sectional fiscal multiplier estimates and demonstrates the need for cross-sectional identification strategies that account for heterogeneity in responses to general-equilibrium forces.

JEL: E52, E62, C32, C33.

---

*Email: m1donald@ucsd.edu.
I am grateful to my advisor, Johannes Wieland, and to my committee members Juan Herreño, Nir Jaimovich, Munseob Lee, Valerie Ramey, Fabian Trottner, and Steve Wu for their invaluable guidance and support. I also thank Titan Alon, Francesco Amodeo, Bence Bardóczy, Saki Bigio, Edoardo Briganti, Dario Caldara, Fabian Eckert, Francesco Ferrante, Giuseppe Fiori, Jim Hamilton, Xinwei Ma, Thomas Mertens, Gaston Navarro, William Peterman, Immo Schott, Yannick Timmer (discussant), Ali Uppal, Christian Wolf, Kaspar Wuthrich, and seminar participants at UCSD, Midwest Macro, the Federal Reserve Board, SED, CEMLA, SOCAE, Surrey, UC3M, Yale, IADB, Norges Bank, IIES, EUI, Banca d'Italia, FEN UChile, ITAM, Brown, LSE, UCL, UPF-CREI, LBS and Banco de España for helpful comments and discussions. The paper also benefited from valuable feedback received during my Dissertation Fellowship at the Federal Reserve Board. A previous version of this paper circulated as "Estimation of Portable Statistics when Agents are Heterogeneous".

1. Introduction

Cross-sectional methods are increasingly used in macroeconomics, international and spatial economics to study how the economy responds to aggregate shocks. When households, firms, or regions differ in their exposure to these shocks, a comparison of their relative responses recovers the elasticity of the cross-sectional outcome to the aggregate shock. Examples include fiscal, trade, and monetary policy, among other economic settings in which heterogeneous exposure to aggregate shocks can be exploited for identification.[1]

These methods are designed to identify partial-equilibrium, or micro-level, elasticities. This requires separating the direct effect of the shock from the aggregate effects it sets off in general equilibrium. The typical empirical strategy combines time fixed effects, which absorb the aggregate effects, with cross-sectional differences in exposure, which provide the variation needed to isolate the direct effect. Throughout this paper, I refer to these cross-sectional elasticities as *portable*[2] elasticities, echoing the terminology introduced by Nakamura and Steinsson (2018). Such elasticities are valuable for their external validity and their role in policy evaluation and counterfactual analysis. By isolating well-identified responses to shocks, they enable progress on important questions that aggregate time series analysis alone cannot resolve.

But this strategy rests on a strong identifying assumption: that exposure to the shock is uncorrelated with sensitivity to aggregate variables that co-move in general equilibrium. In practice, the same characteristics that drive exposure to one policy often also shape sensitivity to others. For example, regions more exposed to fiscal expansions may also be more or less sensitive to interest rates, which typically respond to fiscal policy. In such settings, cross-sectional variation in exposure does not isolate the direct effect of the shock. Identification fails because the estimated elasticity is contaminated by heterogeneous responses to dynamic general equilibrium effects. I refer to such environments as HEDGE settings, short for *heterogeneous exposure to dynamic general equilibrium effects*. This paper asks: How can portable elasticities be reliably identified in such environments?

I provide a new framework that makes it possible to identify portable elasticities in HEDGE economies. The approach combines elements from cross-sectional and time-series analysis to decompose the elasticity estimated using the typical two-way fixed effects (TWFE) design into the portable elasticity of interest and a HEDGE term. I use the

[1] Some influential examples include fiscal and tax policy (Parker et al. 2013; Nakamura and Steinsson 2014; Zidar 2019; Pennings 2021), credit supply (Mian and Sufi 2014), monetary policy (Ottonello and Winberry 2020; Peydró, Polo, and Sette 2021), trade shocks (Autor, Dorn, and Hanson 2013, 2021), investment elasticities (Zwick and Mahon 2017), and wealth effects (Guren et al. 2021; Chodorow-Reich, Nenov, and Simsek 2021), among others.

[2] I use the term *portable* to describe elasticities estimated at different levels of aggregation that capture the cross-sectional, direct effect of an aggregate shock while holding other aggregates fixed. Depending on the context, this may correspond to an individual optimization problem (e.g., marginal propensities to consume) or to a partial-equilibrium elasticity where some local market clears (e.g., cross-sectional multipliers).

framework to estimate cross-sectional fiscal multipliers in the US, allowing states to differ in their exposure to changes in both defense spending and interest rates. The two-year multiplier drops from 1.5 to 1 when applying the decomposition. This finding challenges the view that TWFE estimates of cross-sectional fiscal multipliers are independent of the monetary stance. The estimated portable multiplier is smaller for two reasons: (i) US states that are more exposed to defense spending are also less sensitive to monetary policy shocks, and (ii) monetary policy responds to an expansionary defense spending shock by tightening interest rates. Finally, I develop a two-region TANK model, combining features of Nakamura and Steinsson (2014) and Herreno and Pedemonte (2025), to interpret the empirical findings. The model analogue of the TWFE multiplier can vary sharply across monetary regimes once HEDGE is present, challenging its usefulness for bounding the aggregate multiplier. In contrast, the portable multiplier maintains its ability to provide upper and lower bounds on the aggregate effects of fiscal policy shocks across monetary regimes.

The Identification Challenge. First, I establish the conditions under which HEDGE represents an identification challenge to elasticities estimated using the following TWFE specification:

$$Y_{it+h} = \beta_h s_{ix} \epsilon_{xt} + \lambda_{ih} + \lambda_{th} + e_{it+h}, \tag{1}$$

where Y_{it} is the outcome of unit i at time t (e.g., local GDP), s_{ix} denotes the exposure of unit i to an aggregate variable X, and ϵ_{xt} is an innovation to X in period t. Consider an economy[3] where units differ in their exposure to two aggregate variables: G_t, the variable of interest (e.g., government spending), and R_t (e.g., monetary policy rate). Then it can be shown that the estimate of β_h from a TWFE regression like (1) converges to:

$$\hat{\beta}_h \to \beta_h^P + \gamma \underbrace{\frac{cov\left[s_{ig}, s_{ir}\right]}{v[s_{ig}]}}_{\phi} \underbrace{\frac{cov\left[R_{t+h}, \epsilon_{gt}\right]}{v[\epsilon_{gt}]}}_{\boldsymbol{R}_{g,h}}, \tag{2}$$

where ϵ_{gt} is an innovation to G_t, s_{ik} is unit i's exposure to $k \in \{G, R\}$ and γ measures the differential response to R, conditional on s_{ir}. β_h^P is the portable elasticity at horizon h.[4]

This expression highlights that two objects determine whether the endogenous response of R to ϵ_{gt} is fully differenced out. First, a cross-sectional term—denoted by ϕ—that measures whether exposure to G is correlated with exposure to R. For example, are regions

[3] Concretely, consider a unit-level outcome that evolves according to: $Y_{it} = \beta \times s_{ig} G_t + \gamma \times s_{ir} R_t + u_t + u_i + u_{it}$.

[4] *Hovering the cursor over frequently used in-text mathematical symbols displays brief descriptions of their meaning. This feature is supported only in Adobe Acrobat or PDF-XChange Viewer. They will not appear in most browser-based and (macOS) Preview viewers.*

that receive more military spending also more or less sensitive to interest rate changes? The cross-sectional term will be different from zero when the sensitivities of economic units to G and R are correlated in the cross-section.

Second, a time-series object—denoted by $\boldsymbol{R}_{g,h}$—that captures the impulse response of R to the shock of interest, ϵ_{gt}, at horizon h. For example, in a fiscal setting, this could represent the response of monetary policy to government spending or tax policy shocks. The time series term will be different from zero when the shock of interest triggers an endogenous general equilibrium (GE) response in R.

The TWFE specification implicitly assumes that the cross-sectional heterogeneity used for identification only drives exposure to the shock of interest and not to any aggregate variable that the researcher intends to difference out. This is a strong assumption for many empirical applications where exposure shifters often reflect structural characteristics of economic agents (e.g., demographics, industrial composition, financial constraints). It is plausible that such characteristics drive the response to multiple aggregate variables.

This poses a clear identification challenge: we cannot rely on time-fixed effects alone to control for general equilibrium responses. The estimated coefficients reflect not only the direct effects of the shock of interest but also the effects through the endogenous response in other aggregate variables:

$$\hat{\beta}_h \to \beta_h^P + \underbrace{\gamma \times \phi \times \boldsymbol{R}_{g,h}}_{\Omega_h}. \tag{3}$$

This second component, which I label the HEDGE term and denote by Ω_h, acts as an additional treatment effect. It measures how the endogenous and dynamic response of R to ϵ_{gt} affects the outcome of units that are relatively more exposed to G. Specifically, for a given general equilibrium response $\boldsymbol{R}_{g,h}$, the contribution of the HEDGE term depends on the covariance between exposure to the shock of interest and exposure to the general equilibrium response.

This creates two concerns. First, the sign of the bias introduced by HEDGE is not dictated by economic theory alone, but by the sign of ϕ, the scaled covariance between exposure shifters in the cross-section. As a result, it is possible for general equilibrium forces that attenuate the aggregate effect of a shock, such as contractionary monetary policy in response to a fiscal expansion, to increase the estimated cross-sectional effect, depending on the structure of the data. Second, exposure shifters used in different studies can be correlated with sensitivities to different aggregate variables, leading to variation in estimated effects even when studying the same aggregate shock. These issues highlight the importance, both for interpretation and for empirical consistency, of separating the direct effect from the component driven by HEDGE.

Decomposition Framework. I develop a framework that brings in additional information to identify portable elasticities in HEDGE economies. Its goal is to decompose $\hat{\beta}_h$ in expression (3) into a portable component β_h^P and an HEDGE component Ω_h. Specifically, I develop an estimation procedure for the latter term. The framework is based on the results in McKay and Wolf (2023) and Barnichon and Mesters (2023), who show that reduced-form impulse responses to contemporaneous and news shocks are sufficient to construct dynamic counterfactuals in a general class of linear models. I adapt these results to a different object of interest: the identification of cross-sectional elasticities clean of HEDGE.

The methodology combines cross-sectional and time-series analysis in three steps. First, a cross-sectional step estimates the response of the outcome Y_{it} to an innovation in R, the aggregate variable we intend to difference out, conditional on s_{ig}, the exposure shifters used for identification. Second, a time series step estimates a sequence of counterfactual innovations to R that replicate the impulse response of R to the shock of interest ϵ_{gt}. The final step uses the cross-sectional responses from the first step to evaluate the propagation of the innovations estimated in the second step. This provides an estimate of Ω_h for each horizon h. The estimated portable elasticity at horizon h, $\hat{\beta}_h^P$, is then given by $\hat{\beta}_h - \hat{\Omega}_h$.

Applying the decomposition requires one additional input relative to the TWFE strategy: exogenous variation in the aggregate variable (or variables) that confound identification. The specific nature of this variation — whether contemporaneous or involving news (i.e., anticipated) shocks — depends on the structure of the data-generating process. I distinguish three cases: static, dynamic without anticipation effects, and dynamic with anticipation effects, each imposing different informational requirements for identifying portable elasticities.

Before turning to the implementation of the decomposition, I show that the same source of exogenous variation can also be used to test the presence of HEDGE. The test is straightforward to implement within the TWFE specification: it requires adding an interaction between s_{ig} and the source of exogenous variation in the aggregate variable of interest. For example, one can test whether exposure to fiscal policy is correlated with exposure to monetary policy by adding an interaction between s_{ig} and the interest rate, instrumenting the latter with a well-identified series of monetary shocks. A statistically significant coefficient on this interaction term is evidence of HEDGE, provided the aggregate variable responds to the shock of interest. Importantly, the test relies only on s_{ig}, the source of cross-sectional variation being used for identification, and does not require knowledge of the true exposure shifters for other aggregate variables, such as s_{ir} in the above example. If the test finds evidence of HEDGE, the decomposition framework uses the same exogenous variation to estimate and remove the associated bias.

In static environments, fluctuations are driven by purely transitory shocks, and the framework can be implemented using any observed contemporaneous shock to the GE variable R. Although of little empirical relevance, I present this static setting to discuss how my framework relates to a common robustness check used in the cross-sectional literature. This robustness check, often referred to as *control function approach*, aims to address the HEDGE omitted variable bias by including an interaction term between the exposure shifter to the shock of interest, s_{ig}, and the endogenous aggregate variable R that the researcher suspects is not being differenced out. The underlying assumption is that this interaction term controls for the differential effects of R, without the need to use exogenous variation in R. In static settings, I show that my framework and the control function approach yield equivalent results. They both identify the true portable elasticity.

However, this equivalence result breaks down in dynamic settings, a characteristic feature of most empirical applications. Once dynamics are at play, the control function approach yields inconsistent estimates of portable elasticities because it cannot control for the *dynamic* path of the GE variable *conditional* on the shock of interest. It can only control for dynamic paths of the GE variable that lie in the subspace generated by variation in the included controls. In the simplest case with only a control for R_t, this corresponds to the unconditional autocovariance of R_t.[5] By contrast, the decomposition framework provides an estimation strategy that explicitly accounts for the dynamic behavior of macroeconomic variables. Thus, my approach is robust to general macroeconomic settings and represents an improvement over the control function approach.

The dynamic settings can be separated into two cases. First, I show how to implement the framework in dynamic settings without anticipation or forward-looking effects—referred to as Markov economies. In such cases, identification requires no further information beyond the static case: contemporaneous innovations in R suffice to identify the HEDGE term and the portable elasticity across horizons.

In settings with forward-looking dynamics, where both contemporaneous and expected realizations of variables affect current responses, identification of Ω_h additionally requires news shocks to the expected future path of the GE variable (Barnichon and Mesters 2023; McKay and Wolf 2023). Because such shocks are difficult to obtain, I outline how to apply the framework under limited information and then employ model simulations to assess estimation performance. Across both Markov and forward-looking economies, the framework improves upon the TWFE and control function approaches, even when only contemporaneous innovations to R_t are available.

[5] Intuitively, the impulse response of R depends on the shock that drives it. For example, monetary policy may respond to a demand shock with a hump-shaped path, while orthogonal innovations in R may follow a more AR(1)-like process. A single contemporaneous control cannot distinguish between these shock-contingent responses.

Empirical Application . I illustrate how to implement the framework by estimating US cross-sectional fiscal multipliers using state-level data, perhaps the most influential application of cross-sectional methods in macroeconomics. There is consensus that an advantage of cross-sectionally identified multipliers is their independence from GE effects, in particular those operating through the response of monetary policy. I study whether this is the case: Is monetary policy differenced out in the cross-section of US states? For estimation, I use the dataset on local defense spending built by Dupor and Guerrero (2017) and the series of Romer and Romer (2004) monetary shocks. Using specification (1), which is standard in the literature, I estimate a cumulative 2-year multiplier of 1.5. Applying the decomposition framework, I estimate a *portable* multiplier of 1. This estimate should be interpreted as a *monetary-policy–portable multiplier,* that is the cross-sectional effect of government spending holding the monetary response fixed.[6]

The difference with the TWFE estimate implies a positive HEDGE term. This arises because US states that are more exposed to defense spending shocks are also less sensitive to interest rates and, on average, interest rates increase in response to defense spending shocks. Together, these two imply an upward bias in the TWFE estimate.[7] This finding challenges the assumption that TWFE estimates of the cross-sectional fiscal multiplier are independent from the monetary regime and underscores the importance of GE effects in shaping cross-sectional estimates.

To interpret the empirical estimates, I build a two-region New Keynesian model with heterogeneity in exposure to fiscal and monetary policy. The framework combines the structure of Nakamura and Steinsson (2014), who study government spending in a two-region representative agent setting, with the two-region TANK features of Herreno and Pedemonte (2025), who introduce hand-to-mouth households but abstract from fiscal shocks. The model delivers theoretical analogs of the TWFE cross-sectional multiplier, the HEDGE term, and the portable multiplier and allows me to decompose the cross-sectional effect into distinct channels. In particular, I separate the usual demand forces—expenditure switching across regions and intertemporal substitution driven by relative price changes—from a HEDGE channel that captures intertemporal substitution driven by changes in the common national rate. This channel arises when regions differ in their interest rate sensitivity and acts as an additional treatment that pushes the TWFE cross-sectional multiplier away from its portable value.

[6]While the framework accommodates multiple confounding GE variables, I leave the analysis of additional sources of GE feedback to future work. Importantly, as discussed below, removing this monetary policy channel already restores a central use of cross-sectional multipliers: it allows researchers to combine the estimated effect with alternative monetary policy regimes and still make valid inferences on the aggregate effects of fiscal policy.

[7]These findings can be translated into the notation used to define Ω_h as follows. Increases in the interest rate have contractionary effects on local production, so $\gamma < 0$. The fact that high s_{ig} states are less sensitive to interest rates implies $\phi < 0$ and a monetary tightening following the fiscal shock implies $\boldsymbol{R}_{g,h} > 0$. Taken together, this yields $\Omega_h = \gamma \times \phi \times \boldsymbol{R}_{g,h} > 0$ and an upward bias in the TWFE estimates.

Previous work has argued that cross-sectional multipliers can be interpreted as bounds on the aggregate effects of fiscal policy, serving as an upper bound under conventional monetary policy, where contractionary interest rate responses are differenced out, and as a lower bound at the zero lower bound (Chodorow-Reich 2019; Dupor et al. 2023). This bounding logic relies on the cross-sectional design cleanly separating the direct effect from the aggregate *missing intercept* (Wolf 2023b) operating through monetary policy. With HEDGE, this separation fails: time fixed effects no longer fully remove the monetary response. As a result, the TWFE cross-sectional multiplier can vary widely across monetary regimes and need not bound the aggregate effect. In contrast, the portable multiplier remains stable across monetary regimes and retains its interpretation as a valid bound.

Finally, the model delivers a sufficient statistic that summarizes the precision loss from relying on limited news shocks. Calibrating this statistic to the US setting indicates that the decomposition recovers the portable multiplier with only limited error, even in a forward-looking New Keynesian setting.

Related Literature. This paper contributes to three strands of the macroeconomics literature. First, it is related to the empirical literature that uses cross-sectional methods to identify the effects of aggregate shocks. Nakamura and Steinsson (2018) provide a general review of the state and scope of this literature, while Chodorow-Reich (2020) focuses on the specifics of cross-regional analysis. Nakamura and Steinsson (2018) introduce the concept of portable statistics as well-identified moments or parameters that can be used to distinguish between models and imported across economic settings, precisely due to their partial equilibrium interpretation.

The typical empirical design in this literature relies on time fixed effects[8] to absorb GE effects. Papers that recognize the limitations of time fixed effects usually address the problem through a control-function approach, incorporating an interaction between the exposure shifter used for identification and the aggregate variable that varies in GE. First, I show that, in typical macroeconomic settings, the control function approach generally fails to control for HEDGE. Second, I propose a methodological framework that identifies PE elasticities in HEDGE economies and is robust to the dynamic patterns of macroeconomic data. The estimation framework in this paper extends the use of cross-sectional methods to macroeconomic settings in which cross-sectional variation plus time fixed effects are insufficient to control for dynamic GE effects.

This paper complements a recent literature that studies other limitations of cross-sectional designs to estimate the effects of aggregate shocks. Canova (2024) proposes an estimator robust to heterogeneity in the autoregressive dynamics of local outcomes. Adao, Arkolakis, and Esposito (2019) and Manuel (2025) show how to incorporate spatial links

[8] With a single cross-section, the intercept plays the role of time fixed effects in panel settings.

between regions, which allows indirect shock spillovers (through trade and migration) to be accounted for in local impact estimates. On a different front, Almuzara and Sancibrián (2024) focus on inference in these settings, advancing a procedure to obtain robust standard errors with correct coverage.

The study closest to this paper is Arkhangelsky and Korovkin (2024), who propose a reweighting procedure to identify contemporaneous responses to aggregate shocks in the presence of unobserved aggregate confounders. Their synthetic control approach constructs weights that orthogonalize exposure to the shock of interest from exposure to these confounders in the cross section.

This paper is complementary and takes a different route to the same identification challenge. I focus on environments in which aggregate variables are observed but respond dynamically to the shock. Rather than reweighting the cross section, I use time-series variation to construct counterfactual shocks that offset the GE response of observed aggregate variables.[9] This approach accommodates settings in which the same exposure drives sensitivity to multiple aggregates and where no single reweighting scheme can eliminate all sources of bias.

The decomposition framework that I propose is also closely related to the re-centered instrument in Borusyak and Hull (2023). In their setting, the shift-share instrument is re-centered by subtracting its expected value across possible shock assignments, so that it isolates variation arising from higher-than-expected exposure to exogenous shocks. My approach can be thought of as re-centering the aggregate shock of interest by removing the expected GE response that it induces in other aggregates.

The question of how to translate cross-sectional effects into aggregate effects has led to a growing literature on aggregation. Aggregation strategies include back-of-the-envelope calculations, model-based aggregation, and econometric frameworks that employ panel and factor structures (Sarto 2024; Matthes, Nagasaka, and Schwartzman 2024)[10]. Several papers combine reduced-form estimates at different levels of aggregation with structural assumptions, as in Guren et al. (2021), Chodorow-Reich, Nenov, and Simsek (2021), Guren et al. (2021), Wolf (2023b) and Wolf (2019). I complement this strand of work with a method to identify the free-of-GE micro elasticities that serve as starting points for both back-of-the-envelope and model-based aggregation.

The methodological framework that I propose is closely related to recent advances

[9] To fix ideas, consider a setting with a single aggregate confounder R. The standard identifying condition requires $cov(s_{ig}\epsilon_{gt}, s_{ir}R_t) = 0$, which fails in the presence of HEDGE. My approach can be instead interpreted as constructing a *counterfactual instrument* $\tilde{z}_{it} = s_{ig}(\epsilon_{gt} - \tilde{\epsilon}_{rt})$, where $\tilde{\epsilon}_{rt}$ denotes the counterfactual shock that offsets the GE response of R_t to ϵ_{gt}. This implies that $cov(\tilde{z}_{it}, s_{ir}R_t) = 0$, even when no reweighting of the cross section can achieve this orthogonality. This construction illustrates how time-series information can be used to undo the GE response by orthogonalizing shocks, rather than shares.

[10] These papers propose econometric frameworks that estimate aggregate effects directly using a Bayesian approach in Matthes, Nagasaka, and Schwartzman (2024) and a factor model in Sarto (2024).

in the computation of macroeconomic counterfactuals made by McKay and Wolf (2023), Barnichon and Mesters (2023) and Caravello, McKay, and Wolf (2024). McKay and Wolf (2023) show that reduced-form time series responses to contemporaneous and news policy shocks are sufficient to construct policy rule counterfactuals in a general class of linearized models. Barnichon and Mesters (2023) further show that these impulse responses are also enough to characterize the optimality of the policy rule. My decomposition framework builds on their work and extends it to the cross-sectional setting. I show how to combine cross-sectional and time series reduced-form responses to aggregate shocks to purge cross-sectional elasticities from HEDGE. The decomposition that I propose can be interpreted as computing cross-sectional elasticities under the counterfactual scenario of homogeneous exposure to GE.

The empirical application speaks to the literature on cross-sectional fiscal multipliers. Examples of this literature are Nakamura and Steinsson (2014), Chodorow-Reich et al. (2012), Auerbach, Gorodnichenko, and Murphy (2020), Demyanyk, Loutskina, and Murphy (2019), Dupor and Guerrero (2017), Dupor et al. (2023), Pennings (2022), and Pennings (2021), among others. Chodorow-Reich (2019) provides a comprehensive review of the lessons from this literature. One of the advantages attributed to cross-sectional methods for the estimation of fiscal multipliers is their independence from the state of the economy. In particular, their independence from the stance of monetary policy.[11]

At the same time, it is well documented that sensitivities to interest rates vary substantially between economic units (e.g., households, firms, or regions). For example, Herreno and Pedemonte (2025) find that US cities with lower per capita income are more responsive to monetary policy, a pattern that they show is consistent with a New Keynesian model of a monetary union with heterogeneous shares of hand-to-mouth households. Building on these insights, I revisit the estimation of cross-sectional fiscal multipliers by explicitly accounting for heterogeneity in monetary policy responses. My contribution is twofold. First, I show that heterogeneity in monetary policy responses across US states is correlated with their exposure to defense shocks. This implies that time fixed effects do not *fully* difference out monetary policy responses. Second, I introduce new estimates that control for these differential effects, thereby providing estimates of the cross-sectional fiscal multiplier that can be more confidently interpreted as independent of the monetary policy stance.

[11]The strength of the monetary policy response to fiscal shocks is an important determinant of the size of the aggregate fiscal multiplier. Theoretical and empirical papers on the interaction between fiscal and monetary policy include Christiano, Eichenbaum, and Rebelo (2011), Canova and Pappa (2011), DeLong and Summers (2012), Bachmann and Sims (2012), Riera-Crichton, Vegh, and Vuletin (2015), Auerbach and Gorodnichenko (2015), Farhi and Werning (2016), Leeper, Traum, and Walker (2017), Kato et al. (2018), Ramey and Zubairy (2018), Ramey (2019), Cloyne, Jordà, and Taylor (2020), Hack, Istrefi, and Meier (2023), among many others.

Layout. The remainder of the paper is structured as follows. Section 2 discusses the identification challenge at the core of the paper and presents a diagnostic test for it. Section 3 presents the estimation framework. Section 4 presents the empirical application. Section 5 presents a 2-region New Keynesian model to interpret the empirical findings from the previous section. Lastly, Section 6 concludes.

2. When does HEDGE challenge portability?

First, this section introduces a simple framework to illustrate the identification challenge that arises when units are differentially exposed to aggregate variables that comove with the shock of interest. Next, it presents a diagnostic test to detect whether HEDGE is present with respect to some aggregate variable R that the researcher intends to difference out.

To build intuition, consider a simple HEDGE economy in which units differ in their exposure to an aggregate shock of interest ϵ_{gt} and to one other aggregate variable R_t:

$$Y_{it} = \beta s_{ig}\epsilon_{gt} + \gamma s_{ir}R_t + u_i + u_t + u_{it}, \tag{4}$$

where Y_{it} is the outcome (e.g., local GDP) of unit i in period t and s_{ik} denotes the unit-specific exposure shifter to the aggregate variable k. Unit, time, and unit-time idiosyncratic shocks are denoted with a u plus the corresponding sub-indices. It can be shown that specification (1)[12], which uses unit and time fixed effects, recovers:

$$\hat{\beta}_h \to \beta_h^P + \gamma \underbrace{\frac{cov\left[s_{ig}, s_{ir}\right]}{v[s_{ig}]}}_{\phi} \underbrace{\frac{cov\left[R_{t+h}, \epsilon_{gt}\right]}{v[\epsilon_{gt}]}}_{\boldsymbol{R}_{g,h}}. \tag{5}$$

The above expression clarifies the conditions under which the dynamic effects that operate through R are not fully differenced out. There are two distinct objects: one related to time series variation, $\boldsymbol{R}_g = \{\boldsymbol{R}_{g,h}\}_h$, and one related to the source of cross-sectional variation used for identification, ϕ. The former measures the impulse response of R to the G shock of interest. In other words, $\boldsymbol{R}_g$ captures the dynamic path of R conditional on the realization of the shock of interest. For example, in a fiscal application, this could be the response of federal tax rates to federal government spending shocks.

The cross-sectional term, ϕ, measures the degree to which unit-specific exposure shifters to different aggregate variables comove with each other. It informs whether the driver of exposure to the shock of interest ϵ_{gt} is systematically related to the driver of exposure to R. For example, are regions that receive more government spending systematically more or less sensitive to taxation or to monetary policy? Expression (5) highlights that the

[12] In terms of this example, specification (1) is $Y_{it+h} = \beta_h s_{ig}\epsilon_{gt} + \lambda_{ih} + \lambda_{th} + e_{it+h}$.

assumption in a TWFE strategy like (1) is that the source of cross-sectional heterogeneity used for identification only drives variation in exposure to the shock of interest and not to any of the $\{R_t^k\}_k$ variables that the researcher intends to difference out:

$$cov[s_{ig}s_{ik}] = 0 \quad \forall\ k \ \text{ with } \ cov[R_{t+h}^k, \epsilon_{gt}] \neq 0 \ \text{ for some } h. \tag{6}$$

This assumption may be hard to satisfy in several macroeconomic applications where the type of unit-level characteristics that we use to proxy for heterogeneity in exposure to one aggregate shock are also likely to affect exposure to other aggregate variables. For example, in the fiscal literature, constrained or hand-to-mouth households are an important determinant of the sensitivity of output to fiscal stimulus. These households play a similar role when it comes to the sensitivity of output to monetary policy shocks.

Violations of the identifying assumption lead to omitted-variable bias, whose sign and magnitude can vary depending on the source of heterogeneity. As a result, studies using different sources of cross-sectional variation (or focusing on different time periods) may estimate different elasticities for the same aggregate shock, due to loading on different aggregate variables that move in GE. Such situations decrease the portability of cross-sectionally identified elasticities not just because some of the general-equilibrium responses are part of the estimates, but also because how these GE responses affect estimates depends on the cross-sectional and time-series variation that is being used.

2.1. Test for HEDGE

Next, I present a test to diagnose whether HEDGE poses an identification threat. In economic terms, two conditions must be satisfied for HEDGE with respect to a given aggregate variable R to pose an identification problem. First, R must respond to the shock of interest; that is, it must be one of the macroeconomic variables that move in general equilibrium. Second, at the cross-sectional level, units with different exposures to the shock of interest must exhibit heterogeneous responses to changes in R.

I assume that the researcher has identified a set of aggregate variables $\{R_t^k\}_{k\in K}$ that respond in GE to the shock of interest, in the sense that $cov(R_t^k, \epsilon_{gt}) \neq 0$, and that she intends to difference out. This set can be informed by time-series econometric tools, economic theory, or a combination of both. The test proposed here focuses on the second condition for HEDGE, namely whether the cross-sectional variation used for identification is orthogonal to differential responses to these equilibrium variables.

The goal is then to test whether s_{ig} drives heterogeneity in cross-sectional responses to a given aggregate variable R. Let ϵ_{rt} denote a source of exogenous variation in R (e.g., a monetary policy shock). The test exploits variation in ϵ_{rt} to identify differential responses to R, conditional on s_{ig}. It can be implemented via the following regression:

$$Y_{it+h} = c_h \times s_{ig} R_t + \lambda_{ih} + \lambda_{th} + e_{it+h} \qquad \text{for all} \quad h \in H, \tag{7}$$

where the interaction $s_{ig}R_t$ is instrumented with $s_{ig}\epsilon_{rt}$, and λ_{ih}, λ_{th} are horizon-specific unit and time fixed effects. In regression (7), c_h captures how the response of Y_{it+h} to changes in R_t varies with s_{ig}. For aggregate variables that respond to the shock of interest, a statistically significant coefficient indicates that units more exposed to G also respond differently to changes in R at horizon h, violating the identification assumption of the TWFE approach. In sum, evaluating HEDGE in relation to an aggregate variable R amounts to testing whether these coefficients are jointly zero:

$$H_0: \ c_h = 0 \ \ \forall \ \ h \in H, \qquad\qquad H_A: \ c_h \neq 0 \ \ \text{for some} \ \ h \in H. \tag{8}$$

Rejection of the null hypothesis implies that units with different levels of exposure to the shock of interest, as captured by s_{ig}, exhibit systematically different responses to changes in the aggregate variable R. In this case, estimates from cross-sectional regressions that rely on time fixed effects are likely to confound the direct effect of the shock with general equilibrium responses operating through R. Detecting such patterns in the data is a signal that HEDGE may be an identification threat and that a correction is necessary, such as the decomposition framework proposed in Section 3.

The test can be easily generalized to assess exposure heterogeneity with respect to multiple aggregate variables. Specifically, several interaction terms of the form $s_{ig}R^k_t$, each instrumented with $s_{ig}\epsilon^k_{rt}$, can be included for a set of variables $\{R^k\}_{k\in K}$. A joint test of significance on the corresponding coefficients $\{\hat{c}^k_h\}$ then provides a diagnostic for which general equilibrium channels pose a threat to identification in the cross section.

Importantly, the informational cost of this diagnostic is relatively low. The test requires only a source of exogenous variation in the aggregate variable R that the researcher aims to difference out. This variation is the same as that used to study the effects of R itself, for example, monetary or tax policy shocks, and can often be recovered using structural time series methods such as VARs or narrative identification approaches. In addition, the test does not require knowledge of the true driver of unit-level exposure to R. This is because, in a linear regression framework, what matters for the estimation is the linear relationship between the true exposure shifter and s_{ig}. Therefore, as long as s_{ig} is used for identification and is linearly related to the true exposure shifter, the test remains valid.

Recap. This section clarifies when HEDGE poses a threat to the identification of portable elasticities. Specifically, HEDGE threatens the validity of the TWFE strategy when (i) the GE variable responds to the shock of interest and (ii) the exposure shifter used for identification is correlated with heterogeneity in responses to that GE variable. I provide a simple test to detect such threats by leveraging exogenous variation in the GE variable. A rejection of the null signals that the identifying variation used in the cross-section is also picking up differential responses to general equilibrium forces. In such cases, cross-sectional estimates are likely to conflate the direct effect of the shock with these indirect channels. The next section presents a decomposition framework to estimate portable elasticities by separating them from the HEDGE component.

3. Decomposition Framework

This section presents a decomposition framework to identify portable elasticities in HEDGE economies. The goal is to decompose $\hat{\beta}_h$, the elasticity estimate for horizon h of the TWFE specification, into two distinct components: a portable component, β_h^P, and a HEDGE component, denoted by Ω_h.

The HEDGE component Ω_h measures how the response of a GE variable, namely R, affects the outcome of units with different levels of exposure to the shock of interest. In summary, the framework I propose uses reduced form cross-sectional responses to R to remove the effect of the expected response of R from the TWFE estimates. I outline how to implement this procedure under varying assumptions about the structure of the economic environment. In each case, identifying the HEDGE term requires access to exogenous variation in R, but the nature of this variation depends on the specific setting. I distinguish three stylized cases in terms of their informational requirements and describe how the methodology applies to each.

The first case corresponds to static economic environments. Here, both aggregate and cross-sectional variables are assumed to be driven by purely transitory, serially uncorrelated shocks. Implementing the framework requires access to contemporaneous innovations in the GE variable, R. Although this case has limited empirical relevance for macroeconomic questions, it provides a useful benchmark to compare the proposed framework with the control function approach, which consists of including the interaction $s_{ig}R_t$ as a control in the estimating equation. I show that under the static setting assumptions, the two approaches yield equivalent results. Both consistently estimate portable elasticities. However, the control function approach yields inconsistent estimates when dynamics are introduced, whereas the proposed framework remains valid.

The second case considers dynamic environments that admit a finite-order vector autoregression (VAR) representation (i.e. environments that satisfy a Markov structure). In this environment, implementing the framework requires contemporaneous innovations in R, just as in the static setting. This allows identification of the HEDGE term, Ω_h, both at impact and at different horizons ($h > 0$).

The third and most general case extends the framework to forward-looking environments, where both current and expected future realizations of variables affect present outcomes, such as in DSGE models. In these settings, identifying the HEDGE terms requires information not only on contemporaneous innovations in R, but also on news shocks that shift expectations about the future path of R. This extension is built on the work of McKay and Wolf (2023) and Barnichon and Mesters (2023), who show how time-series reduced-form responses to contemporaneous and news shocks can be used to construct counterfactuals at the aggregate level.

The remainder of this section presents each case in detail. Then, it briefly discusses the argument to establish consistency and how to conduct inference. Detailed explanations are relegated to Appendix A.6 and Appendix A.7, respectively.

3.1. Static Setting

Consider an economy that can be characterized by the following system of equations:

$$\text{(S1)} \qquad \begin{aligned} G_t &= \alpha R_t + u_{gt} \\ R_t &= \delta G_t + u_{rt} \\ u_{kt} &\sim \mathrm{N}(0, \sigma_k^2) \quad \forall\ k = g, r \end{aligned} \qquad\qquad \begin{aligned} Y_{it} &= \beta \times s_{ig} G_t + \gamma \times s_{ir} R_t + u_i + u_t + u_{it}^y \\ s_{ig} &\sim \mathrm{N}(\bar{s}, \sigma_{s_g}^2) \\ s_{ir} &= \phi s_{ig} + u_i^{s_r}, \quad s_{ig} \perp u_i^{s_r} \\ u_i^{s_r} &\sim N(0, \sigma_{s_r}^2), \quad u_{it}^y \sim N(0, \sigma_y^2), \end{aligned}$$

where i denotes economic units (e.g., regions, households, firms) and t time. G_t and R_t are two aggregate variables that are allowed to endogenously respond to each other. u_{gt} and u_{rt} have the interpretation of structural innovations to G and R, respectively. The unit-level outcome Y_{it} is a function of both G_t and R_t plus unit, time and unit-time idiosyncratic innovations. The cross-sectional variation in exposure to G and R is governed by s_{ig} and s_{ir}, respectively. I assume that s_{ir}, the exposure shifter to R_t, is linearly related to s_{ig}, the exposure shifter to G_t, with ϕ governing the strength of this relation. When $\phi = 0$, the exposure shifters to G and R are uncorrelated in the cross section. I further assume that the cross-sectional exposure shifters, s_{ig} and s_{ir}, are time-invariant and, therefore, independent of aggregate shocks.

The economy is static in the sense that cross-sectional and aggregate variables respond only contemporaneously to shocks, with no propagation over time. Given the absence of dynamics, the portable elasticity is non-zero only at impact

$$\beta_h^P = \begin{cases} \beta & \text{if } h = 0 \\ 0 & \text{if } h > 0 \end{cases} \tag{9}$$

and measures the relative effect of a change in G on the unit-level outcome, holding R fixed. The static nature of the problem implies that the cross-sectional elasticity for any horizon $h > 0$ is equal to zero, thus I drop the subscript h for the remainder of this subsection.

For illustration purposes, I consider a simplified setting with $\alpha = 0$, shutting down the endogenous response of G_t to R_t. The estimation steps and results that follow equally apply to a setting with two-way feedback as shown in Appendix A.1. I also normalize the variance of idiosyncratic shocks and the shares s_{ig} to unity.

Consider a researcher who observes an innovation ϵ_{gt} to G_t satisfying Assumption 1.

ASSUMPTION 1. *The observed innovation ϵ_{gt} is a noisy measure of the structural shock u_{gt}:*

$$\epsilon_{gt} = \kappa u_{gt} + \nu_{gt}, \quad \kappa \neq 0,$$

where $\nu_{gt} \sim N(0, \sigma^2_{\nu_g})$ is orthogonal to all structural innovations, in particular

$$cov[\nu_{gt}, u_{gt}] = 0, \quad cov[\nu_{gt}, u_{rt}] = 0.$$

This assumption ensures that ϵ_{gt} isolates variation in G_t orthogonal to the structural innovation in R_t. More generally, ϵ_{gt} may load on other aggregate innovations, in which case the decomposition that follows purges R-driven effects but may still reflect transmission through those channels.[13]

The typical estimation strategy consists of the following TWFE regression:

$$Y_{it} = b \times s_{ig} G_t + \lambda_i + \lambda_t + e_{it}. \tag{10}$$

Because G_t may be endogenous, $s_{ig}\epsilon_{gt}$ is used as an instrument for $s_{ig}G_t$. Here, λ_i and λ_t denote unit and time fixed effects, respectively. For ease of exposition, I set κ such that $cov[G_t, \epsilon_{gt}]/v[\epsilon_{gt}] = 1$. The coefficient then converges in probability to the following expression:[14]

$$\begin{aligned} E[\hat{b}] - \beta^P &= \gamma \times \frac{cov[s_{ir}R_t, s_{ig}\epsilon_{gt}]}{cov[G_t, \epsilon_{gt}]} \\ &= \gamma \times \phi \times \frac{cov[R_t, \epsilon_{gt}]}{cov[G_t, \epsilon_{gt}]} \end{aligned} \tag{11}$$

Here γ is the structural coefficient governing the response of Y_i to a change in R_t, conditional on exposure s_{ir}, and ϕ captures the cross-sectional correlation between s_{ig} and s_{ir}. The product $\tilde{\gamma} = \gamma \times \phi$ therefore captures the response of Y_i to a change in R_t, conditional

[13] See Appendix A.3.4 for details.

[14] Throughout, expectations of estimated coefficients refer to their probability limits.

on s_{ig} rather than s_{ir}:[15]

$$E[\hat{b}] - \beta^P = \tilde{\gamma} \times \frac{cov[R_t, \epsilon_{gt}]}{v[\epsilon_{gt}]} = \Omega, \tag{12}$$

where Ω denotes the bias associated with HEDGE. In order to decompose $\hat{b}$ into a portable and an HEDGE component, I propose an estimator for Ω, such that :

$$E[\hat{\beta}^P] = E[\hat{b} - \hat{\Omega}] = \beta^P. \tag{13}$$

This is the key idea of the framework. The additional piece of information required to estimate Ω is an instrument ϵ_{rt} for R_t that satisfies the following:

ASSUMPTION 2. *The observed innovation ϵ_{rt} is a noisy measure of the structural shock u_{rt}:*

$$\epsilon_{rt} = \upsilon u_{rt} + \eta_{rt}, \quad \upsilon \neq 0,$$

where $\eta_{rt} \sim N(0, \sigma^2_{\eta_r})$ is orthogonal to structural innovations, in particular:

$$cov[\eta_{rt}, u_{gt}] = 0, \quad cov[\eta_{rt}, u_{rt}] = 0.$$

Unlike ϵ_{gt} , the orthogonality restriction on η_{rt} cannot be relaxed. Any loading of η_{rt} on other structural innovations would induce correlation between ϵ_{rt} and these shocks, violating the exclusion restriction necessary for instrument validity.

Equipped with ϵ_{gt} and ϵ_{rt}, the estimation proceeds in three steps. For clarity, I describe the procedure in terms of population moments; the empirical implementation replaces these with their sample counterparts. Throughout, I denote the contemporaneous response of X to the observed innovation ϵ_{kt} by $\boldsymbol{X}_{k,t} = cov[X_t, \epsilon_{kt}]/v[\epsilon_{kt}]$.

1. *Cross-sectional Step.* Estimate the coefficients from a regression of Y_{it} on $s_{ig}\epsilon_{gt}$ and $s_{ig}\epsilon_{rt}$. These coefficients converge in probability to:

$$\beta + \tilde{\gamma} \times \boldsymbol{R}_{g,t}, \qquad \tilde{\gamma} \times \boldsymbol{R}_{r,t},$$

   respectively.
2. *Time Series Step.* The response of R_t to a unit change in ϵ_{rt}, given by $\boldsymbol{R}_{r,t}$, may differ from its response to a unit change in ϵ_{gt}, given by $\boldsymbol{R}_{g,t}$. This step estimates these two time-series impulse responses and then rescales ϵ_{rt} to match the response of R_t to ϵ_{gt}:

[15] Replacing s_{ir} for its expression in terms of ϕ, s_{ig} and $u_i^{s_r}$ in the unit-level outcome yields:

$$\begin{aligned} Y_{it} &= \beta \times s_{ig} G_t + \gamma \times \phi \times s_{ig} R_t + \gamma \times u_i^{s_r} R_t + u_i + u_t + u_{it}^y \\ &= \beta \times s_{ig} G_t + \tilde{\gamma} \times s_{ig} R_t + \gamma \times u_i^{s_r} R_t + u_i + u_t + u_{it}^y, \end{aligned}$$

where $\tilde{\gamma}$ is the parameter associated to the interaction term between s_{ig} and R_t.

$$\boldsymbol{R}_{r,t} \times \tilde{\epsilon}_{rt} = \boldsymbol{R}_{g,t} \Rightarrow \tilde{\epsilon}_{rt} = \frac{\boldsymbol{R}_{g,t}}{\boldsymbol{R}_{r,t}}.$$

3. *Final Step.* Combine the cross-sectional response to ϵ_{rt} with the rescaling from the time-series step to obtain:

$$\overbrace{\tilde{\gamma} \times \boldsymbol{R}_{r,t}}^{\text{CS Step}} \times \underbrace{\frac{\boldsymbol{R}_{g,t}}{\boldsymbol{R}_{r,t}}}_{\text{TS Step}} = \Omega.$$

This constructs the counterfactual response of Y_{it} to the component in R_t induced by ϵ_{gt}, isolating the HEDGE contribution to the baseline estimate.

Next, I present a detailed discussion on the implementation of each step.

Cross-sectional Step. The goal of this step is to identify the reduced form response of Y_i to both G and R, conditional on exposure to the shock of interest s_{ig}. As outlined above, this requires access to an instrument for R_t. For example, if R_t is the interest rate, then a series of well-identified monetary policy shocks would satisfy this requirement. The reduced form responses to changes in G_t and R_t can be estimated as follows:

$$Y_{it} = b \times s_{ig}\epsilon_{gt} + c \times s_{ig}\epsilon_{rt} + \lambda_i + \lambda_t + e_{it}, \tag{14}$$

where each aggregate shock is interacted with s_{ig}, the exposure shifter to G_t. It can be shown that this specification recovers the following two coefficients:

$$\begin{aligned} E[\hat{b}] &= \beta^P + \tilde{\gamma} \times \boldsymbol{R}_{g,t}, \qquad\qquad E[\hat{c}] = \tilde{\gamma} \times \boldsymbol{R}_{r,t}, \\ &= \beta^P + \underbrace{\tilde{\gamma} \times \delta}_{\Omega} \end{aligned} \tag{15}$$

where δ captures the response of R to an innovation in G. First, the coefficient on $s_{ig}\epsilon_{rt}$ isolates variation in R_t that is orthogonal to G_t, and therefore identifies $\tilde{\gamma}$. Second, because this regression only exploits the variation in R_t coming from ϵ_{rt}, the coefficient $\hat{b}$ is still a combination of the portable and HEDGE components as in the TWFE case. In this reduced-form specification, we isolate the variation in R_t driven by ϵ_{rt}, explicitly omit the endogenous variation in R_t induced by G_t.

Alternatively, one could augment the TWFE regression (10) by incorporating $s_{ig}R_t$ as an additional interaction:

$$Y_{it} = b^{CF} \times s_{ig}G_t + c^{CF} \times s_{ig}R_t + \lambda_i + \lambda_t + e_{it}^{CF}. \tag{16}$$

Under the assumptions of this setting, this regression identifies β^P in one step as OLS partials-out the correlation between G_t and R_t. This estimation strategy is often referred to as the *control function approach*. In such a case, there is no need for the additional time-series step that I discuss next. However, the control function approach works under restrictive assumptions on the dynamics of the data generating process, as I clarify in the next subsection. Because these assumptions are hard to justify in macroeconomic environments, I outline the time series step to (i) show that the decomposition and control function approach yield equivalent results in static settings and (ii) ease the transition to dynamic settings.

Time-series Step. The purpose of this step is two-fold. First, it estimates the response of R_t to each aggregate shock. Second, it uses these responses to compute the size of a hypothetical ϵ_{rt} shock that matches the response of R to a unit change in ϵ_{gt}. The time series responses to each shock can be estimated using:

$$R_t = \nu\epsilon_{gt} + a\epsilon_{rt} + e_{rt}, \tag{17}$$

with the estimated coefficients converging to:

$$E[\hat{\nu}] = \boldsymbol{R}_{g,t} = \delta, \qquad E[\hat{a}] = \boldsymbol{R}_{r,t}. \tag{18}$$

Here $\hat{\nu}$ denotes the estimated response of R to a unit change in ϵ_{gt} and $\hat{a}$ the estimated response to a unit change in ϵ_{rt}. Next, I use these estimated responses to calculate a hypothetical shock ϵ_{rt} that changes R_t exactly by $\hat{\nu}$. This hypothetical shock— denoted by $\tilde{\epsilon}_{rt}$ —is given by:

$$\hat{\tilde{\epsilon}}_{rt} = \frac{\hat{\nu}}{\hat{a}} \quad \Rightarrow \quad E[\hat{\tilde{\epsilon}}_{rt}] = \frac{\delta}{\boldsymbol{R}_{r,t}}. \tag{19}$$

Final Step. The final step uses $\hat{c}$, the estimated reduced form response of Y_i to R from the cross-sectional step, to evaluate the effect of $\tilde{\epsilon}_{rt}$, the hypothetical shock to R from the time series step. This yields a consistent estimate of Ω, the HEDGE term:

$$E[\hat{\Omega}] = E[\hat{c} \times \hat{\tilde{\epsilon}}_{rt}] = \tilde{\gamma} \times \delta = \Omega. \tag{20}$$

A consistent estimate for the portable elasticity can be constructed as follows.

$$\hat{\beta}^P = \hat{b} - \hat{c} \times \hat{\tilde{\epsilon}}_{rt} \quad \Rightarrow \quad E[\hat{\beta}^P] = \beta^P, \tag{21}$$

where $\hat{b}$ and $\hat{c}$ are estimated in the cross-sectional step and $\hat{\tilde{\epsilon}}_{rt}$ is estimated in the time series step. As mentioned above, under the assumptions made in this section, this three-step procedure yields an equivalent result to the control function approach.

Monte Carlo Simulations. The performance of the decomposition is illustrated using Monte Carlo simulations. I simulate data from an economy characterized by the system of equations presented above and estimate three different specifications: (i) Baseline, (ii) Decomposition, and (iii) Control function. Baseline corresponds to the TWFE regression. Decomposition corresponds to the results of applying the framework presented above and control function to the results from using specification (16). Figure 1 presents the distribution of the estimated coefficients for each strategy in an economy with $\beta = \gamma = \delta = \phi = .5$ and $\alpha = 0$. The population values for the portable and HEDGE terms are $\beta^P = .5$ and $\Omega = .125$, respectively. There are two takeaways. First, the baseline estimates are biased and centered on $\beta^P + \Omega = .625$. Second, both the decomposition and control function approach consistently estimate β^P. Figure A1 in Appendix A.1 shows that the same results hold for an economy with two-way general equilibrium feedback between G and R (that is, $\alpha \neq 0$).

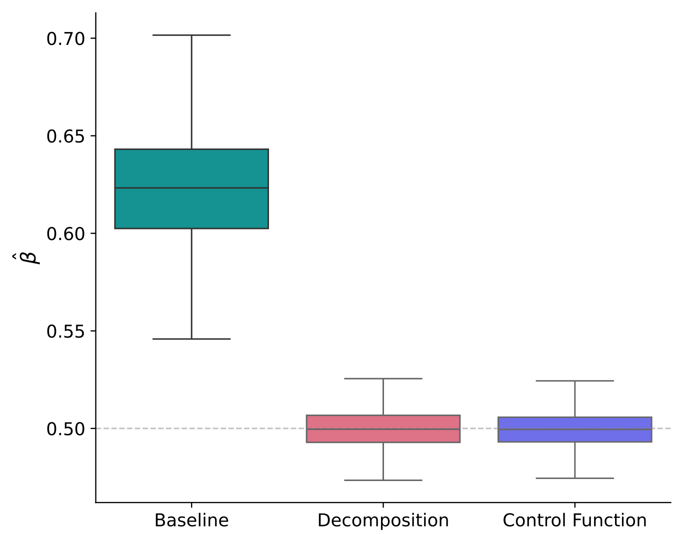


FIGURE 1. Static Monte Carlo Simulations - $\beta^P = .5$

Distribution of point estimates for the *Baseline*, *Decomposition* and *Control Function* estimation strategies. Based on 1000 repetitions with sample size of $n = 100$ and $t = 300$ and parameters set to $\delta = \beta = \gamma = \phi = .5$.

3.2. Dynamic Markov Settings - No anticipation effects

This subsection studies dynamic environments where aggregate variables evolve according to a finite-order VAR, while cross-sectional outcomes respond to current and lagged realizations of these aggregates and may exhibit their own persistence. I refer to these as Markov environments. These settings lie between static and forward-looking models, allowing for persistence while abstracting from anticipation effects.

In this environment, the control function approach generally fails to deliver consistent estimates, even though it is valid in static settings. The Markov structure makes it possible to isolate the source of this failure. It also provides a transparent benchmark that serves as a stepping stone to the forward-looking environments studied in Section 3.3.

In Markov environments, aggregate and cross-sectional variables evolve according to the following system of equations.

$$\begin{bmatrix} G_t \\ R_t \end{bmatrix} = \sum_{j=1}^{P} M_j \begin{bmatrix} G_{t-j} \\ R_{t-j} \end{bmatrix} + B \begin{bmatrix} u_{gt} \\ u_{rt} \end{bmatrix},$$

(S2)

$$Y_{it} = \sum_{s=0}^{S} \beta_s s_{ig} G_{t-s} + \sum_{k=0}^{K} \gamma_k s_{ir} R_{t-k} + \sum_{l=1}^{L} \psi_l Y_{i,t-l} + u_i + u_t + u_{it}^{y},$$

where G_t and R_t follow a $VAR(p)$ process, allowing for responses to each other's structural shocks u_{gt} and u_{rt}. The unit-level outcome Y_{it} follows an $AR(L)$ process and loads heterogeneously on contemporaneous and lagged realizations of both G_t and R_t, with the same exposure shifters applied uniformly across all lags. All other features of the environment are as in the static setting described in Section 3.1.

I assume that agents do not anticipate future realizations of aggregate shocks. Under this Markov structure, sequences of shocks that replicate a given path of aggregate variables ex post also reproduce the corresponding outcome responses.

For exposition purposes, I use as an example an economy with $P = 1$, $S = K = L = 0$ and where G does not respond to R. As in the static setting, I normalize the standard deviation of idiosyncratic shocks to unity. This yields familiar expressions that are helpful in building intuition. Formally, the economy is characterized by:

$$\begin{aligned} G_t &= \rho_g G_{t-1} + u_{gt}, \\ R_t &= \rho_r R_{t-1} + \delta G_t + u_{rt}, \\ Y_{it} &= \beta \times s_{ig} G_t + \gamma \times s_{ir} R_t + u_i + u_t + u_{it}^{y}. \end{aligned}$$

I refer to this as the $AR(1)$ case. This simplified example departs from the static setting only by introducing persistence in the aggregate variables. The parameters ρ_g and ρ_r govern the strength of persistence in G and R, respectively.

The goal is to estimate β_h^P, the dynamic differential response of Y_i to a shock to G_t, holding the time-path of R fixed. For the $AR(1)$ example, this implies:

(22) $$\beta_h^P = \beta \times \rho_g^h \qquad \forall \;\; h,$$

where β_h^P measures the cumulative captures the direct effect of a shock to G_t on Y_{it+h}, holding the entire path of R fixed. Consider a researcher who observes an innovation in G, denoted by ϵ_{gt}, that satisfies Assumption 3.

ASSUMPTION 3. *The observed innovation ϵ_{gt} is a noisy measure of the structural shock u_{gt}:*

$$\epsilon_{gt} = \kappa u_{gt} + \eta_{gt}, \quad \kappa \neq 0,$$

where $\eta_{gt} \sim N(0, \sigma_\nu^2)$ is orthogonal to contemporaneous, past, and future structural innovations:

$$cov[\eta_{gt}, u_{ks}] = 0 \quad \forall\ k \in \{g, r\},\ s.$$

Throughout, I denote the impulse response of X to an observed innovation ϵ_{kt} by $\boldsymbol{X}_k = \{\boldsymbol{X}_{k,h}\}_h$. Here $\boldsymbol{X}_{k,h}$ denotes the impulse response of X to ϵ_{kt} at horizon h. As before, I normalize κ such that $\boldsymbol{G}_{g,0} = 1$.

The TWFE estimation strategy can be implemented through the following panel local projection for each horizon h:

$$Y_{it+h} = b_h \times s_{ig} G_t + \lambda_{ih} + \lambda_{th} + e_{it+h}, \tag{23}$$

where $s_{ig}G_t$ is instrumented with $s_{ig}\epsilon_{gt}$ to address potential endogeneity concerns in G_t. Here, b_h measures the cumulative effect of a change in G_t on Y_{it+h} for the unit with average exposure. Horizon-specific unit and time fixed effects are captured by λ_{ih} and λ_{th}, respectively. The estimate for b_h converges to:

$$E[\hat{b}_h] = \underbrace{\beta_h^P}_{\text{Portable Effect}} + \underbrace{\Omega_h}_{\text{HEDGE term}}. \tag{24}$$

The first term is the portable component that measures the effect of interest, holding the path of R fixed. The second term Ω_h captures the HEDGE component that is now horizon-specific. In the AR(1) case, these terms can be expressed as:[16]

$$\begin{aligned} E[\hat{b}_h] &= \beta\ \boldsymbol{G}_{g,h} + \gamma\phi \boldsymbol{R}_{g,h} \\ &= \underbrace{\beta \times \rho_g^h}_{\beta_h^P} + \underbrace{\gamma\phi \times \delta \sum_{j=0}^{h} \rho_r^{h-j} \rho_g^j}_{\Omega_h}. \end{aligned} \tag{25}$$

If $\phi = 0$ or $\delta = 0$, the HEDGE terms vanish and the benchmark strategy recovers the

[16] The assumptions on ϵ_{gt} imply that the first-stage coefficient $\boldsymbol{G}_{g,0} = 1$, so I omit it for notational simplicity.

true portable effect. We can break down Ω_h into two terms:

$$\Omega_h = \tilde{\gamma} \times \boldsymbol{R}_{g,h}. \tag{26}$$

As in the static setting, the first term $\tilde{\gamma} = \gamma \times \phi$ measures the differential response of Y_i to a change in R, conditional on s_{ig}. The second term measures the dynamic response of R to a shock to G.

Next, I outline how to implement each step of the decomposition to estimate Ω_h in Markov settings. The additional input relative to the TWFE strategy is the same as in the static setting: an instrument for the contemporaneous realization of R. In what follows, I assume that, on top of the shock of interest ϵ_{gt}, we observe a contemporaneous innovation to R_t, denoted by ϵ_{rt}, that satisfies Assumption 4.

ASSUMPTION 4. *The observed innovation ϵ_{rt} is a noisy measure of the structural shock u_{rt}:*

$$\epsilon_{rt} = \upsilon u_{rt} + \eta_{rt}, \quad \upsilon \neq 0,$$

where $\eta_{rt} \sim N(0, \sigma^2_{\eta_r})$ is orthogonal to contemporaneous, past, and future structural innovations:

$$cov[\eta_{rt}, u_{ks}] = 0 \quad \forall\ k \in \{g, r\},\ s.$$

This assumption implies that ϵ_{rt} isolates exogenous variation in R_t that is orthogonal to contemporaneous, past, and future innovations to G, as well as, past and future innovations to R.[17]

Cross-sectional Step. The goal of this step is to identify the dynamic reduced-form responses of Y_i to G and R, conditional on s_{ig}. For identification of the R responses, we exploit the variation in R induced by the innovation ϵ_{rt}. The following local projection can be used to estimate these reduced-form responses:

$$Y_{it+h} = b_h \times s_{ig}\epsilon_{gt} + c_h \times s_{ig}\epsilon_{rt} + \lambda_{ih} + \lambda_{th} + e_{it+h}, \tag{27}$$

where each aggregate shock is interacted with s_{ig}, the exposure shifter to G. First, this specification identifies the same b_h as the baseline regression.[18] This is because the

[17]This assumption allows for correlation between ϵ_{gt} and ϵ_{rt}. The HEDGE term can be still be identified in this case as OLS partials out any shared variation.

[18]In the general case in which the first-stage coefficient on ϵ_{gt} is different from one, then the reduced form regression identifies a rescaled version of the expression in (25). The rescaled version is given by $\kappa_g \times \hat{b}_h^{TW}$ where κ_g is the coefficient of the first stage and $\hat{b}_h^{TW}$ is the *2SLS* coefficient of (25). Given that the decomposition results are robust to arbitrary first-stage values, I omit them from the main outline. The important point is that introducing the interaction between s_{ig} and ϵ_{rt} does not directly address the omitted variable bias due to HEDGE.

additional interaction only captures variation in R that is not correlated with G. For the AR(1) case, this implies

$$E[\hat{b}_h] = \beta\ \boldsymbol{G}_{g,h}\ +\ \tilde{\gamma}\ \boldsymbol{R}_{r,h}. \tag{28}$$

Second, this regression also identifies the response of Y_i to R, conditional on s_{ig} *and* conditional on the path of R after a change in ϵ_{rt}:

$$E[\hat{c}_h] = \tilde{\gamma}\ \boldsymbol{R}_{r,h}. \tag{29}$$

Inspection of $\hat{b}_h$ and $\hat{c}_h$ clarifies that the omitted variable bias due to HEDGE is a function of $\boldsymbol{R}_g$, while the estimated reduced-form response of Y_i to R is a function of a different source of variation, namely $\boldsymbol{R}_r$.

In this example, expressing the impulse responses in terms of parameters yields:

$$E[\hat{b}_h] = \underbrace{\beta \times \rho_g^h}_{\beta_h^P} + \underbrace{\tilde{\gamma} \times \delta \sum_{j=0}^{h} \rho_r^{h-j} \rho_g^j}_{\Omega_h}, \qquad\qquad E[\hat{c}_h] = \tilde{\gamma} \times \rho_r^h. \tag{30}$$

Here, the estimated coefficient on the interaction term $s_{ig}\epsilon_{rt}$ is a function of the differential sensitivity parameter $\tilde{\gamma}$ and the auto-regressive structure of R. To better understand the link between $\hat{c}_h$ and Ω_h, we can rewrite the latter as:

$$\Omega_h = \left[\tilde{\gamma}\rho_r^h + \tilde{\gamma} \sum_{j=1}^{h} \rho_r^{h-j} \rho_g^j\right] \times \delta \qquad \forall\ \ h \geq 0. \tag{31}$$

Consider the special case where G_t is an impulse (i.e. $\rho_g = 0$) so that the second term in the expression above drops:

$$\Omega_h = \tilde{\gamma}\rho_r^h \times \delta. \tag{32}$$

In this case, $\hat{c}_h$ identifies a rescaled version of Ω_h, where the scale depends on δ, the parameter that governs the contemporaneous response of R_t to G_t:

$$\Omega_h = \tilde{\gamma}\rho_r^h \times \delta \qquad \textit{vs.} \qquad E[\hat{c}_h] = \tilde{\gamma}\rho_r^h. \tag{33}$$

Intuitively, this scaling captures that in this example a change of one unit in ϵ_{rt} changes R_t by one unit, whereas this change is size δ for a change of one unit in ϵ_{gt}. Therefore, the only piece of information missing to compute Ω_h in this scenario is δ. Or, more generally, the ratio between the impact response of R to each aggregate shock, as in the static setting example.

Next, the time-series step provides a general strategy to take into account the difference between the impulse responses $\boldsymbol{R}_g$ and $\boldsymbol{R}_r$, which serves as a bridge between the HEDGE term and $\hat{c}_h$.

Time-series Step. This step provides the additional inputs required to move from $\hat{c}_h$ to $\hat{\Omega}_h$. From the previous step, we know how to evaluate the cross-sectional propagation of changes in R caused by ϵ_{rt} using $\hat{c}_h$. Following the methodology of Sims and Zha (2006), one can find a combination of innovations in R, denoted by $\tilde{\boldsymbol{\epsilon}}_r = \{\tilde{\epsilon}_{rt+h}\}_h$, which taken together reproduce $\boldsymbol{R}_g$ the impulse response of R to a G shock. Then we can use the reduced-form estimates from the cross-sectional step to evaluate the response of Y_i to the sequence of innovations $\tilde{\boldsymbol{\epsilon}}_r$.

By construction, this sequence of innovations exactly replicates $\boldsymbol{R}_g$ in the *ex post* sense. This is because at each point in time $h > 0$ we hit the economy with additional surprises to R, *unexpected* from the perspective of $h = 0$. Under the Markov assumption, innovations are not anticipated, so enforcing the path $\boldsymbol{R}_g$ ex post through a sequence of unexpected shocks yields the same outcome response as if that path had been anticipated ex ante. Section 3.3 relaxes the assumption of no anticipation effects following the work by McKay and Wolf (2023) and Barnichon and Mesters (2023).

Applying the approach of Sims and Zha (2006) requires estimates of the impulse response of R to the two aggregate shocks, ϵ_{gt} and ϵ_{rt}. These can be estimated using traditional time series methods, for example, local projections (Jordà 2005):

$$R_{t+h} = \nu_h \epsilon_{gt} + a_h \epsilon_{rt} + e_{rt+h}. \tag{34}$$

The regression coefficients $\hat{\nu}_h$ and $\hat{a}_h$ converge to the following expression:

$$E[\hat{\nu}_h] = \boldsymbol{R}_{g,h}, \qquad E[\hat{a}_h] = \boldsymbol{R}_{r,h}. \tag{35}$$

In the AR(1) example, this is equivalent to:

$$E[\hat{\nu}_h] = \delta\rho_r^h + \delta\sum_{j=1}^{h}\rho_r^{h-j}\rho_g^j, \qquad E[\hat{a}_h] = \rho_r^h. \tag{36}$$

In this example, unless $\rho_g = 0$, the path of R will differ between shocks even after rescaling by δ.

We can use $\hat{\boldsymbol{R}}_g = \{\hat{\nu}_h\}_h$ and $\hat{\boldsymbol{R}}_r = \{\hat{a}_h\}_h$ to find the sequence of innovations $\tilde{\boldsymbol{\epsilon}}_r$. Before presenting the general formula, I illustrate how the procedure works using the AR(1) example. Finding the innovation for period $h = 0$ requires:

$$\hat{a}_0 \hat{\tilde{\epsilon}}_{rt+0} = \hat{v}_0 \qquad \Rightarrow \qquad E[\hat{\tilde{\epsilon}}_{rt+0}] = E\left[\frac{\hat{v}_0}{\hat{a}_0}\right] = \delta. \tag{37}$$

where $\hat{\tilde{\epsilon}}_{rs}$ denotes the estimated counterfactual shock for period s. In other words, we look for the shock size to R that matches the estimated impact effect of a unit change in ϵ_{gt}. From the perspective of $h = 1$, this innovation implies that R changes by:

$$E[\hat{a}_1 \times \hat{\tilde{\epsilon}}_{rt+0}] = \delta\rho_r \qquad \neq \quad E[\hat{v}_1] = \delta\rho_r + \delta\rho_g. \tag{38}$$

Although the innovation for period $h = 0$ matches the impact response $\hat{v}_0$, it does not match the response at $h = 1$ that is given by $\hat{v}_1$. Therefore, we feed in an additional innovation that hits the economy in period $h = 1$ to account for the remaining difference between the two paths:

$$\hat{v}_1 = \hat{a}_1 \times \hat{\tilde{\epsilon}}_{rt+0} + \hat{a}_0 \times \hat{\tilde{\epsilon}}_{rt+1}, \tag{39}$$

where to evaluate the effect of this second innovation we use the contemporaneous coefficient $\hat{a}_0$. This implies that the estimated $\tilde{\epsilon}_{rt+1}$ satisfies:

$$E[\hat{\tilde{\epsilon}}_{rt+1}] = \delta\rho_g. \tag{40}$$

The sequence of ex-post innovations that replicates the path $\hat{\boldsymbol{R}}_g$ up to horizon H is characterized by the following result.

PROPOSITION 1. *Under Assumption 4, $a_0 \neq 0$. Hence, for any horizon H, there exists a unique sequence of innovations $\{\tilde{\epsilon}_{rt+h}\}_{h=0}^{H}$ such that*

$$v_h = \sum_{j=0}^{h} a_{h-j} \tilde{\epsilon}_{rt+j} \quad \forall\, h \in \{0, \dots, H\}.$$

Moreover, since $\hat{a}_0 \xrightarrow{p} a_0$, it follows that $\hat{a}_0 \neq 0$ with probability approaching one, so that the empirical analogue

$$\hat{v}_h = \sum_{j=0}^{h} \hat{a}_{h-j} \hat{\tilde{\epsilon}}_{rt+j}$$

is well-defined.

This result applies to any economy that satisfies (S3), including settings with two-way feedback between G and R.

Final Step. Once we estimate $\tilde{\epsilon}_r$, we can compute $\hat{\Omega}_h$ as follows:

$$\hat{\Omega}_h = \sum_{k=0}^{h} \hat{c}_{h-k}\hat{\tilde{\epsilon}}_{rt+k} \qquad \forall\ h\ \in\ H. \tag{41}$$

That is, we estimate $\hat{\Omega}_h$ using the reduced-form estimates from the cross-sectional step to evaluate the propagation of the sequence of innovations $\tilde{\epsilon}_r$. Crucially, the cross-sectional estimates $\hat{c}_h$ are evaluated using a sequence of shocks that generate the same variation in R that we used to identify those cross-sectional coefficients in the first place. Under the assumptions of this subsection, this yields a consistent estimate of the HEDGE term at each horizon h. Lastly, the portable elasticity estimate is computed as follows.

$$\hat{\beta}_h^P = \hat{b}_h - \hat{\Omega}_h. \tag{42}$$

where $\hat{b}_h$ is the estimated coefficient on $s_{ig}\epsilon_{gt}$ from the cross-sectional step.

Control Function Approach in Dynamic Settings. Inclusion of the interaction $s_{ig}R_t$ as a control, in general, does not deliver a consistent estimate of β_h^P. The reason is that the additional interaction term can only control for time paths of R that can be spanned by variation in the included controls, instead of variation in ϵ_{gt}. Because these two sources of variation may generate different time paths for R, the control function approach is not a valid robustness check for HEDGE in general macroeconomic environments. Intuitively, the dynamic response of R depends on the shock that drives it. For example, monetary policy may respond to a demand shock with a hump-shaped path due to inertia, while orthogonal fluctuations in R may follow a more AR(1)-like process. A contemporaneous control $s_{ig}R_t$ cannot distinguish between these state-contingent paths.

Concretely, the control function approach specification is given by:

$$Y_{it+h} = b_h^{CF} \times s_{ig}G_t + c_h^{CF} \times s_{ig}R_t + \lambda_{ih}^{CF} + \lambda_{th}^{CF} + e_{it+h}^{CF}, \tag{43}$$

where $s_{ig}G_t$ is instrumented with $s_{ig}\epsilon_{gt}$, as in the baseline TWFE strategy, and the superscript CF is used to denote the estimates from the control function approach. In the AR(1) example, it can be shown that:

$$E[\hat{b}_h^{CF}] - \beta_h^P \propto \tilde{\gamma} \times \left(\boldsymbol{R}_{g,h} - \delta\boldsymbol{R}_{\tilde{R},h}\right), \tag{44}$$

where $\delta\ \boldsymbol{R}_{\tilde{R},h}$ is proportional to the path of R spanned by the included controls (in this case, the contemporaneous realization R_t).[19] Specifically, δ, captures the contemporaneous response of R to the G shock.

Expression (44) highlights that, unless $\boldsymbol{R}_{g,h} = \delta\boldsymbol{R}_{\tilde{R},h}$, the control function approach does not provide a consistent estimate of β^P_h. For example, if the path of R resembles an AR(2) conditional on ϵ_{gt} but an AR(1) unconditionally then $\boldsymbol{R}_{g,h} \neq \delta\boldsymbol{R}_{\tilde{R},h}$. In Appendix A.3.1, I derive the expected value of $\hat{b}^{CF}_h$ for alternative DGPs to further illustrate the mechanics of the control function approach. Moreover, while the control function approach is inconsistent in general, its performance relative to its (biased) population value can vary substantially with model mis-specification. In particular, when the set of additional controls, such as lagged dependent variables or lagged shocks, is misspecified, the resulting estimates can deviate considerably from that population value.

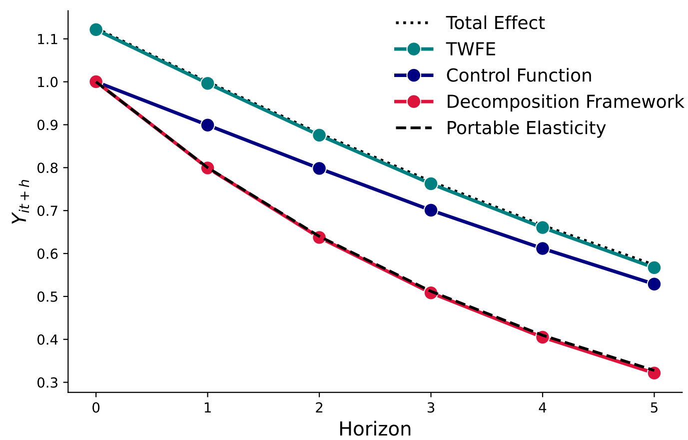


FIGURE 2. Dynamic Setting Without Anticipation Effects - Monte Carlo Results

Solid lines show the estimated cross-sectional responses to a G shock for three specifications - *TWFE, Decomposition Framework* and *Control Function* - together with the population portable and total ($\beta^P_h + \Omega_h$) elasticities. Shaded areas show 90% confidence bands. Based on 1000 repetitions with N = 300 and T = 1000. The DGP is generated using the same parameters as for the static case plus $\rho_g = .8$, $\rho_r = .8$, $\rho_y = .0$.

[19]Plugging in the expressions for $\boldsymbol{R}_{g,h}$ and $\boldsymbol{R}_{\tilde{R},h}$ yields:

$$\boldsymbol{R}_{g,h} - \delta\boldsymbol{R}_{\tilde{R},h} = \delta\rho_r^h + \delta\sum_{j=1}^{h}\rho_r^{h-j}\rho_g^j - \delta\underbrace{\rho_r^h}_{\boldsymbol{R}_{\tilde{R},h}} = \delta\sum_{j=1}^{h}\rho_r^{h-j}\rho_g^j.$$

In this AR(1) example, the control function approach partially addresses the problem by taking care of the correlation between ϵ_{gt} and R_{t+h} resulting from the pure auto-regressive structure of R. However, it cannot account for the dynamic effects of ϵ_{gt} on R_{t+h} that go through $\{G_{t+k}\}_{k=1}^h$. Note that, for example, if R responded to G with a lag then the control function approach would, in expectation, yield no improvements over the TWFE strategy for $h > 0$.

Monte Carlo Simulations. The performance of the decomposition is illustrated through a series of Monte Carlo simulations. I simulate data for different economies that satisfy the Markov assumptions of system (S3) and estimate three different impulse response functions: i) TWFE, ii) Control Function, iii) Decomposition Framework. These correspond respectively to specifications (23) and (43), and to the portable elasticity obtained from the decomposition strategy discussed above.

I present two sets of results. First, in Figure 2 I show results for the AR(1) example. Full details on the parameters used for the simulation can be found in the Appendix A.3.2. The figure shows the estimated cross-sectional responses to ϵ_{gt}. The dashed black line corresponds to the population portable elasticity, while the dotted black line corresponds to the population total effect (i.e., $\beta^P_h + \Omega_h$). The decomposition framework provides a consistent estimate of the portable elasticity at all horizons. In contrast, neither the TWFE strategy nor the control function approach identify the portable elasticity.

Second, I study the performance of the decomposition and control function approaches in a variety of data-generating processes that fit into (S3). The sample size is set to $N = 300$ and $T = 1000$, and the number of repetitions to $J = 1000$. Appendix A.3.2 details the set of data-generating processes and their parameter values. For each estimation strategy k, DGP z and horizon h, I compute the absolute difference between the population portable elasticity, β^P_{zh}, and the average estimate across all repetitions and then normalize it by β^P_{zh} to make it comparable across DGPs.

$$\theta^k_{zh} = \frac{\mid J^{-1} \sum_{j=1}^{J} \hat{b}^k_{jzh} - \beta^P_{zh} \mid}{\mid \beta^P_{zh} \mid}. \tag{45}$$

Consistency requires $\theta^k_{zh} \to 0$. Table 1 shows the mean and standard deviation of θ^k_{zh} by horizon and estimation method. The decomposition approach consistently recovers the true portable effects as opposed to the control function approach and TWFE baseline strategy.

TABLE 1. Absolute Mean Relative Bias by Estimation Method and Horizon

Estimation Strategy	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 5$
Decomposition	0.002	0.002	0.007	0.013	0.021	0.031
	(0.003)	(0.002)	(0.007)	(0.014)	(0.023)	(0.035)
Control Function	0.445	0.259	0.245	0.282	0.379	0.516
	(0.494)	(0.223)	(0.182)	(0.300)	(0.417)	(0.519)
TWFE	0.146	0.226	0.333	0.463	0.603	0.743
	(0.130)	(0.192)	(0.295)	(0.385)	(0.480)	(0.580)

Standard errors shown in parentheses. Values are reported as fractions of the population portable elasticity (1 = 100 %). The sample size is set to $N = 300$ and $T = 1000$ and the number of repetitions to $J = 1000$. See Appendix A.3.2 for further details on each parametrization.

Extensions. Appendix A.3.3 presents the derivations and results for settings with two-way general equilibrium feedback (2W-GE) between aggregate variables. This analysis yields an important insight. I show that, in general, when there is 2W-GE, the cross-sectional elasticities estimated using either two-stage least squares (2SLS) or reduced-form analysis will be contaminated by GE effects, even in the absence of HEDGE. In other words, in such settings, the HEDGE conditions discussed at the beginning of the paper are sufficient to break identification but not necessary. The TWFE strategy will, in general, not identify a clean partial equilibrium elasticity in these settings. This happens because units that are more exposed to G will be more exposed to both the *autonomous* change in G triggered by ϵ_{gt} and to the posterior change in response to R. This is an overlooked source of omitted variable bias that this paper brings to light and which the decomposition framework *also* addresses.

Appendix A.3.4 shows that the decomposition framework can accommodate settings where both R_t and ϵ_{gt} are correlated with a third aggregate variable A_t (for example, TFP, exchange rate), but A_t affects all units in the same way. This echos empirical applications for which the available source of variation ϵ_{gt} is not as good as random at the aggregate level, but it does satisfy exogeneity with respect to A_t in the cross section. Note that this is equivalent to saying that there is no HEDGE with respect to A_t.

Appendix A.4 discusses three alternative estimation strategies that one might expect to address HEDGE, and demonstrates through a Monte Carlo simulation exercise that they fail to do so. The strategies considered are (i) interactive fixed effects (Bai 2009); (ii) alternative version of the CFA including $s_{ir}R_t$ rather than $s_{ig}R_t$; (iii) alternative CFA instrumenting $s_{ig}R_t$ with $s_{ig}\epsilon_{rt}$.

3.3. Forward-Looking Dynamic Setting

This section generalizes the decomposition framework to settings with forward-looking dynamics. This is akin to the dynamics implied by DSGE models, which I use as a laboratory to test the performance of the framework. The difference with respect to the preceding section is that now there can be anticipation effects: agents respond both to current and expected realizations of macroeconomic variables. The key implication is that estimating the differential response to the conditional path the GE variable, R, requires feeding the system a counterfactual sequence of R shocks that reproduces $\boldsymbol{R}_g$ both ex post and in expectation.

As shown by McKay and Wolf (2023) and Barnichon and Mesters (2023), constructing counterfactuals that are robust to forward-looking dynamics requires access to shocks that shift both the contemporaneous and expected future realization of macroeconomic variables. First, I show how to implement the decomposition when a researcher has access to these types of shocks and then I discuss implementation in cases where access to shocks is limited. In settings with limited access to news shocks, the decomposition framework provides an approximation to the population portable elasticity. The conditions that determine the precision loss resulting from limited information depend on the data-generating process and the shocks under analysis. As an illustration, in Section 5.3, I study the determinants of the decomposition accuracy in the context of a two-region New Keynesian model used to estimate cross-sectional fiscal multipliers.

Consider an economy in which the unit-level outcome can be characterized as follows.

$$\begin{aligned} Y_{it} = &\sum_{f=0}^{F} \beta_f s_{ig} E_t[G_{t+f}] + \sum_{f=0}^{F} \gamma_f s_{ir} E_t[R_{t+f}] \\ &+ \sum_{s=1}^{S} \beta^{\ell}_{s} s_{ig} G_{t-s} + \sum_{k=1}^{K} \gamma^{\ell}_{k} s_{ir} R_{t-k} + \sum_{l=1}^{L} \psi_l Y_{it-l} + u_i + u_t + u^{y}_{it}, \end{aligned} \tag{46}$$

where $E_t[X_{t+f}]$ is the expectation at time t for the realization of X in period $t+f$. The parameter β_f, for $f > 0$, captures the present response to the expected realization of G in period $t+f$. Similarly for γ_f. I assume that only the first F expected realizations of aggregate variables matter for the present responses and use the superscript ℓ to denote the responses to lagged realizations of the aggregate variables.[20] In addition, I assume that exposure shifters are variable-specific but not horizon specific (i.e., s_{ig} determines the cross-sectional response to both contemporaneous, expected, and lagged realizations of G).[21]. The aggregate processes are as in the Markov economy with the addition that they are subject to both contemporaneous and news (or anticipated) structural innovations:

$$\begin{bmatrix} G_t \\ R_t \end{bmatrix} = \sum_{j=1}^{P} M_j \begin{bmatrix} G_{t-j} \\ R_{t-j} \end{bmatrix} + B \begin{bmatrix} \boldsymbol{u}_{gt} \\ \boldsymbol{u}_{rt} \end{bmatrix}, \tag{47}$$

where $\boldsymbol{u}_{kt}$ is the vector of contemporaneous and news shocks to k in period t.[22]

ASSUMPTION 5. *For each aggregate variable $k \in \{g, r\}$, the innovation vector $\boldsymbol{u}_{kt}$ admits the*

[20] Alternatively, the DGP in (46) can be interpreted as an approximation of a true DGP with $F^* = \infty$ and $\gamma_f \approx 0$ for $f > F$.

[21] The estimation framework does not require knowledge of the true exposure shifters to R so assuming a constant s_{ir} is without loss of generality. Horizon-specific shifters to R, $s_{ir,j}$ for $j \in \{F, K\}$, imply that ϕ becomes horizon-specific which will be reflected in the estimated cross-sectional responses to different news shocks (See cross-sectional step below). The assumption that s_{ig} is a good proxy for exposure to present and expected realizations of G is in line with most empirical work. If the exposure shifters to the contemporaneous realization of G, s_{ig}, and to the expected realizations of G, $s_{ig,f}$, differ, then $s_{ig} \times \epsilon_{gt}$ will be a strong instrument for $s_{ig} \times G_t$ but possibly weaker for $s_{ig} \times E_t[G_{t+f}]$. This is independent of HEDGE and, therefore, relaxing the assumption of constant s_{ig} exceeds the scope of this paper.

[22] For example, the first entry of $\boldsymbol{u}_{kt}$ corresponds to a contemporaneous shock, while the second one corresponds to news about the realization of k in period $t+1$.

decomposition

$$\boldsymbol{u}_{kt} = \{u^t_{kt+f}\}^F_{f=0},$$

where u^t_{kt} is a contemporaneous shock and, for $f \geq 1$, u^t_{kt+f} is a news shock revealed at time t about the realization of variable k at time $t+f$.

All innovations are mean zero and mutually orthogonal:

$$cov[u^t_{kt+f}, u^t_{k't+f'}] = 0 \quad \textit{for all } (k, f, t) \neq (k', f', t').$$

This orthogonality assumption implies that each news shock shifts expectations about a specific future realization of the aggregate variable without affecting other horizons or contemporaneous innovations.

The rest of the economy is the same as in the Markov setting of Subsection 3.2.

The object of interest is the dynamic effect of a shock to G_t on the cross-sectional outcome, holding the realized and *expected* path of R fixed. Consider the following panel local projection with two-way fixed effects:

$$Y_{it+h} = b_h \times s_{ig} G_t + \lambda_{ih} + \lambda_{th} + e_{it+h}, \tag{48}$$

where $s_{ig}\epsilon_{gt}$ is used as an instrument for $s_{ig}G_t$ to address potential endogeneity concerns. To simplify notation, I assume that $\beta_j = \beta \ \forall \ f, \ell$ and shut down the parameters on lagged realizations of aggregate and local variables in the process for Y_{it}. Then, $\hat{b}_h$ converges to the following expression:

$$E[\hat{b}_h] = \beta^P_h + \underbrace{\sum_{f=0}^{F} \gamma_f \phi \frac{cov[E_t[R_{t+h+f}], \epsilon_{gt}]}{v[\epsilon_{gt}]}}_{\Omega_h}, \tag{49}$$

where the first term is the portable component that captures the effect of a shock to G_t on Y_{it+h}, including the effects through changes in future G, holding $E_t[R_{t+h}]$ for $h \geq 0$ constant. The second term is the HEDGE component which is a function of (i) the contemporaneous and expected changes in R_t conditional on a shock to G_t and (ii) the cross-sectional sensitivities to contemporaneous and expected future realizations of R, conditional on s_{ig}. Next, I outline how the cross-sectional and time-series steps can be generalized to accommodate forward-looking dynamics.

Cross-sectional Step . This section outlines the estimation strategy used to identify the dynamic reduced-form responses of Y_i to G_t and to both current and expected changes in R. The latter set of parameters measures the response of Y_{it+h} to the expectation formed at time t about the realization of R_{t+f}. For example, for $h = f = 0$, this parameter measures the contemporaneous response of Y_{it} to a contemporaneous innovation in R. Identifying these parameters requires exogenous variation in the period-t expectation about the realization of R_{t+f}. In other words, we need to observe news (or anticipated) shocks about the realization of R_{t+f} that agents learn about in period t. Let $\epsilon^t_{r,t+f}$ denote an observed innovation that captures news received at time t about the realization of R_{t+f}. Analogously to the Markov setting, I assume that these observed innovations are noisy measures of the corresponding structural news shocks:

$$\epsilon^t_{r,t+f} = \upsilon_f u^t_{rt+f} + \eta^t_{r,t+f}, \quad \upsilon_f \neq 0,$$

where the noise terms $\eta^t_{r,t+f}$ are orthogonal to all structural innovations. For example, an unexpected monetary policy announcement this quarter about a possible one-period interest rate hike next quarter could be used to identify the coefficients for $f = 1$, at different horizons.

Suppose that we have access to news shocks up to horizon F, then we can estimate the following reduced-form local projections:

$$Y_{it+h} = b_h \times s_{ig}\epsilon_{gt} + \sum_{f=0}^{F} c_{fh} \times s_{ig}\epsilon^t_{r,t+f} + \lambda_{ih} + \lambda_{th} + e_{it+h}.$$

Here $\hat{c}_{f,h}$ is the reduced form response of Y_{it+h} to the news received at t about the realization of R_{t+f}, conditional on s_{ig}. As before, $\hat{b}_h$ is the reduced-form response of Y_{it+h} to a change in G_t, conditional on s_{ig}.

Time-Series Step . The time-series step estimates the impulse response of R to (i) the shock of interest, and (ii) the same set of contemporaneous and news shocks to R used in the cross-sectional step.

$$R_{t+h} = \nu_h \epsilon_{gt} + \sum_{f=0}^{F} a_{fh}\epsilon^t_{r,t+f} + e_{rt+h}, \tag{50}$$

where $\epsilon^t_{r,t+f}$ is the news shock to R defined above. The coefficients a_{fh} measure the response of R_{t+h} to the information received in period t about the realization of R_{t+f}. Under rational expectations, $\hat{\nu}_h$ and $\hat{a}_{fh}$ are unbiased estimates of $cov[E_t[R_{t+h}], \epsilon_{gt}]$ and $cov[E_t[R_{t+h}], \epsilon^t_{r,t+f}]$, respectively.

Next, I use the impulse responses to the $F+1$ news shocks to find the combination of date-t shocks that replicates the estimated impulse response of R to ϵ_{gt}, $\hat{\boldsymbol{R}}_g = \{\hat{\nu}_h\}_h$. Let $\hat{A}$ be an $(H+1) \times (F+1)$ matrix whose (h, f)-th entry is $\hat{a}_{hf}$, the estimated response of R_{t+h} to a time-t news shock about R_{t+f}.

The estimated sequence of counterfactual date-t news shocks $\hat{\tilde{\boldsymbol{\epsilon}}}^t = \{\hat{\tilde{\epsilon}}^t_{r,t+f}\}^F_{f=0}$ is defined as the solution to:

$$\hat{\boldsymbol{R}}_g = \hat{A}\hat{\tilde{\boldsymbol{\epsilon}}}^t. \tag{51}$$

PROPOSITION 2.

i. *If $\hat{\boldsymbol{R}}_g$ lies in the column space of $\hat{A}$, then there exists a sequence $\hat{\tilde{\boldsymbol{\epsilon}}}^t$ that exactly replicates $\hat{\boldsymbol{R}}_g$. Moreover, if $\hat{A}$ has full column rank, this solution is unique.*

ii. *If $\hat{\boldsymbol{R}}_g$ does not lie in the column space of $\hat{A}$, the solution is given by the least squares projection:*

$$\hat{\tilde{\boldsymbol{\epsilon}}}^t = \arg\min_{e \in R^{F+1}} \left\| \hat{\boldsymbol{R}}_g - \hat{A}e \right\|^2.$$

iii. *A sufficient condition for the uniqueness of the solution in (ii) is that $\hat{A}$ has full column rank, in which case:*

$$\hat{\tilde{\boldsymbol{\epsilon}}}^t = (\hat{A}'\hat{A})^{-1}\hat{A}'\hat{\boldsymbol{R}}_g.$$

The rank condition on $\hat{A}$ is an empirical property of the estimated impulse responses. In particular, while Assumption 5 provides a structural interpretation of the shocks, uniqueness of the counterfactual decomposition depends on whether the observed news shocks span sufficiently rich variation in the data. The system is overidentified when $H > F$, exactly identified when $H = F$, and underidentified when $H < F$. In typical applications, $H > F + 1$, so an exact replication is not guaranteed. In this case, the solution corresponds to the projection of $\hat{\boldsymbol{R}}_g$ onto the space spanned by the news shocks observed at time t.

Final Step. Lastly, I compute $\hat{\Omega}_h$ by using the estimates from the cross-sectional step to evaluate the propagation of the estimated sequence of news shocks $\tilde{\boldsymbol{e}}^t$:

$$\hat{\Omega}^t_h = \sum_{f=0}^{F} \hat{c}_{fh} \times \hat{\tilde{\epsilon}}^t_{r,t+f}, \tag{52}$$

where the superscript t is used to denote the estimated HEDGE term of the ex-ante decomposition. The estimate for β^P_h follows from the following:

$$\hat{\beta}^{P,t}_h = \hat{b}_h - \hat{\Omega}^t_h. \tag{53}$$

Implementation with limited access to news shocks. In the cases where we observe $\tilde{F} < F$ news shocks, the decomposition yields an approximation to the population portable elasticities. In general, the loss of precision increases in the strength of forward-looking dynamics beyond $\tilde{F}$ (i.e., the importance of the parameters γ_f for $f > \tilde{F}$) and in the difference between the observed path of R after the shock of interest and the path enforced by the estimated sequence of shocks $\tilde{\boldsymbol{e}}^t$. In Section 5.3, I use a two-region New Keynesian model to explore which observable moments can be used to study the performance of the framework in settings with limited access to news shocks.

DSGE Monte Carlo Simulations. To assess the performance of the decomposition framework, I conduct DSGE Monte Carlo (Ramey 2016) simulations using data generated from a two-region New Keynesian model with government spending shocks. The model extends Nakamura and Steinsson (2014) by allowing heterogeneity across regions in both their exposure to government spending, G_t, and their sensitivity to the interest rate, R_t. Without loss of generality, I assume that $cov[s_{ig}, s_{ir}] < 0$, so that the region more exposed to G_t is less sensitive to R_t. Government spending follows an AR(1) process with innovation ϵ_{gt}, while the nominal interest rate, i_t, is determined by a Tayor rule and subject to its own shock ϵ_{it}. As a result, the real interest rate R_t responds positively and persistently to fiscal shocks (that is, $cov[R_{t+h}, \epsilon_{gt}] > 0$). This calibration implies a positive HEDGE term. Appendix D details the full model, and Table A9 summarizes the calibration.

The simulated data are used to estimate cross-sectional fiscal multipliers using the estimation strategies discussed above. Given the forward-looking nature of standard New-Keynesian dynamics, I assume access to a finite number of news shocks (i.e., $\tilde{F} < F = \infty$) when applying the ex-ante decomposition.

I compare two parametrizations that differ in how R_t responds to a monetary shock ϵ_{it}, while holding fixed the response of R_t to ϵ_{gt}. In both scenarios, fiscal shocks induce a persistent increase in R_t. In the first scenario—labeled *misaligned*—monetary shocks induce only a transitory response in R_t, while in the second—labeled *aligned*—the response is calibrated to more closely mirror the one following fiscal shocks.[23] In other words, these scenarios play with the difference between $\hat{\boldsymbol{R}}_g$ and $\hat{\boldsymbol{R}}_r^0$, the estimated response of R to a contemporaneous innovation to the monetary rule.

Figure 3 presents the simulation results. The left panel corresponds to the *misaligned* scenario, and the right panel corresponds to the *aligned* scenario. In each case, I show the estimated multipliers for four estimation strategies: (i) the TWFE approach, (ii) the control function approach, (iii) the ex-ante decomposition using only a contemporaneous innovation to R, (iv) the ex-ante decomposition using four news shocks[24] to R.

The *misaligned* scenario combines the strong forward-looking dynamics of the model (Del Negro, Giannoni, and Patterson 2023; McKay, Nakamura, and Steinsson 2016) with significantly different paths for R after each aggregate shock. In this setting, the decomposition framework improves on the TWFE estimator, but suffers from some loss of

[23]In Appendix A.5, Figure A3 plots the impulse responses of R for each scenario.

[24]That is, on top of the contemporaneous innovation, I assume that the researcher observes news shocks about the realization of R in period $t + s$ for $s = \{1, 2, 3\}$.

precision. This is because the estimation is constrained by the limited number of shocks (i.e. $\tilde{F} = \{1, 4\}$). Nevertheless, the decomposition estimates, using one or four shocks, both outperform the control function approach, which in this calibration yields estimates that closely resemble those of the TWFE benchmark.

In the *aligned* scenario, the model still features forward-looking dynamics, but the interest rate responses to fiscal and monetary shocks are more closely aligned. Here, the decomposition performs substantially better. In particular, the ex-ante decomposition with one or four news shocks recovers a portable elasticity that is very close to the population value at all horizons. While the control function estimator also improves under this calibration, it continues to under-perform relative to the decomposition approach.

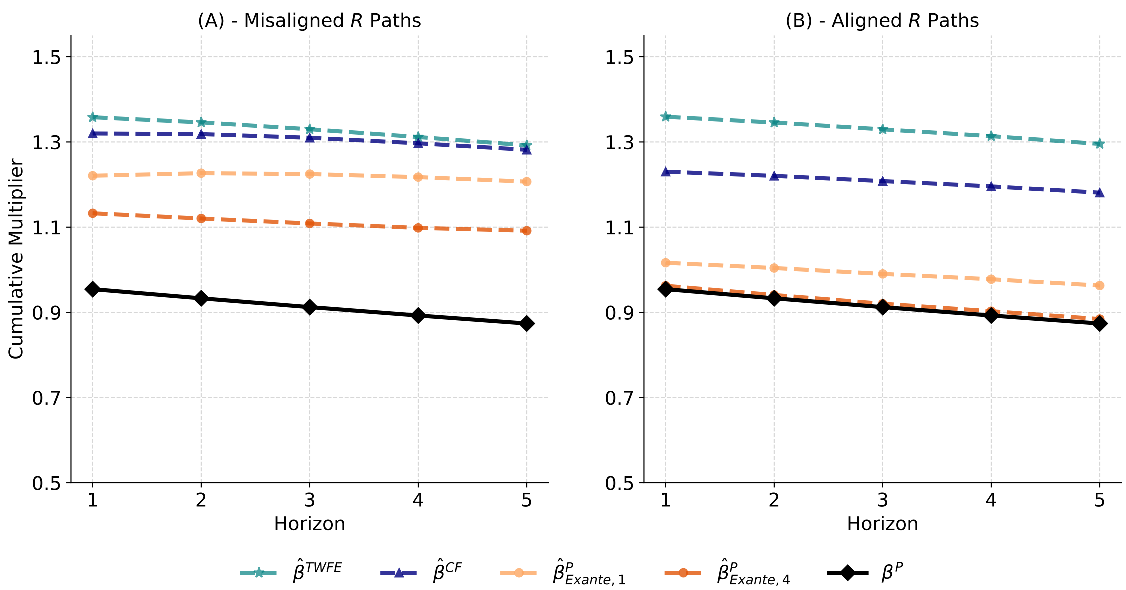


FIGURE 3. DSGE Monte Carlo Simulations - Results

The left (*right*) panel corresponds to a calibration where the path of R is *misaligned* (*aligned*) across aggregate shocks (See Figure A3). The solid black line corresponds to β^P the population portable elasticities. The dashed lines correspond to different estimation methods. $\hat{\beta}^{TWFE}$: TWFE approach; $\hat{\beta}^{CF}$: control function approach; $\hat{\beta}^{P}_{Exante,k}$: ex-ante decomposition using k shocks.

Extensions. Appendix A.5 presents several extensions of the simulation analysis. First, I show that the decomposition framework consistently recovers the portable elasticity as the number of available news shocks increases, under both parameterizations. Second, I examine environments with weaker forward-looking dynamics to confirm that the precision loss from limited information is smaller in those settings. Finally, in Subsection 5.3, I explore, within the limited-information setting, which observable data moments are informative about the accuracy of the framework and whether the method tends to over- or under-estimate the population portable elasticity, using the structure of the two-region

New Keynesian model.

3.4. Asymptotic Properties and Inference

This section summarizes the asymptotic properties of the estimators and the corresponding inference procedure. A formal discussion of consistency is provided in Appendix A.6. The argument proceeds in three steps. First, consistency of the panel local projection estimators is established under large-N, fixed-T asymptotics. Second, consistency of the time-series local projection estimators is established under large-T asymptotics. Third, $\hat{\beta}_h^P$ is shown to be a continuous mapping of these estimators and therefore consistent by the Continuous Mapping Theorem under mild non-degeneracy conditions.

The estimated portable elasticities at horizon h depend on both panel and time-series coefficients across horizons. For inference, I jointly estimate the cross-sectional and time-series parameters using a just-identified GMM system, which facilitates inference across moment conditions. This approach allows me to explicitly account for the covariance between panel and time-series moment conditions induced by shared aggregate shocks by clustering at the time level.

For the covariance between panel moment conditions associated with different parameters, the framework can flexibly accommodate alternative clustering schemes (e.g., time, unit, or two-way clustering). For the time-series moments, I use a HAC-consistent estimator as the baseline.

Standard errors for the HEDGE terms and the portable elasticities are obtained via the Delta method applied to the cluster-robust covariance matrix of the moment conditions. Monte Carlo simulations show that this procedure delivers (i) standard errors equivalent to those from the control function approach in static settings, and (ii) correct coverage in dynamic environments. A detailed exposition of the inference procedure is provided in Appendix A.7.

Recap. This section outlined an estimation strategy to decompose cross-sectionally identified elasticities into a portable component and an HEDGE component. The framework combines cross-sectional and time-series analysis. In the cross-sectional step, it estimates the differential responses to innovations in the variable of interest, G, and to innovations in the variable(s) that the researcher wants to difference out, R, conditional on exposure to the main variable of interest s_{ig}. In the time-series step, it estimates the impulse response function of R to both the shock of interest, ϵ_{gt}, and to innovations to R itself, ϵ_{rt}. Finally, the output of these steps is combined to estimate the HEDGE component. I discuss how to apply each of these steps under different assumptions about the data-generating process and about the information available to the researcher. Using DSGE Monte Carlo simulations, I show that the decomposition framework identifies the portable elasticity. In cases where the researcher has access to limited information, the decomposition framework works as an approximation to the true portable elasticity, and it represents an improvement on both baseline estimates that ignore HEDGE and the correction implied by the control function approach.

4. Empirical Application - US Cross-sectional Fiscal Multiplier

To illustrate the use of the decomposition framework, I revisit the estimation of the US cross-sectional fiscal multiplier through the lens of an HEDGE economy. In particular, I evaluate the possibility that states in the US differ not only in how exposed they are to fiscal shocks but also to changes in interest rates. I focus on the link between government spending and interest rates for two reasons. First, the strength of the monetary policy response to government spending shocks (i.e., the degree of *monetary offset*) is an important determinant of the aggregate fiscal multiplier. The response of monetary policy is undoubtedly among the set of aggregate variables that we aim to difference out by using cross-sectional methods (Nakamura and Steinsson 2014). Importantly, there is consensus in the profession that available estimates of the cross-sectional fiscal multiplier are independent of monetary policy, and the method proposed in this paper is designed to test whether this is indeed the case. Second, it is well documented that interest rate sensitivities vary substantially across economic regions (e.g. Carlino and DeFina (1998); Almgren et al. (2022); Herreno and Pedemonte (2025)), making it a natural starting point to explore the sensitivity of cross-sectional fiscal multipliers to HEDGE.

The empirical results should therefore be interpreted as isolating a single, yet economically central source of GE feedback—monetary policy. In this sense, the decomposition recovers a *monetary-policy–portable multiplier*, i.e., the cross-sectional effect of government spending holding the monetary response fixed. While other aggregate channels may also generate heterogeneous general equilibrium effects, I abstract from them in this application and leave their analysis to future work. Importantly, as shown below, removing this channel already restores the ability to use cross-sectional estimates to study counterfactual monetary policy regimes.

For estimation, I use the dataset built by Dupor and Guerrero (2017). The authors construct a panel dataset of US state-level defense contracts between 1951 and 2014.[25] Gross State Product data are sourced from BEA and available from 1963 onward. Following Sarto (2024), I exclude four small-population states with extreme variation in Gross State Product. Including all states yields similar point estimates with larger variance.[26] Both

[25] Data between 1951-2009 is sourced from two reports: the Prime Contract Awards by State report and the Atlas/Data Abstract for the US and Selected Areas. Both were compiled by the Directorate for Information Operations and Reports. Data for 2010-2014 are sourced from USASpending.gov.

[26] The excluded states are North Dakota (0.23% of the US population), South Dakota (0.27%), Wyoming (0.18%), and Alaska (0.23%). These states display multiple observations with GSP yearly changes over 20%.

output and defense contracts are deflated using the national Consumer Price Index and scaled by state population. I complement this dataset with data on the Federal Funds rate, total nominal federal tax receipts, nominal GDP, and the series of Romer and Romer (2004) monetary policy shocks as extended by Wieland and Yang (2020). The frequency of analysis is yearly and covers the period 1969-2007.[27]

I construct output and defense spending variables to estimate cumulative fiscal multipliers following the recommendations of Ramey and Zubairy (2018). Concretely, the cumulative percentage change in variable X_{it} relative to a baseline level of output Y_{it-1} is computed as:

$$X^C_{it+h} = \frac{[\sum_{j=0}^{h} X_{it+j}] - (h+1)X_{it-1}}{Y_{it-1}} \times 100,$$

where the subscript i indexes US states and t indexes years. The aggregate variables are constructed analogously. I measure the state-level exposure to defense spending, s_{igt}, following Nakamura and Steinsson (2014):

$$s_{igt} = \frac{G^{pc}_{it}/Y^{pc}_{it}}{G^{pc}_{t}/Y^{pc}_{t}},$$

where G^{pc}_{it} is per-capital real defense spending in state i during year t and Y^{pc}_{it} is per-capita real output in state i during year t. G^{pc}_{t} and Y^{pc}_{t} are the corresponding national counterparts. The idea is that a state is relatively more exposed to defense spending if its share of local defense spending to GDP is higher than the corresponding national share. I follow Nakamura and Steinsson (2014) and define s_{ig} as the average of this ratio over the first five years of their sample (i.e., 1966-1970).

First, I estimate the cross-sectional fiscal multiplier using the TWFE regression.

$$Y^C_{it+h} = \beta^{TW}_h \times G^C_{it+h} + \lambda_{ih} + \lambda_{th} + \text{Controls} + e_{it+h}, \tag{54}$$

where Y^C_{it+h} is the cumulative change in real output per capita in state i, G^C_{it+h} is the cumulative change in per capita defense contracts in state i, and λ_{ih}, λ_{th} are horizon-specific fixed effects for each state and year. I follow the literature and instrument G^C_{it+h} with $s_{ig}G^C_t$, where s_{ig} denotes state i's exposure to aggregate defense spending. The instrument interacts this predetermined exposure with the contemporaneous aggregate change in defense contracts, G^C_t, thus isolating variation in regional spending driven by changes in aggregate defense spending. I control for the lag of real per-capita output growth. The parameter of interest is β^{TW}_h, which measures the response of local output to an increase

Figure A18 in the Appendix presents results including all states.

[27]The sample is limited by the time-span of the Romer-Romer monetary policy shocks.

in local government spending equivalent to 1% of local output - the cross-sectional fiscal multiplier. The superscript *TW* is used to denote that this estimate corresponds to the TWFE specification.

The standard practice in this literature is to cluster standard errors at the state (unit) level. However, as shown in Almuzara and Sancibrián (2024), the appropriate clustering dimension in shift-share settings with a common aggregate shock corresponds to the time dimension. Common aggregate shocks induce correlation in residuals across units within a given period. I conduct Monte Carlo simulations in Appendix A.7 that confirm this result: clustering at the unit level leads to substantial undercoverage at all horizons.

Accordingly, I report baseline results using time-level clustering for the panel estimates and follow the inference procedure outlined in Appendix A.7. This approach generally delivers larger standard errors than unit-level clustering.[28]

The left panel of Figure 8 plots the estimation results. The TWFE multiplier is 1.5 at the two-year horizon and above one at all shorter and longer horizons.

Next, I present evidence of HEDGE with respect to monetary policy responses and implement the decomposition framework as outlined in Section 3.2 using a sequence of counterfactual contemporaneous shocks and as outlined in Section 3.3 using only date *zero* shocks. In the following, I provide further details on each step.

4.1. HEDGE Test

I test for HEDGE in monetary policy responses by estimating the following local projections:

$$Y^{C}_{it+h} = c_h \times s_{ig} i_t + \lambda_{ih} + \lambda_{th} + \text{Controls} + e_{it+h} \qquad \text{for} \quad h = 1, 4, \tag{55}$$

where i_t is the Federal Funds Rate. For estimation, I instrument $s_{ig} i_t$ with $s_{ig} RR_t$, where RR_t are the Romer-Romer monetary shocks.

Figure 4 shows that the estimated response is positive and statistically significant at all horizons. This finding rejects the null of homogeneous exposure to interest rate changes and suggests that time fixed effects alone may not fully absorb monetary policy shocks in this setting. In particular, the results indicate that states more exposed to defense spending, as measured by s_{ig}, are less responsive to changes in interest rates.

One possible explanation relates to differences in the cyclicality of government and private demand. If government purchases are less cyclical than private consumption, states with higher exposure to defense spending may be less sensitive to interest rate fluctuations. Another possibility is cross-state income differences: Figure A10 in Appendix B shows that per capita income is positively correlated with s_{ig}, and to the extent that

[28] For comparison, Figure A16 reports results using unit-level clustering.

higher income is associated with weaker liquidity constraints, this pattern is consistent with existing evidence for the U.S. (Herreno and Pedemonte 2025) and for Europe (Almgren et al. 2022).

Robustness. Appendix C.1 presents detailed results for each of the following robustness checks. First, in Figure A12, I show that the estimated response is robust to including the interaction between government spending and state-level exposure $s_{ig}G_t^C$. Second, I re-estimate equation (55) using the monetary shock series from Aruoba and Drechsel (2024) for the period 1980-2007. Figure A13 shows a positive estimated response, although less precisely estimated, possibly due to the shorter sample. Third, I estimate state-specific output elasticities to the Romer and Romer (2004) shock using local projections. Figure A14, shows that the estimated state-level elasticities are positively correlated with state-level exposure to G and that the correlation is statistically significant.

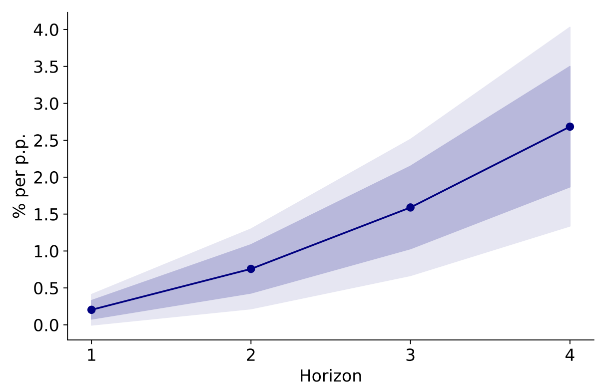


FIGURE 4. HEDGE Test - Response of Y_i to i_t, conditional on s_{ig}

Shaded area represents 68% (darker) and 90% (lighter) confidence bands. Plot corresponds to the cumulative percentage differential response of local output to a percentage point increase in the Federal Funds rate, evaluated at $s_{ig} = 1$.

The previous test indicates that in this empirical setting the cross-sectional condition for HEDGE to represent an identification challenge is met: s_{ig} drives cross-sectional responses to both defense and monetary policy shocks. Next, I show that the time-series condition is also met, namely, that the interest rate moves in response to a defense spending shock, which motivates the use of the decomposition framework.

4.2. Time-Series Step

The following step estimates the time series response of the interest rate to defense spending shocks and monetary policy shocks. First, this will inform whether the interest rate is effectively a GE variable in this empirical setting. I find this to be the case, so I use these reduced-form impulse responses to construct two different sequences of counterfactual monetary surprises. The first one, following Subsection 3.2, finds the sequence of ex-post contemporaneous surprises to the interest rate that replicate its observed path after a defense spending shock. I refer to this as the ex-post decomposition (Sims and Zha 2006). Second, following Subsection 3.3, I find the magnitude of a period *zero* monetary surprise that better matches the estimated path of the interest rate after a defense spending shock in the least squares sense. I refer to this as the ex-ante decomposition (McKay and Wolf 2023; Barnichon and Mesters 2023).

First, I run the following local projections:

$$i_{t+h} = \nu_h G_t^C + a_h RR_t + \text{Controls} + e_{t+h}, \tag{56}$$

where i_t is the Federal Funds rate, G_t^C is the one-year change in national defense contracts relative to GDP and RR_t are the series of Romer and Romer (2004) monetary policy shocks. The set of controls includes lagged realizations of the dependent variable, the Romer-Romer shock, per-capita real output (in logs), the average federal tax rate, CPI inflation, and per-capita real defense spending (in logs). Standard errors are HAC robust.

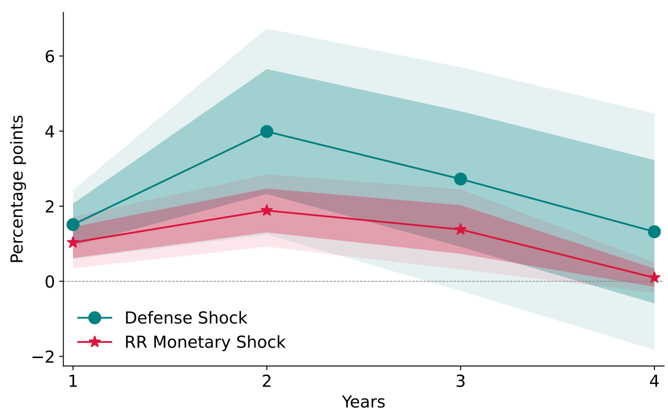


FIGURE 5. Time-series Step - Interest Rate Responses

Shaded area represents 68% (darker) and 90% (lighter) confidence bands. The green line shows the response of the nominal interest rate to a defense spending shock equal to 1% of lagged national output. The red line shows the response to a Romer–Romer monetary policy shock of one standard deviation.

Figure 5 shows the impulse response functions of the interest rate to each aggregate shock. In response to a defense spending shock equivalent to 1% of lagged national output, the interest rate increases by around 1.5 percentage points on impact.[29] A one-standard-deviation Romer–Romer shock raises the interest rate by 1 percentage point on impact, with a peak response of approximately 1.9 percentage points in the second year.

[29]The response to a one standard deviation increase in G_t^C is of approximately 0.42 percentage points.

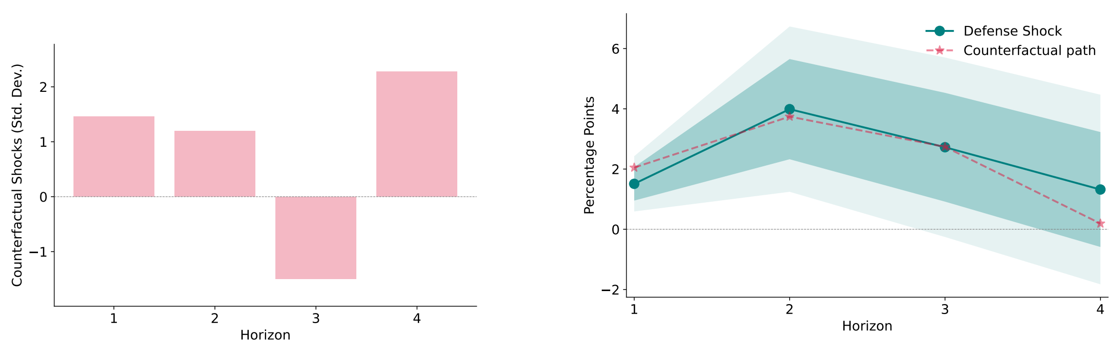
Importantly, the shape of these two impulse responses is relatively similar, which, as I discuss in Subsection 5.3, indicates that the loss of precision in applying the decomposition without access to news shocks should be relatively small.

In order to implement the ex-post decomposition, I employ the estimated dynamic response to the Romer-Romer monetary shock to compute up to $H = 4$ counterfactual monetary surprises that replicate the interest rate response to the defense spending shock. Figure 6A displays the resulting sequence of shocks, denoted $\tilde{\boldsymbol{e}}$.

For the ex-ante decomposition, I only use a single counterfactual shock $\tilde{\boldsymbol{e}}^0$ that hits the economy at $h = 0$ and propagates according to $\hat{a}_h$. This involves solving the minimization problem (57) to find the size of the Romer–Romer shock that best replicates the path of the interest rate following the defense shock. Setting $F = 0$, the minimization problem becomes:

$$\tilde{\boldsymbol{e}}^0 = \arg\min_{e \in R} \sum_{h=0}^{H} \left(\hat{v}_h - \hat{a}_{h0}\, e\right)^2, \tag{57}$$

where $\hat{a}_{h0} = \hat{a}_h$ are the estimated responses to a Romer-Romer shock. I find that a Romer-Romer shock of $\tilde{\boldsymbol{e}}^0 = 1.98$ standard deviations minimizes (57). Figure 6B plots the resulting ex-ante counterfactual path for the interest rate, alongside its estimated response to the defense shock. The counterfactual path remains within one standard deviation of the target response at all horizons.



A. Ex-post Counterfactual R Shocks

B. Ex-ante Counterfactual Path of R

FIGURE 6. Time Series Step - Counterfactual Computations

Panel (A) shows the sequence of estimated counterfactual monetary shocks for the ex-post decomposition. The units are standard deviations of Romer-Romer monetary shocks. Panel (B) shows the counterfactual path for R implemented by the ex-ante decomposition with $\tilde{\boldsymbol{e}}^0 = 1.98$. *Defense Shock* is the estimated response of the interest rate to the defense shock with its corresponding 68% and 90% confidence bands.

4.3. Cross-sectional step

I jointly estimate the cross-sectional responses to defense spending shocks and interest rate shocks using the following reduced-form specification:

$$Y^C_{it+h} = b_h \times s_{ig} G^C_t + c_h \times s_{ig} RR_t + \lambda_{ih} + \lambda_{th} + \text{Controls} + e_{it+h}, \tag{58}$$

where RR_t is the series of Romer and Romer (2004) monetary shocks and G^C_t is the one-year change in aggregate defense contracts. I include the same set of controls as in the TWFE specification. The coefficient c_h captures the differential response of cumulative output to a one standard deviation in the monetary shock for the region with exposure equal to the national average. I separately estimate the first-stage regression for local government spending using:

$$G^C_{it+h} = \theta_h \times s_{ig} G^C_t + \lambda_{ih} + \lambda_{th} + Controls + e_{it+h}, \tag{59}$$

where θ_h captures the first-stage coefficient at horizon h. I add as controls all variables included in the reduced form specification (58). I proceed in two steps to align the units from the time-series and cross-sectional reduced-form estimates.

Figure 7 shows the percentage response of local output to a monetary shock of one standard deviation, evaluated at $s_{ig} = 1$. The estimated differential response is positive and significant at all horizons, consistent with the findings in the HEDGE test above. The corresponding two-stage least squares responses to the defense shock, $\hat{b}_h/\hat{\theta}_h$, are nearly identical to the TWFE estimates and are therefore relegated to Figure A15 in Appendix C.2.[30]

4.4. Portable Multiplier Estimate

Lastly, I combine the results of the previous two steps to construct estimates for the portable cross-sectional multiplier up to $H = 4$. Formally, $\hat{\beta}^P_h$ is given by:

$$\hat{\beta}^P_{h,Ex\,post} = \frac{\hat{b}_h - \sum_{k=0}^{h} \hat{c}_{h-k} \times \tilde{e}_{r,t+k}}{\hat{\theta}_h}, \qquad \hat{\beta}^P_{h,Exante} = \frac{\hat{b}_h - \hat{c}_h \times \tilde{\boldsymbol{e}}^0}{\hat{\theta}_h}, \tag{60}$$

where $\tilde{e}_{r,t+k}$ is the ex-post monetary surprise in period $t+k$ and $\tilde{\boldsymbol{e}}^0$ is the period *zero* monetary shock. Dividing by the first stage coefficient expresses results in terms of changes to local government spending as is usual in this literature. Figure 8 shows the results for each version of the decomposition and for the TWFE specification.[31] The control function

[30] These estimates may differ from $\hat{\beta}^{TW}_h$ if there is in-sample correlation between the Romer-Romer shock and any of the other regressors included in the specification.

[31] See Figure A16 for unit-level clustering.

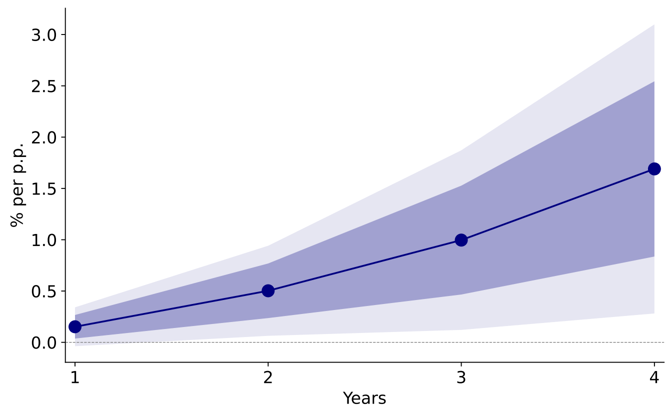


FIGURE 7. Response of Y_i to RR_t, conditional on s_{ig}

Shaded area represents 68% (darker) and 90% (lighter) confidence bands. Standard errors are computed as described in Appendix A.7 with time-level clustering for panel moment conditions. The solid line shows the cumulative percentage response of local output to a one standard deviation Romer-Romer shock, evaluated at $s_{ig} = 1$. The specification controls for lagged output growth.

approach, shown for completeness, delivers estimates that are virtually identical to those of the TWFE strategy. The portable multiplier is below the TWFE multiplier at all horizons for both versions. This is for two reasons. First, it reflects the fact that $\hat{c}_h > 0$ which means that US states that are more exposed to defense spending also react less to changes in interest rates. Second, the interest rate is estimated to increase in response to a defense spending shock. Together, both forces push the cross-sectional multipliers estimated using the TWFE strategy above the estimated portable multipliers. For example, at the two-year horizon, the TWFE multiplier is 1.5 while the estimated portable multiplier is 1.

Lastly, Figure 9 shows the estimated HEDGE terms for each version of the decomposition expressed in the same units as the cross-sectional multiplier:

$$\tilde{\Omega}_{h,Expost} = \frac{\hat{\Omega}_{h,Expost}}{\hat{\theta}_h}, \qquad \tilde{\Omega}_{h,Exante} = \frac{\hat{\Omega}_{h,Exante}}{\hat{\theta}_h}, \tag{61}$$

where $\hat{\theta}_h$ are the first stage coefficients from (59). The estimated HEDGE terms are statistically different from zero at the 90% level for $h > 1$ for both versions of the decomposition. For comparison, Figure A17 reports results clustering at the unit-level.

Robustness. Appendix C.2 presents results from a series of robustness checks. First, Figure A16 and Figure A17 report the baseline specification using unit-level clustering for panel moments instead of time-level clustering. As expected, this choice yields narrower confidence bands. Second, I show that results are robust to using the real Federal Funds rate instead of the nominal rate. Figure A19 shows that the point estimates for

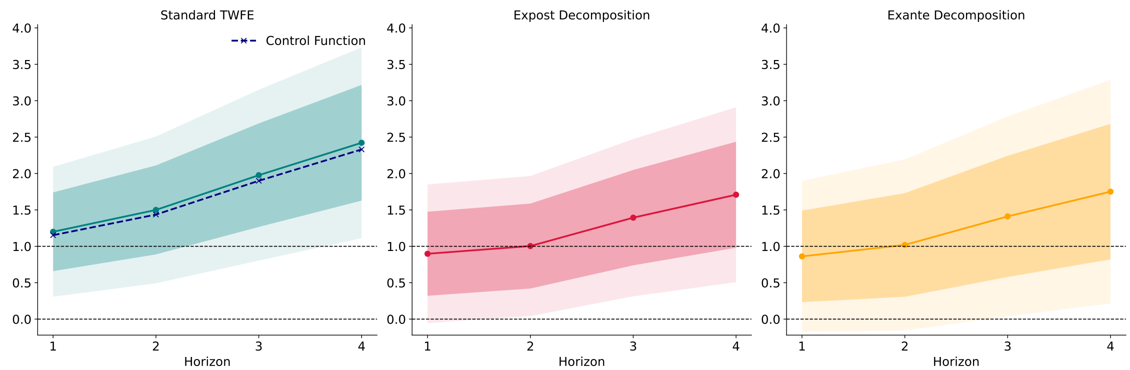


FIGURE 8. Decomposition Results for the US Cross-Sectional Multiplier

Shaded area represents 68% (darker) and 90% (lighter) confidence bands. Standard errors are computed as described in Appendix A.7 with time-level clustering for panel moment conditions. Point estimates show the cumulative percentage change in local output when local defense spending increases by 1% of local output. *Ex-post Decomposition* shows results for the decomposition outlined in Subsection 3.2. *Ex-ante Decomposition* shows results for the decomposition outlined in Subsection 3.3 using a single shock. The blue dotted line shows the results from the control function approach.

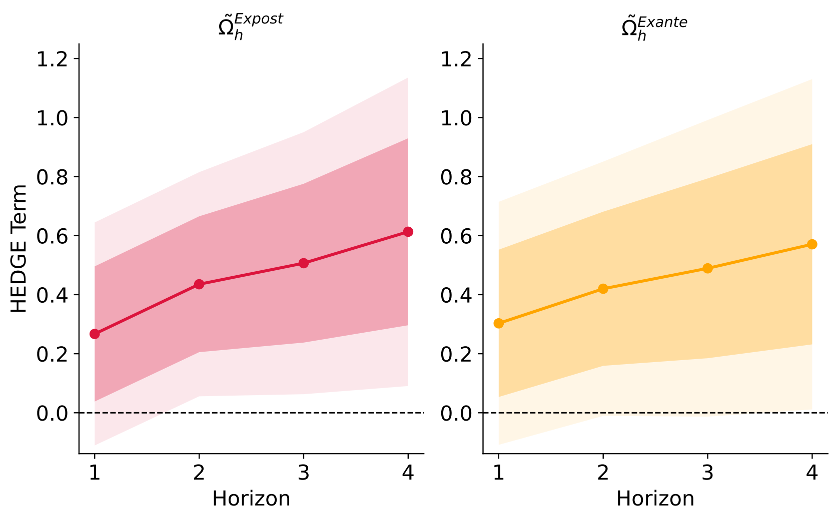


FIGURE 9. Decomposition Results - Estimated HEDGE Terms

The plot shows the point estimate and confidence bands for $\tilde{\Omega}_{h,Expost}$ on the left-hand side and for $\tilde{\Omega}_{h,Exante}$ on the right-hand side. Standard errors are computed as described in Appendix A.7 with time-level clustering for panel moment conditions.

the portable multiplier remain largely unchanged, although the ex-post decomposition becomes noisier. Third, Figure A20 shows that the results are robust to using the period average value of s_{igt} instead of the average of 1966-1970. Fourth I present results using two alternative sets of controls in the cross-sectional regressions: (i) no controls, (ii) baseline plus lagged shock(s). Figure A21 shows results when no controls are included - in line with the baseline choices of, for example, Nakamura and Steinsson (2014); Dupor and Guerrero

(2017); Auerbach, Gorodnichenko, and Murphy (2020). This choice results in significantly larger estimates of the cross-sectional fiscal multiplier relative to my baseline specification, which controls for lagged output growth. The fact that controlling for lagged output dampens the cross-sectional multiplier was first noted by Ramey (2020). Overall, both the TWFE and portable multipliers shift up, yet the magnitude of the estimated HEDGE term remains stable. Figure A22 corresponds to the baseline specification augmented by adding lags of the corresponding aggregate shocks, interacted with s_{ig}. Adding lagged shocks further shifts downward both the TWFE and portable estimated multipliers, in particular, both drop below 1 at the two-year horizon. However, the estimated HEDGE term is of the same magnitude as in the baseline specification, albeit much less precisely estimated. Fifth, I present the results for seven alternative control choices detailed in Table A8. Figure A23A plots the baseline estimate for the TWFE alongside each of the seven alternatives. Figures A23B and A23C repeat this exercise for the ex-ante portable multiplier and the HEDGE term, respectively. Across specifications, the estimated HEDGE terms fall within one-standard deviation of the baseline estimates.

Recap. This section revisits the estimation of cross-sectional fiscal multipliers using state-level data for the US between 1969-2007. I show that taking into account the heterogeneous effects of monetary policy across US states matters for the estimation of cross-sectional fiscal multipliers. In particular, I estimate a portable multiplier that is meaningfully lower than baseline TWFE estimates. This finding underscores the limitations of relying solely on time fixed effects to difference-out GE effects, in particular, those operating through systematic monetary responses.

5. Model

This section develops a theoretical counterpart to the empirical decomposition of Section 4 using a two-region New Keynesian environment. The model provides a structural interpretation of the empirical objects by expressing the TWFE cross-sectional multiplier—denoted by $\mathscr{M}_{CS}$—the HEDGE term Ω, and the portable multiplier—denoted by $\mathscr{M}_{CS}^{P}$—in terms of model-implied responses to a fiscal policy shock. I use the model to decompose $\mathscr{M}_{CS}$ into three channels: substitution between regional goods; intertemporal substitution driven by regional inflation differentials; and the HEDGE channel, which arises when regions differ in their responsiveness to aggregate real interest rate movements. The HEDGE term enters as a residual general-equilibrium force that moves the model-analogue of the TWFE multiplier away from its portable counterpart.

Next, I show that HEDGE can cause cross-sectional multipliers to vary widely across monetary regimes, sometimes in directions opposite to those of the aggregate multiplier. This challenges the common interpretation of $\mathscr{M}_{CS}$, the TWFE cross-sectional multiplier, as partial equilibrium objects, a view often justified by the inclusion of time fixed effects. Crucially, when regions differ in their exposure to general-equilibrium forces, $\mathscr{M}_{CS}$ reflects not only direct policy effects, but also heterogeneous responses to GE adjustments. As a result, the property highlighted in prior work (Chodorow-Reich 2019; Dupor et al. 2023) that cross-sectional multipliers can serve as upper or lower bounds on the aggregate multiplier, depending on the monetary environment, is no longer guaranteed when HEDGE is present and not taken into account. In contrast, the portable multiplier $\mathscr{M}^{P}_{CS}$ - for which this paper provides an identification strategy - retains several key properties: GE invariance, validity as a bound on the aggregate multiplier, and usefulness for model comparison.

Finally, I treat the two-region New Keynesian model as a useful approximation to the empirical environment and ask: which observable data moments best predict the precision loss in estimating the HEDGE term under limited information? The model reveals that a sufficient statistic, the cumulative discrepancy between the realized and counterfactual path of R, is highly informative about the estimation error in the HEDGE term. I then calibrate the model to match the empirical value of this moment. This allows me to evaluate how the decomposition performs in a data-generating process that mirrors the empirical setting in Section 4. The results suggest that, conditional on the model environment, the method delivers accurate estimates of the portable multiplier even in the presence of strong forward-looking dynamics.

Environment. I now present a general model structure, beginning with its sequence-space representation and structural blocks. The derivations follow McKay and Wolf (2023); Wolf (2023a), and the two-region[32] model is based on Nakamura and Steinsson (2014). The economy consists of two regions and is characterized by a set of $n_k \times 2$ endogenous regional variables $\{K_t\}$ and a set of n_w endogenous aggregate variables $\{W_t\}$. There are two structural aggregate shocks $s \in \{w_1, w_2\}$ that represent changes in aggregate variables $w_{1t}, w_{2t} \in W_t$. The time path of a shock s is denoted by ϵ_s, and $\boldsymbol{\epsilon} = (\epsilon_{w_1}, \epsilon_{w_2})$ denotes the full path of the shock. Each entry j in ϵ_s corresponds to an innovation in w_{st} that is announced at time $t = 0$ and implemented at time $t = j$.[33] Let $dx = \{x_t\}_{t=0}^{\infty}$ denote the perfect-foresight path of a variable x, and let $dx_{|\boldsymbol{\epsilon}}$ be the corresponding path conditional on the shock sequence $\boldsymbol{\epsilon}$. The regional block of the model is characterized by

$$H_k dK + H_w dW = \mathbf{0}, \tag{62}$$

[32]The framework can be generalized to arbitrary N regions.

[33]For example, if ϵ_i is the vector of innovations to the nominal interest rate, then the first entry captures a contemporaneous surprise, while the second represents news received at $t = 0$ about a change in the rate at $t = 1$.

where the vector dK collects the time paths of $k_r \in K_t$ and similarly for dW. The matrices H_k and H_w condense the relationships that characterize the regional block.[34] The aggregate block is characterized by

$$A_k dK + A_w dW + A_\epsilon \boldsymbol{\epsilon} = \mathbf{0}. \tag{63}$$

The above formulation assumes that the structural shocks *only* affect the regional block through their effect on the set of aggregate variables W_t. This assumption is important because it ensures that these shocks serve as valid instruments for the observed realizations of the policy or aggregate variables.

I denote by $\boldsymbol{H}_z^x$ the partial equilibrium Jacobian of z with respect to x. This matrix maps changes in the time-path of variable x to changes in variable z, holding all else equal. Each column j of this matrix gives the path of z in response to a one-unit innovation to x announced at $t = 0$ and implemented at $t = j$. The general equilibrium Jacobian of z with respect to the structural shock ϵ_{w_1} is denoted by $\boldsymbol{J}_z^{\epsilon_{w_1}}$. Each column j of this matrix contains the general equilibrium response of an endogenous variable z to a change in w_1 announced in period $t = 0$ and implemented in period $t = j$.

5.1. Anatomy of the NK Cross-sectional Fiscal Multiplier

I use these objects to characterize the determinants of the cross-sectional fiscal multiplier in a standard two-region New Keynesian model. I take as a baseline the model in Nakamura and Steinsson (2014) and extend it to allow regional heterogeneity in household responses to the real interest rate. While the full model features both Ricardian and hand-to-mouth households, the derivations below assume a representative household in each region for tractability. The main features of the model are summarized here, with full details presented in the Appendix D.

The economy consists of two regions of equal size n, denoted H (Home) and F (Foreign). In each region, a representative household has separable preferences over labor and consumption and CES preferences over domestically and foreign-produced goods, each of which is itself a CES aggregate of locally produced varieties. Financial markets are complete, which implies perfect risk sharing across regions.

The federal government purchases domestic and foreign varieties using the same CES demand structure of households. Aggregate government spending is subject to an idiosyncratic shock ϵ_g that increases spending heterogeneously across space: each additional dollar of spending increases purchases of the Home (Foreign) good by $s_H \times n$ ($s_F \times n$) dollars. The government levies lump sum taxes to ensure the budget balances in each

[34] I omit regional shocks and unobserved variables for notational simplicity but the environment can be generalized to accommodate those features.

period.

On the supply side, local firms produce differentiated varieties using local labor as the sole input. Firms compete monopolistically and set prices à la Calvo, giving rise to two regional New Keynesian Phillips curves. Lastly, the nominal interest rate is set by a monetary authority and is subject to an idiosyncratic shock ϵ_i.

Aggregate government spending, G, follows an exogenous process governed by:

$$d\boldsymbol{G} = [\boldsymbol{A}_G^G]^{-1}\epsilon_g, \tag{64}$$

where $\boldsymbol{A}_G^G$ allows for a general auto-regressive structure. This implies the following process for regional government consumption g_r:

$$\begin{aligned} d\boldsymbol{g_r} &= \boldsymbol{H}_{g_r}^G d\boldsymbol{G} \\ &= s_r\ d\boldsymbol{G} \qquad \forall\ \ r = H, F. \end{aligned} \tag{65}$$

The scalar s_r captures the exposure of the region r to aggregate government purchases. The nominal interest rate is set according to the following inertial Taylor rule:

$$d\boldsymbol{i} = \boldsymbol{A}_i^Y d\boldsymbol{Y} + \boldsymbol{A}_i^\Pi d\boldsymbol{\Pi} + \boldsymbol{A}_i^{\epsilon_i}\epsilon_i, \tag{66}$$

where Y and Π denote aggregate output and national inflation, respectively (i.e., the weighted sum of the corresponding regional counterparts). Total consumption in region r is determined by the Euler equation:

$$d\boldsymbol{c_r} = \boldsymbol{H}_{c_r}^r d\boldsymbol{r} + \boldsymbol{H}_{c_r}^q d\boldsymbol{q} \quad \forall\ \ r = H, F. \tag{67}$$

Consumption depends on the path of the national real interest rate, $r \in \boldsymbol{W}$, and the real exchange rate, $q \in \boldsymbol{K}$, which captures regional inflation differentials. The real exchange rate reflects the gap between the national and local real rates induced by differences in regional CPI inflation.

The path of local output can be expressed as a function of regional demand, regional government purchases, and relative prices across regions.

$$d\boldsymbol{Y_r} = \boldsymbol{H}_{y_r}^{c_H} d\boldsymbol{c_H} + \boldsymbol{H}_{y_r}^{c_F} d\boldsymbol{c_F} + \boldsymbol{H}_{y_r}^{g_r} d\boldsymbol{g_r} + \boldsymbol{H}_{y_r}^q d\boldsymbol{q} \quad \forall\ \ r = H, F. \tag{68}$$

The first two terms on the right-hand side capture the effects of changes in regional consumption demand. The third term captures the effect of regional government purchases on local output. The last term measures the effect of changes in q, the real exchange rate between regions.

Let $\boldsymbol{\epsilon}_g$ denote a time path of structural shocks such that $\epsilon_g = [1, 0, ..0]$ and $\epsilon_i = \boldsymbol{0}$.

In other words, there is a contemporaneous one-time increase in ϵ_g that propagates according to (64) and no further shocks hit the economy. We want to compute the cross-sectional output multiplier, $\mathscr{M}_{CS}$, in response to $\boldsymbol{\epsilon}_g$:

$$\mathscr{M}_{CS} = \frac{d\boldsymbol{y}_{\boldsymbol{H}|\epsilon_g} - d\boldsymbol{y}_{\boldsymbol{F}|\epsilon_g}}{d\boldsymbol{g}_{\boldsymbol{H}|\epsilon_g} - d\boldsymbol{g}_{\boldsymbol{F}|\epsilon_g}}. \tag{69}$$

where division between vectors is understood element-wise across horizons. Since $d\boldsymbol{g}_r = s_r d\boldsymbol{G}$, we have:

$$d\boldsymbol{g}_H - d\boldsymbol{g}_F = (s_H - s_F)d\boldsymbol{G}, \tag{70}$$

which yields

$$\mathscr{M}_{CS} = \frac{d\boldsymbol{y}_{\boldsymbol{H}|\epsilon_g} - d\boldsymbol{y}_{\boldsymbol{F}|\epsilon_g}}{(s_H - s_F)\, d\boldsymbol{G}_{|\epsilon_g}}, \tag{71}$$

where $\mathscr{M}_{CS}$ measures the relative change in local output per unit of differential government spending between regions after an aggregate fiscal shock.[35] Substituting in the structural expressions for local output and consumption yields a decomposition into three components: two relative price effects and one interest rate sensitivity term.

$$\mathscr{M}_{CS} = d\mathbf{1} + \frac{1}{(s_H - s_F)\, d\boldsymbol{G}_{|\epsilon_g}} \left[(\Theta_{IS} + \Theta_{ES}) d\boldsymbol{q}_{|\boldsymbol{\epsilon}_g} + \vartheta [\boldsymbol{H}^r_{c_H} - \boldsymbol{H}^r_{c_F}] d\boldsymbol{r}_{|\epsilon_g} \right], \tag{72}$$

where $d\mathbf{1}$ is a vector of ones with the same dimension as $d\boldsymbol{G}$. In this model, the cross-sectional fiscal multiplier can deviate from unity through three channels.[36]

First, an *expenditure switching* channel, which captures how households reallocate consumption between Home and Foreign goods in response to changes in relative prices. Its strength is governed by Θ_{ES}, a function of the elasticity of substitution across Home and Foreign goods and the degree of home bias in preferences. A positive fiscal shock raises the relative price of goods produced in the more exposed region, leading households to substitute toward the less exposed region. This substitution dampens the local output response, reducing the multiplier.

The second is a *relative intertemporal substitution* channel that reflects the effect of differences in local real rates on consumption. These differences arise due to heterogeneous consumption baskets across regions - for example, from home-biased preferences - which

[35] Without loss of generality, the above formula can be modified to compute either cumulative multipliers or multipliers relative to the change in government purchases on impact.

[36] In the full model, the presence of *hand-to-mouth* households adds an additional *Keynesian* channel that, generally, amplifies the cross-sectional multiplier.

cause region-specific CPI inflation responses. This channel operates through movements in the real exchange rate and its strength is governed by Θ_{IS}. Under home bias, the relative intertemporal channel also has a negative effect on the size of the multiplier. The fiscal shock generates expected relative disinflation in the more exposed region, raising its relative real interest rate and lowering its relative consumption.[37] Like the ES channel, this channel tends to reduce the multiplier.

The third channel, which I refer to as the *HEDGE* channel, also operates through intertemporal substitution but captures differences in sensitivity to the national real rate between regions. This channel is non-zero only when regions differ in their interest rate sensitivity. For example, with CRRA utility and separable preferences, this would be the case with region-specific intertemporal elasticities of substitution (IES). Its strength is governed by two components: (i) ϑ, a function of the degree of home bias in preferences; and (ii) the heterogeneity in real rate sensitivity. Unlike the other two channels, HEDGE can amplify or dampen the cross-sectional multiplier, depending on the direction of the real rate movement and the correlation between exposure to fiscal policy and interest rate sensitivity.

The portable multiplier is defined as the TWFE cross-sectional multiplier net of the HEDGE channel:

$$\mathscr{M}_{CS}^{P} = \mathscr{M}_{CS} - \underbrace{\frac{\vartheta[\boldsymbol{H}_{c_H}^{r} - \boldsymbol{H}_{c_F}^{r}]}{(s_H - s_F)} \frac{d\boldsymbol{r}_{|\epsilon_g}}{d\boldsymbol{G}_{|\epsilon_g}}}_{\text{HEDGE Channel}} . \tag{73}$$

This mirrors the decomposition in equation (2) and identifies the model-based analogue to the empirical HEDGE term, Ω_h, and to the portable elasticity, β_h^P. The first term in the HEDGE channel captures the relative sensitivity to the real rate, normalized by fiscal exposure. The second captures the general equilibrium response of the real rate to government spending shocks.

To compute the model analogue of the empirical HEDGE term Ω_h, I isolate the contribution of the real interest rate channel to the differential in cross-regional output $d\boldsymbol{y}_{\boldsymbol{H}|\epsilon_g} - d\boldsymbol{y}_{\boldsymbol{F}|\epsilon_g}$. I ask: What portion of the differential would arise if the only effect of the fiscal shock were the general equilibrium path of the real rate, holding all other channels constant? To answer this, I construct a counterfactual sequence of monetary policy shocks that reproduces the real rate path $d\boldsymbol{r}_{|\epsilon_g}$ induced by the original fiscal shock. This allows me to simulate the output effects of the HEDGE channel in isolation and recover the

[37] On impact, the spending shock increases relative producer price inflation for the exposed region. Regional convergence requires that this initial difference in inflation rates across regions be compensated for by expected relative disinflation in the future. That is, there is a decrease in expected future producer price inflation in the exposed region relative to the less exposed region. With home bias, this translates into a decrease in expected consumer price inflation for the exposed region because its consumption basket is loads more heavily on its locally produced good.

model-implied path of Ω.

Under the assumption that structural shocks affect regional outcomes only through their impact on aggregate variables, the HEDGE term can be computed using the general equilibrium response of local output to a time-path of monetary innovations $\tilde{\boldsymbol{\epsilon}}_i$ that replicates $d\boldsymbol{r}_{|\boldsymbol{\epsilon}_g}$. Letting $\boldsymbol{G}_r^{\epsilon_i}$ denote the general equilibrium Jacobian of the real rate with respect to the monetary shock ϵ_i, I compute the following:

$$\tilde{\boldsymbol{\epsilon}}_i = [\boldsymbol{G}_r^{\epsilon_i}]^{-1} d\boldsymbol{r}_{|\boldsymbol{\epsilon}_g}, \tag{74}$$

where $\tilde{\boldsymbol{\epsilon}}_i$ is the vector of contemporaneous and news shocks that implements the desired real rate path. Given this sequence, the region-specific output responses are:

$$d\boldsymbol{y}_{r|\tilde{\boldsymbol{\epsilon}}_i} = \boldsymbol{G}_{y_r}^{\epsilon_i} \tilde{\boldsymbol{\epsilon}}_i \quad \forall \ r = H, F. \tag{75}$$

The implied model counterpart to Ω_h is given by:

$$\boldsymbol{\Omega} = \frac{d\boldsymbol{y}_{H|\tilde{\boldsymbol{\epsilon}}_i} - d\boldsymbol{y}_{F|\tilde{\boldsymbol{\epsilon}}_i}}{(s_H - s_F) d\boldsymbol{G}_{|\boldsymbol{\epsilon}_g}}, \tag{76}$$

where $\boldsymbol{\Omega} = \{\Omega_h\}_{h=0}^{H}$ denotes the vector of model-implied HEDGE terms across horizons. Subtracting this vector from the TWFE multiplier $\mathscr{M}_{CS}$ yields $\mathscr{M}_{CS}^{P}$, the model-analogue of the portable cross-sectional multiplier.

5.2. Pitfalls of Failing to Control for HEDGE

The previous discussion highlighted how the HEDGE channel may dampen or amplify the cross-sectional fiscal multiplier relative to a counterfactual economy where regions are equally responsive to interest rate changes. In this section, I use the model to study the implications of failing to control for HEDGE on (i) the dispersion of cross-sectional fiscal multipliers and on (ii) their relation to the aggregate fiscal multiplier - denoted by $\mathscr{M}_A$.

The first point relates to the monetary policy invariance that has been attributed to the available estimates of the cross-sectional multiplier. In HEDGE economies, the New Keynesian model can generate cross-sectional fiscal multipliers $\mathscr{M}_{CS}$, which are just as dispersed as the aggregate multipliers.

The exercise is as follows. Consider three different monetary policy rules that have different implications for the covariance between government spending shocks and real interest rates. These are the same three rules studied by Nakamura and Steinsson (2014). In the first case, the monetary authority follows an inertial Taylor rule with a weight on inflation of $\phi_\pi = 1.5$ and on the output gap of $\phi_y = .1$. This monetary rule implies an aggressive response to the inflationary pressures that follow a government spending shock. For the second case, I assume that the monetary authority keeps the national real rate fixed at its steady-state level in response to government spending shocks (i.e. $i_t - E_t\left[\Pi_{t+1}^{agg}\right] = r_{ss} = i_{ss}$). In the third case, the monetary authority keeps the nominal interest rate fixed at its steady-state level in response to government spending shocks (i.e. $i_t = i_{ss}$). This last scenario is akin to an economy that hits the Zero Lower Bound (ZLB).

In addition, I consider two scenarios for the covariance between exposure to government spending shocks and to interest rate changes. In both cases, the shock is to ϵ_g but the difference comes from how the aggregate shock loads in each region. I refer to the case where the more interest rate sensitive region is more exposed to government spending as *high IES*, and to the reverse case as *low IES*. Essentially, I solve the model for each combination of monetary rule and spending shock structure to compute the corresponding cross-sectional and aggregate fiscal multipliers. The calibration is the one used for the DSGE Monte Carlo exercise in Subsection 3.3, detailed in Table A9.

TABLE 2. Fiscal Multipliers & Monetary Policy with HEDGE

	Government Spending Shock	
Monetary Rules	*High IES*	*Low IES*
Taylor Rule		
Portable Cross Sectional Multiplier - $\mathcal{M}_{CS}^{P}$	.95	.95
TWFE Cross Sectional Multiplier - $\mathcal{M}_{CS}$	.6	1.3
Aggregate Multiplier - $\mathcal{M}_{A}$	.64	.64
Fixed Nominal Rate		
Portable Cross Sectional Multiplier - $\mathcal{M}_{CS}^{P}$	.95	.95
TWFE Cross Sectional Multiplier - $\mathcal{M}_{CS}$	1.07	.82
Aggregate Multiplier - $\mathcal{M}_{A}$	1.12	1.14
Fixed Real Rate		
Portable Cross Sectional Multiplier - $\mathcal{M}_{CS}^{P}$	.95	.95
TWFE Cross Sectional Multiplier - $\mathcal{M}_{CS}$	.95	.95
Aggregate Multiplier - $\mathcal{M}_{A}$	1	1

The table shows one year multipliers for alternative monetary rules and cross-sectional loadings of the government spending shock.

Table 2 presents the model-analogues of the TWFE and portable cross-sectional multipliers, as well as the aggregate multiplier for each scenario. The portable cross-sectional multiplier is, by definition, invariant to monetary policy and to the structure of the government spending shock. However, the TWFE cross-sectional multiplier depends on both the monetary response and on which region is more exposed to the spending shock. These multipliers range between .6 and 1.3, whereas the aggregate multiplier, which also depends on the monetary rule, ranges between .64 and 1.14. This exercise illustrates that HEDGE can introduce substantial variability in cross-sectional multipliers depending on the direction of GE responses and the correlation between cross-sectional exposures.

The results in Table 2 also show how general equilibrium forces that attenuate the aggregate effect of a fiscal shock, like the aggressive monetary response implied by the Taylor rule case, can actually increase the estimated cross-sectional effect, conditional on the structure of the government spending shock. For example, a researcher studying the effect of a government spending shock tilted to the *low IES* region in a ZLB environment would estimate a smaller cross-sectional multiplier than a researcher doing so under normal monetary policy. The opposite is true with a shock tilted to the *high IES* region. The counterintuitive and context-specific way in which HEDGE affects cross-sectional estimates brings us to the second point of this section. Namely, how informative are TWFE cross-sectional estimates of the aggregate fiscal multiplier?

Previous work has argued that under conventional monetary policy, such as a Taylor rule that responds aggressively to inflation, the cross-sectional multiplier is likely to serve as an upper bound on the aggregate multiplier. The reason is that TWFE cross-sectional estimators do not include the contractionary effects of monetary tightening, which typically dominate the dampening effects of expenditure switching or relative intertemporal substitution. Conversely, at the ZLB, monetary policy is unresponsive, so cross-sectional estimates tend to understate the aggregate effects, acting as a lower bound. Table 2 shows that these conclusions may no longer hold once HEDGE is present and is not accounted for. In particular, changes in the bias of the government spending shock or the differential interest rate sensitivities across regions can flip these bounding relationships in either direction.

This highlights a key point: When general equilibrium responses to monetary policy contaminate cross-sectional estimates, the sign and magnitude of the resulting bias become context-specific. As a result, TWFE cross-sectional multipliers cannot be reliably interpreted as bounds on aggregate effects. In contrast, the portable multiplier $\mathscr{M}^{P}_{CS}$, which I show how to identify empirically, retains its validity to bound the aggregate fiscal multiplier regardless of the monetary policy regime.

5.3. Performance of the Decomposition with Limited Information

In general, access to news shocks is limited. For this reason, I use the model to propose a series of data moments that can be used to assess the performance of the framework in settings with limited access to news shocks. In particular, I simulate data from the model to study how the method performs in relation to the following two questions:

a. How close is the estimated HEDGE term to its theoretical value?
b. Are the estimated portable elasticities an upper or lower bound for the theoretical portable elasticities?

Throughout the exposition, I consider a researcher who applies the ex-ante decomposition using a single date *zero* innovation in the GE variable.

How close is the estimated HEDGE term to its theoretical value?. I measure the overall discrepancy between the estimated and true HEDGE term as follows.

$$D_{\Omega} = H^{-1} \sum_{h=0}^{H} \frac{\Omega_{t+h} - \hat{\Omega}_{t+h}}{\Omega_{t+h}}. \tag{77}$$

D_{Ω} measures the average percent deviation between the estimated and true HEDGE term up to horizon H. I express it as a percent to facilitate comparison across calibrations with HEDGE terms of different size.

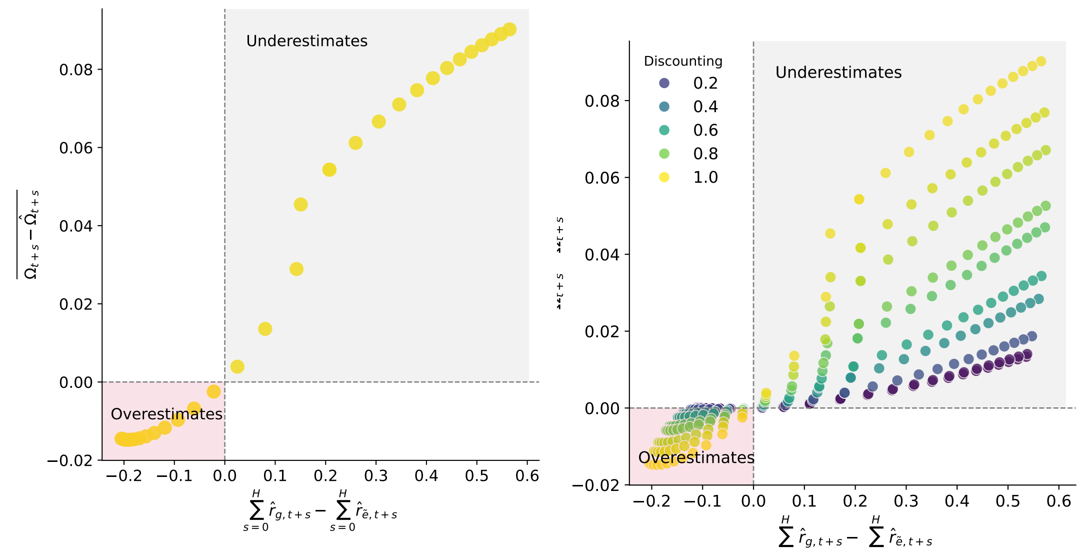
A sufficient statistic for the overall performance of the decomposition is the difference between the integral of the path for the GE variable in response to the shock of interest and the integral of its counterfactual path using a single date *zero* shock, namely:

$$D_R = \sum_{h=0}^{H} \boldsymbol{R}_{g,h} - \sum_{h=0}^{H} \tilde{\epsilon}_{r,0} \boldsymbol{R}_{r,h}. \tag{78}$$

where $\tilde{\epsilon}_{r,0}$ the size of the counterfactual shock that solves the least squares projection in Proposition 2. The intuition behind this result is straightforward: the better we enforce the deviation in R from its pre-shock value, the closer the estimation is to the true HEDGE term. Figure 10A plots D_R against D_{Ω} for different model calibrations. These calibrations differ only on the assumed persistence for the shock processes. The calibration associated with a smaller D_R yields a better overall performance as measured by D_{Ω}.

Are the estimated portable elasticities an upper or lower bound for the theoretical portable elasticities?. Figure 10A shows that the sign of D_R is informative about whether the estimated HEDGE term understates or overstates the true HEDGE term. In particular, when D_R is positive and the deviations of R from the pre-shock values are smaller under the counterfactual than the realized path, then $\hat{\boldsymbol{\Omega}}$, on average, underestimates the true $\boldsymbol{\Omega}$. The reverse is true when D_R is negative. The intuition behind this result is that using a counterfactual path that does not *perturb* the economy enough leads to an underestimation of the HEDGE term. In such a case, the estimated portable elasticity is an upper bound for the true portable elasticity, in absolute terms.

The preceding discussion evaluated performance taking as given the strength of forward-looking dynamics in the economy. Next, I repeat this exercise using simulated data from economies with different degrees of forward-looking forces. I extend the baseline model to allow for a discount term, $\alpha \in (0, 1]$, in the Euler equation (McKay, Nakamura, and Steinsson 2017, 2016). The lower α, the more muted the effect of expected future consumption on current consumption decisions - with $\alpha = 1$ in the baseline model. Figure 10B shows the simulation results for the same set of calibrations as Figure 10A but considering different values for α. First, the two data moments discussed above apply equally well



A. Degree of Misalignment - No Discounting

B. Strength of Anticipation Effects

FIGURE 10. Performance with Limited Information

The x-axis measures the cumulative differential between the path of R after the shock of interest and the counterfactual implied from a decomposition using only a single date zero shock. The y-axis shows the average (across time) difference between the true model-implied HEDGE term and its estimated value. Positive values on the y-axis imply that on average the estimated HEDGE term is smaller than the true value. Panel (A) plots results for a model without discounting in the Euler equation. Panel (B) plots results for models with different degrees of discounting in the Euler Equation. A higher value (e.g. yellow dots) implies stronger forward-looking dynamics.

regardless of the strength of forward-looking forces. Second, the plot shows that given a value for D_R, the cost of using limited information decreases as forward-looking forces become less important.

In summary, the previous analysis shows how data moments that are easy to compute can be used to inform the performance of the method, given the modeling assumptions.

5.4. Empirical Results Through the Lens of the Model

I take advantage of the mapping between the model and the estimation framework to evaluate the precision of my empirical findings from Section 4. Concretely, I calibrate the model to match the empirical value of D_R in order to answer: *How well does the decomposition with a single shock perform in this model economy?*

To perform the model-based decomposition, I extend the two-region RANK model of Nakamura and Steinsson (2014) to a two-region economy with two types of households: Ricardian and hand-to-mouth (H2M). The environment thus combines the key features of Nakamura and Steinsson (2014) with the two-region TANK model of Herreno and Pedemonte (2025), who analyze heterogeneous regional responses to monetary policy

TABLE 3. Calibration - Matching D_R in the Model and Data

n	Size of Home Region	.5	Equal sized regions
s_H, s_F	Regional exposure to ϵ_g	.8, 1.2	ϵ_g tilted towards Foreign region
$\bar{\sigma}$	Average IES	1	Nakamura and Steinsson (2014)
σ_H	Home IES	1.2	Home relatively more i sensitive
σ_F	Foreign IES	.8	Foreign relatively less i sensitive
λ_r	Share of *Hand-to-mouth*	.2	Common across regions
ρ_{ϵ_i}	Persistence of Monetary Policy Shock	.75	Set to match D_R

This table corresponds to the calibration used in Subsection 5.4. It details the calibration for parameter values that are either new to the model or that differ from Nakamura and Steinsson (2014). The remaining parameters are detailed in Table A9.

in a setting without fiscal shocks. I assume an equal share of hand-to-mouth households across regions. These households play a key role in disciplining the level of the portable multiplier, which I target to be around one at year two, consistent with the empirical estimates in Section 4.

To generate cross-regional heterogeneity in the sensitivity to monetary policy, I introduce region-specific intertemporal elasticities of substitution for Ricardian households. This allows the model to replicate the observed variation in interest rate sensitivity and its correlation with fiscal exposure. In this calibration, the gap between the TWFE and portable cross-sectional multipliers, i.e., the HEDGE term, is driven by the combination of the real rate response to the fiscal shock and the heterogeneity in Ricardians' IES across regions.

In the following, I outline the rest of the calibration details. Most of the parameters are set to their values in Nakamura and Steinsson (2014). These parameters are detailed in Table A9 in Appendix D. Table 3 summarizes the parameters that differ from those in Nakamura and Steinsson (2014) or are newly introduced in my extended model. I consider equally sized regions such that $n = .5$. These equally sized regions are heterogeneously exposed to national government spending shocks instead of region-specific government spending shocks. This means that when government spending increases, it increases everywhere but with different intensities across space, in line with the time-series variation used in Section 4. I set the exposure shifters to national government spending shocks to $s_H = .8$ and $s_F = 1.2$. Therefore, government spending shocks are biased towards the Foreign region as $s_F > s_H$. These parameters have no direct counterpart in Nakamura and Steinsson (2014) because the authors use region-specific government spending shocks.[38] To generate differential sensitivities to interest rate changes, I allow Ricardian households to have region-specific IES. I set the average IES for Ricardian households equal to 1

[38]Alternatively, one can think of region-specific shocks as setting $s_k = 1$ and $s_{k'} = 0$. In addition, this parameter becomes irrelevant for homogeneous regions, and the only requirement for identification of the cross-sectional multiplier is $s_H \neq s_F$.

as in Nakamura and Steinsson (2014) and the region-specific IES to $\sigma_H = 1.2$ and $\sigma_F = .8$. The Home region is relatively more sensitive to changes in the interest rate. This calibration reproduces the correlation between the sensitivities to fiscal and monetary policy shocks found in Section 4. This exercise targets the sign of the correlation between fiscal and monetary sensitivity rather than the dynamic path of cross-sectional interest rate responses, a task that I leave for future work. Lastly, the share of hand-to-mouth households is set at $\lambda_r = .2$ for both regions to target the level of the portable multiplier close to unity on the two-year horizon.

Crucial for this exercise are the parameters that govern the response of the real rate to government spending and monetary policy shocks. To minimize my departure from the calibration in Nakamura and Steinsson (2014), I match D_R by choosing the persistence for the monetary policy shock, leaving the process for government spending unchanged. This requires the monetary policy shock to follow an AR(1) with persistence $\rho_{\epsilon_i} = .75$. Formally,

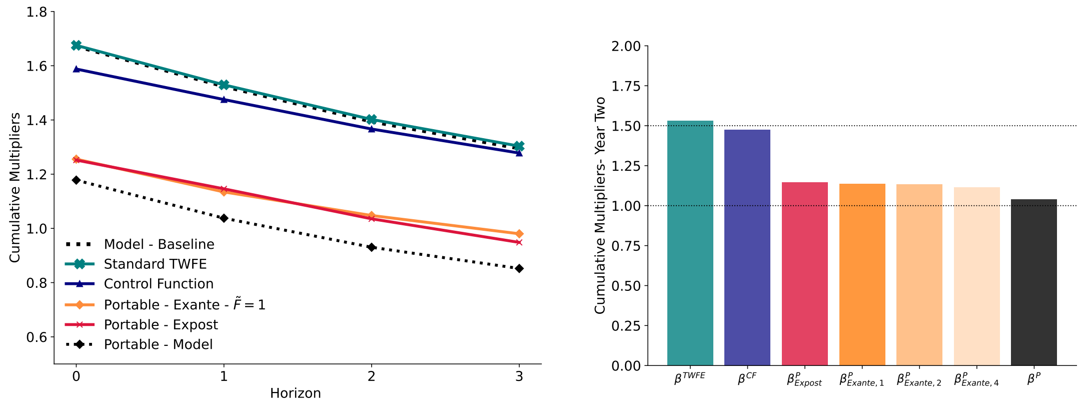
$$i_t = \rho_i i_{t-1} + (1 - \rho_i)[i_{ss} + \phi_\pi \Pi_t^{agg} + \phi_y \hat{Y}_t^{agg}] + u_{it}, \qquad u_{it} = \rho_{\epsilon_i} u_{it-1} + \epsilon_{it}. \tag{79}$$

The nominal interest rate follows an Taylor rule subject to an AR(1) shock, and national government spending follows an exogenous AR(1) process, as in Nakamura and Steinsson (2014).

Next, I simulate data from the model and apply the decomposition framework assuming that only a contemporaneous interest rate shock is available. Figure 11 displays the results. The left panel plots the cumulative fiscal multipliers across horizons, comparing the true model-implied portable multiplier (black dotted line) with several alternative estimators. Both the TWFE estimator and the control function approach exhibit upward bias relative to the true portable multiplier, driven by the positive HEDGE term. In contrast, estimates obtained by applying either the ex post decomposition or the ex ante decomposition with a single shock correct for most of the omitted variable bias, bringing estimated multipliers much closer to the true portable effect.

Importantly, both versions of the decomposition perform well even in a setting with strong forward-looking dynamics (McKay, Nakamura, and Steinsson 2016; Del Negro, Giannoni, and Patterson 2023). This is consistent with the findings of the previous subsection, which showed that a small D_R - that is, a small discrepancy between the realized and counterfactual path of R -implies a limited precision loss under restricted information.

Figure 11B zooms in on the year-two multiplier and quantifies the extent to which each method corrects for the HEDGE term. While the TWFE and control function estimators overstate the true portable multiplier by approximately 0.5 points, the decomposition-based estimators align much more closely with the model-implied benchmark. In particular, allowing for additional news shocks to the monetary policy rule ($F = 2$ or $F = 4$) yields only modest improvements relative to the single-shock decomposition. This highlights that when the counterfactual path of R can be closely approximated, even with limited instruments, the decomposition method delivers accurate estimates of the portable elasticities.



A. Estimated Multipliers in the Model B. Zoom in - 2-year Cumulative Multiplier

FIGURE 11. Decomposition Results using Model Simulated Data

Panel (A) shows the estimates of the cross-sectional fiscal multipliers for different estimation strategies, together with the true model implied values. Panel (B) zooms in at the 2-year horizon multipliers shown in the left panel. Results are based on 100 repetitions.

6. Conclusion

This paper addresses the identification challenge that arises in cross-sectional empirical strategies when units' exposure to multiple aggregate variables is driven by common exposure shifters. In such scenarios, the typical empirical design that relies on cross-sectional variation in exposure combined with time fixed effects fails to separately identify portable or partial equilibrium effects from general equilibrium driven responses. This challenges the *portability* of these estimated elasticities, limiting their external validity and their usefulness for policy analysis.

To overcome this problem, I introduce a decomposition framework that explicitly separates portable effects from general equilibrium driven effects by jointly exploiting cross-sectional and time-series variation. The method clarifies the conditions under which portable statistics can be consistently estimated and implemented under varying informational assumptions ranging from static to forward-looking dynamic environments.

Applying this framework to US data, I estimate the cross-sectional fiscal multiplier and demonstrate that accounting for general equilibrium effects operating through monetary responses significantly reduces the estimated cross-sectional multiplier. This finding challenges the conventional assumption in this literature that monetary policy responses are fully absorbed by time-fixed effects, highlighting the importance of explicitly accounting for monetary-fiscal interactions when estimating cross-sectional multipliers.

More broadly, the framework can be applied beyond fiscal policy to settings where cross-sectional estimates may be shaped by general-equilibrium feedback—such as monetary, trade, or credit-supply shocks. By clarifying how to recover portable elasticities in HEDGE economies, it provides a foundation for reinterpreting and extending existing work that combines micro data with aggregate shocks. Future research can use this approach to obtain clean empirical counterparts of the objects required for model calibration, policy evaluation, and counterfactual analysis.

PAULA DONALDSON
University of California, San Diego

References

Adao, Rodrigo, Costas Arkolakis, and Federico Esposito. 2019. "General equilibrium effects in space: Theory and measurement.", National Bureau of Economic Research.

Almgren, Mattias, José-Elías Gallegos, John Kramer, and Ricardo Lima. 2022. "Monetary policy and liquidity constraints: Evidence from the euro area." *American Economic Journal: Macroeconomics* 14 (4): 309–340.

Almuzara, Martín, and Víctor Sancibrián. 2024. "Micro responses to macro shocks." *FRB of New York Staff Report* (1090).

Arkhangelsky, Dmitry, and Vasily Korovkin. 2024. "On Policy Evaluation With Aggregate Time-Series Instruments.", CEPR Discussion Papers.

Aruoba, S Borağan, and Thomas Drechsel. 2024. "Identifying monetary policy shocks: A natural language approach.", National Bureau of Economic Research.

Auerbach, Alan, and Yuriy Gorodnichenko. 2015. "How powerful are fiscal multipliers in recessions?" *NBER Reporter* (2): 21–24.

Auerbach, Alan, Yuriy Gorodnichenko, and Daniel Murphy. 2020. "Local fiscal multipliers and fiscal spillovers in the USA." *IMF Economic Review* 68: 195–229.

Autor, David, David Dorn, and Gordon H Hanson. 2021. "On the persistence of the China shock.", National Bureau of Economic Research.

Autor, David H, David Dorn, and Gordon H Hanson. 2013. "The China syndrome: Local labor market effects of import competition in the United States." *American economic review* 103 (6): 2121–2168.

Bachmann, Rüdiger, and Eric R Sims. 2012. "Uncertainty and the effectiveness of fiscal policy in the US and Europe." *American Economic Review* 102 (3): 60–65.

Bai, Jushan. 2009. "Panel data models with interactive fixed effects." *Econometrica* 77 (4): 1229–1279.

Barnichon, Régis, and Geert Mesters. 2023. "A sufficient statistics approach for macro policy." *American Economic Review* 113 (11): 2809–2845.

Borusyak, Kirill, and Peter Hull. 2023. "Nonrandom exposure to exogenous shocks." *Econometrica* 91 (6): 2155–2185.

Canova, Fabio. 2024. "Should we trust cross-sectional multiplier estimates?" *Journal of Applied Econometrics* 39 (4): 589–606.

Canova, Fabio, and Evi Pappa. 2011. "Fiscal Policy, Pricing Frictions and Monetary Accommodation." *Economic Policy* 26 (1): 555–598.

Caravello, Tomás E, Alisdair McKay, and Christian Wolf. 2024. "Evaluating policy counterfactuals: A var-plus approach.", National Bureau of Economic Research.

Carlino, Gerald, and Robert DeFina. 1998. "The differential regional effects of monetary policy." *Review of economics and statistics* 80 (4): 572–587.

Chodorow-Reich, Gabriel. 2019. "Geographic Cross-Sectional Fiscal Spending Multipliers: What Have We Learned?" *American Economic Journal: Economic Policy* 11 (2): 1–34.

Chodorow-Reich, Gabriel. 2020. "Regional data in macroeconomics: Some advice for practitioners." *Journal of Economic Dynamics and Control* 115: 103875.

Chodorow-Reich, Gabriel, Laura Feiveson, Zachary Liscow, and William Gui Woolston. 2012. "Does state fiscal relief during recessions increase employment? Evidence from the American Recovery and Reinvestment Act." *American Economic Journal: Economic Policy* 4 (3): 118–145.

Chodorow-Reich, Gabriel, Plamen T Nenov, and Alp Simsek. 2021. "Stock market wealth and the real economy: A local labor market approach." *American Economic Review* 111 (5): 1613–1657.

Christiano, Lawrence, Martin Eichenbaum, and Sergio Rebelo. 2011. "When is the government spending multiplier large?" *Journal of Political Economy* 119 (1): 78–121.

Cloyne, James S, Òscar Jordà, and Alan M Taylor. 2020. "Decomposing the fiscal multiplier.", National Bureau of Economic Research.

Del Negro, Marco, Marc P Giannoni, and Christina Patterson. 2023. "The forward guidance puzzle." *Journal of Political Economy Macroeconomics* 1 (1): 43–79.

DeLong, J Bradford, and Lawrence H Summers. 2012. "Fiscal policy in a depressed economy." *Brookings Papers on Economic Activity* 43 (1): 233–297.

Demyanyk, Yuliya, Elena Loutskina, and Daniel Murphy. 2019. "Fiscal stimulus and consumer debt." *Review of Economics and Statistics* 101 (4): 728–741.

Dupor, Bill, and Rodrigo Guerrero. 2017. "Local and aggregate fiscal policy multipliers." *Journal of Monetary Economics* 92: 16–30.

Dupor, Bill, Marios Karabarbounis, Marianna Kudlyak, and M Saif Mehkari. 2023. "Regional consumption responses and the aggregate fiscal multiplier." *Review of Economic Studies* 90 (6): 2982–3021.

Farhi, Emmanuel, and Iván Werning. 2016. "Fiscal multipliers: Liquidity traps and currency unions." In *Handbook of macroeconomics*, vol. 2, 2417–2492: Elsevier.

Guren, Adam M, Alisdair McKay, Emi Nakamura, and Jón Steinsson. 2021. "Housing wealth effects: The long view." *The Review of Economic Studies* 88 (2): 669–707.

Guren, Adam, Alisdair McKay, Emi Nakamura, and Jón Steinsson. 2021. "What do we learn from cross-regional empirical estimates in macroeconomics?" *NBER Macroeconomics Annual* 35 (1): 175–223.

Hack, Lukas, Klodiana Istrefi, and Matthias Meier. 2023. "Identification of systematic monetary policy."

Herreno, Juan, and Mathieu Pedemonte. 2025. "The geographic effects of monetary policy shocks." *Review of Economics and Statistics*: 1–46.

Jordà, Òscar. 2005. "Estimation and inference of impulse responses by local projections." *American economic review* 95 (1): 161–182.

Kato, Akihisa, Wataru Miyamoto, Thuy Lan Nguyen, and Dmitriy Sergeyev. 2018. "The Effects of Tax Changes at the Zero Lower Bound: Evidence from Japan." *AEA Papers and Proceedings* 108: 513–518.

Leeper, Eric M, Nora Traum, and Todd B Walker. 2017. "Clearing up the fiscal multiplier morass." *American Economic Review* 107 (8): 2409–2454.

Manuel, Ed. 2025. "Identification in General Equilibrium: The Missing Intercept Problem Revisited.", Working Paper.

Matthes, Christian, Naoya Nagasaka, and Felipe Schwartzman. 2024. "Estimating The Missing Intercept."

McKay, Alisdair, Emi Nakamura, and Jón Steinsson. 2016. "The power of forward guidance revisited." *American Economic Review* 106 (10): 3133–3158.

McKay, Alisdair, Emi Nakamura, and Jón Steinsson. 2017. "The discounted euler equation: A note." *Economica* 84 (336): 820–831.

McKay, Alisdair, and Wolf. 2023. “What can time-series regressions tell us about policy counterfactuals?” *Econometrica* 91 (5): 1695–1725.

Mian, Atif, and Amir Sufi. 2014. “What explains the 2007–2009 drop in employment?” *Econometrica* 82 (6): 2197–2223.

Nakamura, Emi, and Jón Steinsson. 2014. “Fiscal stimulus in a monetary union: Evidence from US regions.” *American Economic Review* 104 (3): 753–792.

Nakamura, Emi, and Jón Steinsson. 2018. “Identification in macroeconomics.” *Journal of Economic Perspectives* 32 (3): 59–86.

Ottonello, Pablo, and Thomas Winberry. 2020. “Financial heterogeneity and the investment channel of monetary policy.” *Econometrica* 88 (6): 2473–2502.

Parker, Jonathan A, Nicholas S Souleles, David S Johnson, and Robert McClelland. 2013. “Consumer spending and the economic stimulus payments of 2008.” *American economic review* 103 (6): 2530–2553.

Pennings, Steven. 2021. “Cross-region transfer multipliers in a monetary union: Evidence from social security and stimulus payments.” *American Economic Review* 111 (5): 1689–1719.

Pennings, Steven. 2022. “Locally financed and outside financed regional fiscal multipliers.” *Economics Letters* 213: 110389.

Peydró, José-Luis, Andrea Polo, and Enrico Sette. 2021. “Monetary policy at work: Security and credit application registers evidence.” *Journal of Financial Economics* 140 (3): 789–814.

Ramey, Valerie. 2020. “Discussion of Guren, McKay, Nakamura, Steinsson "What Do We Learn from Cross-Sectional Empirical Estimates in Macroeconomics?” *NBER Macro Annual* 35.

Ramey, Valerie A. 2016. “Macroeconomic shocks and their propagation.” *Handbook of macroeconomics* 2: 71–162.

Ramey, Valerie A. 2019. “Ten years after the financial crisis: What have we learned from the renaissance in fiscal research?” *Journal of Economic Perspectives* 33 (2): 89–114.

Ramey, Valerie A, and Sarah Zubairy. 2018. “Government Spending Multipliers in Good Times and in Bad: Evidence from US Historical Data.” *Journal of Political Economy* 126 (2): 850–901.

Riera-Crichton, Daniel, Carlos A Vegh, and Guillermo Vuletin. 2015. “Fiscal solvency and procyclical fiscal policy.” *Journal of Development Economics* 117: 305–319.

Romer, Christina D, and David H Romer. 2004. “A new measure of monetary shocks: Derivation and implications.” *American economic review* 94 (4): 1055–1084.

Sarto, Andrés Pablo. 2024. “Recovering macro elasticities from regional data.”

Sims, Christopher A, and Tao Zha. 2006. “Were there regime switches in US monetary policy?” *American Economic Review* 96 (1): 54–81.

Wieland, Johannes F, and Mu-Jeung Yang. 2020. “Financial dampening.” *Journal of Money, Credit and Banking* 52 (1): 79–113.

Wolf, Christian. 2019. “The Missing Intercept in Cross-Regional Regressions.”, Working Paper.

Wolf, Christian. 2023a. “Fiscal stimulus and the systematic response of monetary policy.” In *AEA Papers and Proceedings,* vol. 113: 388–393.

Wolf, Christian. 2023b. “The missing intercept: A demand equivalence approach.” *American Economic Review* 113 (8): 2232–2269.

Zidar, Owen. 2019. “Tax cuts for whom? Heterogeneous effects of income tax changes on growth and employment.” *Journal of Political Economy* 127 (3): 1437–1472.

Zwick, Eric, and James Mahon. 2017. "Tax policy and heterogeneous investment behavior." *American Economic Review* 107 (1): 217–248.

Appendix

Table of Contents

Appendix A. Decomposition Framework - Additional Derivations and Results

A.1. Static Setting - Further Details

Back to reference in main text.

This appendix presents the derivations for a generalized static setting with two-way feedback between aggregate variables. In the following, I reproduce the system of equations that characterize a static economic:

$$\text{(S1)} \quad \begin{aligned} G_t &= \alpha R_t + u_{gt} \\ R_t &= \delta G_t + u_{rt} \\ u_{kt} &\sim \mathrm{N}(0, \sigma_k^2) \quad \forall\ k = g, r \end{aligned} \qquad \begin{aligned} Y_{it} &= \beta \times s_{ig} G_t + \gamma \times s_{ir} R_t + u_i + u_t + u_{it}^y \\ s_{ig} &\sim \mathrm{N}(\bar{s}, \sigma_{s_g}^2) \\ s_{ir} &= \phi s_{ig} + u_i^{s_r} \\ u_i^{s_r} &\sim N(0, \sigma_{s_r}^2), \quad u_{it}^y \sim N(0, \sigma_y^2). \end{aligned}$$

With two-way feedback, the process for the aggregate variable G and R can be rewritten as:

$$G_t = \frac{1}{1-\delta\alpha} u_{gt} + \frac{\alpha}{1-\delta\alpha} u_{rt}, \tag{A1}$$

$$R_t = \frac{\delta}{1-\delta\alpha} u_{gt} + \frac{1}{1-\delta\alpha} u_{rt}. \tag{A2}$$

Next, I outline the general formulas for each step of the decomposition.

Cross-sectional Step. The reduced-form regression from the cross-sectional step is given by:

$$Y_{it} = b s_{ig} \epsilon_{gt} + c s_{ig} \epsilon_{rt} + \lambda_i + \lambda_t + e_{yt}. \tag{A3}$$

The regression coefficients $\hat{b}$ and $\hat{c}$ converge to:

$$E[\hat{b}] = \frac{\beta}{1-\alpha\delta} + \frac{\tilde{\gamma}\delta}{1-\alpha\delta}, \tag{A4}$$

$$E[\hat{c}] = \frac{\beta\alpha}{1-\alpha\delta} + \frac{\tilde{\gamma}}{1-\alpha\delta}. \tag{A5}$$

Note that, due to the presence of two-way feedback, we cannot decompose $\hat{b}$ into two additive terms as before. The first term on the right-hand side is a rescaled version of the portable elasticity where the rescaling captures the contemporaneous feedback between aggregate variables. Similarly for the HEDGE term. However, this is not an impediment to apply the decomposition as I show next.

Time Series Step. The estimated responses of R_t to each aggregate shock are given by:

$$\text{(A6)} \qquad R_t = \nu\epsilon_{gt} + a\epsilon_{rt} + e_{rt}, \qquad E[\hat{\nu}] = \frac{\delta}{1-\delta\alpha}, \qquad E[\hat{a}] = \frac{1}{1-\delta\alpha}.$$

The hypothetical ϵ_{rt} shock - denoted by $\tilde{\epsilon}_{rt}$ - satisfies the following:

$$\text{(A7)} \qquad \tilde{\epsilon}_{rt} = \frac{\hat{\nu}}{\hat{a}} = \delta.$$

Last Step. As pointed out above, with two-way feedback we cannot in general neatly decompose the TWFE estimate into two additive terms but we can still consistently estimate the portable elasticity in the same way as before. Evaluating the effect of $\tilde{\epsilon}_{rt}$ using $\hat{c}$ from the cross-sectional step and subtracting it from $\hat{b}$ yields:

$$\text{(A8)} \qquad \begin{aligned} \hat{\beta}^P &= \hat{b} - \hat{c} \times \tilde{\epsilon}_{rt} \\ &= \frac{\beta}{1-\alpha\delta} + \frac{\tilde{\gamma}\delta}{1-\alpha\delta} - \frac{\beta\alpha\delta}{1-\alpha\delta} - \frac{\tilde{\gamma}\delta}{1-\alpha\delta} \\ &= \beta. \end{aligned}$$

Monte Carlo Simulations with Two-Way Feedback. The performance of the decomposition is illustrated using Monte Carlo simulations. I simulate data from an economy characterized by the system of equations presented above with $\alpha \neq 0$ and estimate three different specifications: (i) Baseline, (ii) Decomposition, and (iii) Control Function. Baseline corresponds to the two-way fixed-effects regression. Decomposition and control function correspond to the results of applying the framework presented above or the control function approach. Figure A1 presents the distribution of the estimated coefficients for each strategy in an economy with $\beta = \gamma = \delta = \phi = \alpha = .5$. The population value for the portable effect is .5 and the omitted variable bias due to HEDGE is .33. There are two takeaways. First, the baseline estimates are biased and centered on $\beta^P + \Omega = .83$. Second, both the decomposition and control function approach consistently estimate β^P.

A.2. Decomposition a la Sims and Zha (2006) - Further Details

This appendix presents derivations for the *ex-post* decomposition in a more general economy than the AR(1) example in the main text.

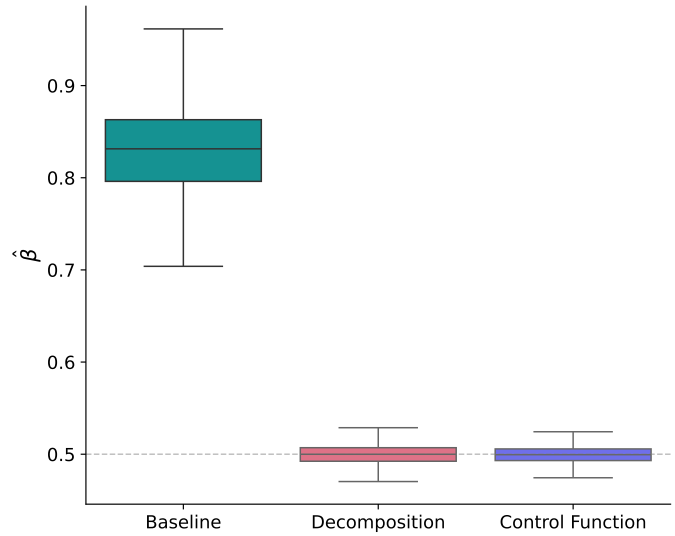


FIGURE A1. Static Monte Carlo Simulations with Two-Way Feedback

Distribution of point estimates for the *Baseline*, *Decomposition Framework* and *Control Function* estimation strategies. Based on 1000 repetitions with sample size of n = 100 and t = 300 and parameters set to $\delta = \beta = \gamma = \phi = \alpha = .5$.

Back to reference in the main text.

Let the data-generating process be characterized by

$$\begin{aligned}
s_{ig} &= u_i^{s_g}, \qquad s_{ir} = \phi s_{ig} + u_i^{s_r}, \\
G_t &= \sum_{j=1}^{J} \rho_{gj} G_{t-j} + u_{gt}, \\
R_t &= \sum_{m=1}^{M} \rho_{rm} R_{t-m}, + \sum_{n=0}^{N} \delta_n G_{t-n} + u_{rt}, \\
Y_{it+h} &= \sum_{s=0}^{S} \beta_s s_{ig} G_{t+h-s} + \sum_{k=0}^{K} \gamma_k s_{ir} R_{t+h-k} + \sum_{l=1}^{L} \psi_l Y_{i,t+h-l} + u_{it+h}^{y} \quad \forall h \geq 0, \\
u_{kt} &\sim \mathrm{N}(0, \sigma_k^2), \; u_i^{s_k} \sim \mathrm{N}(0, \sigma_{s_k}^2) \; \forall \; k = g, r, \quad u_{it}^{y} \sim \mathrm{N}(0, \sigma_y^2).
\end{aligned} \tag{A9}$$

The derivations that follow set $\psi_l = 0$ for notational tractability. This is without loss of generality.

The total, portable and HEDGE effects of a shock to G_t in period $t + h$ are given by:

$$\beta_h^T = \underbrace{\sum_{s=0}^{S} \beta_s \frac{\partial G_{t+h-s}}{\partial u_{gt}}}_{\beta_h^P} + \underbrace{\sum_{k=0}^{K} \gamma_k \phi \frac{\partial R_{t+h-k}}{\partial u_{gt}}}_{\Omega_h}. \tag{A10}$$

where the effects are evaluated at $s_{ig} = 1$. The total derivative of the cross-sectional outcome at horizon h with respect to u_{rt} is given by:

$$\gamma_h^T = \sum_{k=0}^{K} \gamma_k \phi \frac{\partial R_{t+h-k}}{\partial u_{rt}}. \tag{A11}$$

Reduced Form Analysis. Consider the following panel local projections with two-way fixed effects:

$$Y_{it+h} = b_h \times s_{ig} \epsilon_{gt} + c_h \times s_{ig} \epsilon_{rt} + \lambda_{ih} + \lambda_{th} + e_{it+h}, \tag{A12}$$

where ϵ_{gt} and ϵ_{rt} are the innovations to G and R, respectively. The regression coefficient $\hat{b}_h$ converges to:

$$E[\hat{b}_h] = \sum_{s=0}^{S} \beta_s \frac{cov[G_{t+h-s}, \epsilon_{gt}]}{v[\epsilon_{gt}]} + \sum_{k=0}^{K} \gamma_k \phi \frac{cov[R_{t+h-k}, \epsilon_{gt}]}{v[\epsilon_{gt}]}, \tag{A13}$$

$\forall\ s \in S$ such that $h - s \geq 0$ and $\forall\ k \in K$ such that $h - k \geq 0$. The regression coefficient $\hat{c}_h$ converges to:

$$E[\hat{c}_h] = \sum_{k=0}^{K} \gamma_k \phi \frac{cov[R_{t+h-k}, \epsilon_{rt}]}{v[\epsilon_{rt}]}, \tag{A14}$$

$\forall\ k \in K$ such that $h - k \geq 0$. If $\frac{cov[R_{t+h}, \epsilon_{rt}]}{v[\epsilon_{rt}]} = \frac{cov[R_{t+h}, \epsilon_{gt}]}{v[\epsilon_{gt}]}$ for all $h \in H$, then the cross-sectional step directly provides us with an estimate of Ω_h at each horizon. However, whenever the conditional path of R given each aggregate shock differs, an additional correction is needed. Implementing this correction requires information from time-series analysis. Basically, we need estimates for the path of R conditional on the realization of each aggregate shock.

The time-series step estimates the path of R conditional on each aggregate shock using:

$$R_{t+h} = v_h \epsilon_{gt} + a_h \epsilon_{rt} + e_{rt+h}. \tag{A15}$$

The coefficients v_h and a_h converge to

$$E[\hat{a}_h] = \frac{cov[R_{t+h}, \epsilon_{rt}]}{v[\epsilon_{rt}]} = \sum_{m=1}^{M} \rho_m \frac{cov[R_{t+h-m}, \epsilon_{rt}]}{v[\epsilon_{rt}]} + \frac{cov[u_{rt+h}, \epsilon_{rt}]}{v[\epsilon_{rt}]}, \tag{A16}$$

$$E[\hat{v}_h] = \frac{cov[R_{t+h}, \epsilon_{gt}]}{v[\epsilon_{gt}]} = \sum_{n=0}^{N} \beta_{rg,n} \frac{cov[R_{t+h-n}, \epsilon_{gt}]}{v[\epsilon_{gt}]}, \tag{A17}$$

for all $m \in M$ such that $h - m \geq 0$ and for all $n \in N$ such that $h - n \geq 0$. Following the notation in the body of the paper, I denote with $\boldsymbol{R}_{g,h}$ and $\boldsymbol{R}_{r,h}$ the impulse response of R to each innovation at horizon h.

The coefficients $\hat{c}_h$ from the cross-sectional step identify the differential response of the unit level outcome to a unit increase in ϵ_{rt}. This ϵ_{rt} shock implies a path for R given by $\hat{a}_h$ for each h. However, to construct the HEDGE term we need the path for R given by the coefficients $\hat{v}_h$. Following Sims and Zha (2006), I outline how to replicate $\boldsymbol{R}_g$ using a series of contemporaneous innovations to R dated $s \geq 0$ for all $s \in H$.

For $h = 0$, we want to find the size of the ϵ_{r0} shock that equalizes both paths on impact:

$$\boldsymbol{R}_{g,0} = \boldsymbol{R}_{r,0}\tilde{\epsilon}_{r0} \qquad \Rightarrow \qquad \tilde{\epsilon}_{r0} = \frac{\boldsymbol{R}_{g,0}}{\boldsymbol{R}_{r,0}}. \tag{A18}$$

An R shock of size $\tilde{\epsilon}_{r0}$ yields an interest rate response equal to $\boldsymbol{R}_{g,0}$. Next, we can evaluate $E[\hat{c}_0]$ for an innovation to R of size $\tilde{\epsilon}_{r0}$, instead of unity:

$$\begin{aligned} E[\hat{c}_0 \times \hat{\tilde{\epsilon}}_{r0}] &= \tilde{\gamma}_0 \boldsymbol{R}_{r,0}\tilde{\epsilon}_{r0} \\ &= \tilde{\gamma}_0 \boldsymbol{R}_{g,0} \\ &= \Omega_0. \end{aligned} \tag{A19}$$

The differential response of Y_{it} to an innovation to R of size $\tilde{\epsilon}_{r0}$ equals the response that operates through the endogenous response of R to a G shock. In other words, the HEDGE term for $h = 0$.

For $h = 1$, it will be the case that, unless $\boldsymbol{R}_g$ and $\boldsymbol{R}_r$ are multiples of each other:

$$\boldsymbol{R}_{g,1} \neq \boldsymbol{R}_{r,1}\tilde{\epsilon}_{r0}. \tag{A20}$$

This means that the initial rescaling by $\tilde{\epsilon}_{r0}$ is not enough to equalize the conditional paths of R at horizons $h > 0$. The next step is to feed in a contemporaneous innovation at date $h = 1$ that accounts for the remaining difference:

$$\tilde{\epsilon}_{r1} = \boldsymbol{R}_{g,1} - \boldsymbol{R}_{r,1}\tilde{\epsilon}_{r0} \tag{A21}$$

$$= \boldsymbol{R}_{g,1} - \boldsymbol{R}_{r,1}\frac{\boldsymbol{R}_{g,0}}{\boldsymbol{R}_{r,0}}, \tag{A22}$$

where $\tilde{\epsilon}_{r1}$ is an innovation to R taking place at $t+1$. This gives us two R shocks—$\{\tilde{\epsilon}_{r0}, \tilde{\epsilon}_{r1}\}$—that propagate to the cross-sectional outcome as follows

$$
\begin{aligned}
E[\hat{c}_1 \times \hat{\tilde{\epsilon}}_{r0} + \hat{c}_0 \times \hat{\tilde{\epsilon}}_{r1}] &= \left[\tilde{\gamma}_0 \boldsymbol{R}_{r,1} + \tilde{\gamma}_1 \boldsymbol{R}_{r,0}\right] \boldsymbol{R}_{g,0} + \tilde{\gamma}_0 \boldsymbol{R}_{r,0} \left[\boldsymbol{R}_{g,1} - \boldsymbol{R}_{r,1} \boldsymbol{R}_{g,0}\right] \\
&= \tilde{\gamma}_0 \boldsymbol{R}_{r,1} \boldsymbol{R}_{g,0} + \tilde{\gamma}_1 \boldsymbol{R}_{g,0} + \tilde{\gamma}_0 \boldsymbol{R}_{g,1} - \tilde{\gamma}_0 \boldsymbol{R}_{r,1} \boldsymbol{R}_{g0} \\
&= \tilde{\gamma}_1 \boldsymbol{R}_{g,0} + \tilde{\gamma}_0 \boldsymbol{R}_{g,1} \\
&= \Omega_1.
\end{aligned} \tag{A23}
$$

In period $h = 1$, it is *as if* the economy was responding to a shock of size $\tilde{\epsilon}_{r0}$ that took place a period ago plus to a shock of size $\tilde{\epsilon}_{r1}$ taking place contemporaneously. The combination of these responses identifies Ω_1.

As before, for $h = 2$ we construct a contemporaneous shock at date $h = 2$ equivalent to the difference between $\boldsymbol{R}_{g,2}$ and the R path implied by the sequence of shocks fed in so far:

$$
\tilde{\epsilon}_{r2} = \boldsymbol{R}_{g,2} - \boldsymbol{R}_{r,2} \tilde{\epsilon}_{r0} - \boldsymbol{R}_{r,1} \tilde{\epsilon}_{r1}, \tag{A24}
$$

where $\boldsymbol{R}_{r,2} \tilde{\epsilon}_{r0}$ is the response of R at $h = 2$ to the shock that took place at $t = 0$ and $\boldsymbol{R}_{r,1} \tilde{\epsilon}_{r1}$ is the response of R at $h = 2$ to the shock that took place in the previous period. Note that because $R_{r,0} = 1$ I omit it from the left-hand side. Once we have $\tilde{\epsilon}_{r2}$ we can construct an estimate for Ω_2 as:

$$
E[\hat{c}_2 \times \hat{\tilde{\epsilon}}_{r0} + \hat{c}_1 \times \hat{\tilde{\epsilon}}_{r1} + \hat{c}_0 \times \hat{\tilde{\epsilon}}_{r2}] = \Omega_2. \tag{A25}
$$

The two formulas below generalize the computation of the sequence of shocks $\tilde{\epsilon}_{rh}$ and the estimation of Ω_h for any $h \geq 0$.

$$
\hat{\nu}_h = \sum_{k=0}^{h} \hat{a}_{h-k} \hat{\tilde{\epsilon}}_{rk} \quad \forall\, h \in H, \tag{A26}
$$

$$
\widehat{\Omega}_h = \sum_{k=0}^{h} \hat{c}_{h-k} \hat{\tilde{\epsilon}}_{rk} \quad \forall\, h \in H. \tag{A27}
$$

A.3. Dynamic Setting without Anticipation Effects - Further Details

A.3.1. Control Function Approach in Dynamic Settings

Back to reference in main text.

This Appendix presents detailed derivations for the control function estimator.

The reason why the control function approach fails to control for the dynamic path of the GE variable in response to a shock of interest ϵ_{gt} is independent of whether the setting is a panel or a time-series regression. Given this, in this Appendix I provide details on the mechanics of the control function approach using an aggregate setting.

Throughout, I use the Frisch–Waugh–Lovell (FWL) theorem to characterize the expected value of the elasticities estimated under the control function approach. The FWL theorem implies:

$$\hat{b} = \frac{cov[y^{\perp z}, x^{\perp z}]}{v[x^{\perp z}]}, \tag{A28}$$

where:

$$x^{\perp z} = x - \theta z, \qquad \theta = \frac{cov[x, z]}{v[z]},$$

$$y^{\perp z} = y - \theta_y z, \qquad \theta_y = \frac{cov[y, z]}{v[z]}.$$

AR(1) Process for G and R. First, I consider the following DGP:

$$G_t = \rho_g G_{t-1} + u_{gt} \tag{A29}$$

$$R_t = \rho_r R_{t-1} + \delta G_t + u_{rt} \tag{A30}$$

$$Y_t = \beta G_t + \gamma R_t + u_{yt} \tag{A31}$$

where I assume the researcher observes an innovation to G, ϵ_{gt}, that satisfies Assumption (3). The control function approach can be implemented with the following local projection:

$$Y_{t+h} = b_h G_t + c_h R_t + e_{t+h} \tag{A32}$$

where G_t is instrumented with ϵ_{gt}. The FWL theorem implies the following expression for the 2SLS estimate $\hat{b}_h$:

$$E[\hat{b}_h] = \frac{cov[Y_{t+h}^{\perp R_t}, \epsilon_{gt}^{\perp R_t}]}{cov[G_t^{\perp R_t}, \epsilon_{gt}^{\perp R_t}]}. \tag{A33}$$

Then expanding and re-arranging the numerator yields:
(A34)

$$cov[Y_{t+h}^{\perp R_t}, \epsilon_{gt}^{\perp R_t}] = \beta\rho_g^h\Big(cov[G_t, \epsilon_{gt}] - \frac{cov[G_t, R_t]}{v[R_t]}cov[R_t, \epsilon_{gt}]\Big) + \gamma\Big(cov[R_{t+h}, \epsilon_{gt}] - \frac{cov[R_{t+h}, R_t]}{v[R_t]}cov[R_t, \epsilon_{gt}]\Big)$$
$$= \beta\rho_g^h\Big(v[\epsilon_{gt}] - \delta\frac{cov[G_t, R_t]}{v[R_t]}v[\epsilon_{gt}]\Big) + \gamma\Big(cov[R_{t+h}, \epsilon_{gt}] - \frac{cov[R_{t+h}, R_t]}{v[R_t]}cov[R_t, \epsilon_{gt}]\Big),$$

where I used the fact that $cov[R_t, \epsilon_{gt}] = \delta v[\epsilon_{gt}]$. Similarly, the denominator can be expressed as:

$$cov[G_t^{\perp R_t}, \epsilon_{gt}^{\perp R_t}] = v[\epsilon_{gt}] - \delta\frac{cov[G_t, R_t]}{v[R_t]}v[\epsilon_{gt}]. \tag{A35}$$

Putting both pieces together, yields:

$$E[\hat{b}_h] = \beta\rho_g^h + \frac{\gamma v[\epsilon_{gt}^{\perp R_t}]}{cov[G_t^{\perp R_t}, \epsilon_{gt}^{\perp R_t}]}\Big(\boldsymbol{R}_{g,h} - \boldsymbol{R}_{r,h} \times \boldsymbol{R}_{\tilde{R},h}\Big) \tag{A36}$$

where:

$$\boldsymbol{R}_{\tilde{R},h} \equiv \frac{cov[R_{t+h}, R_t]}{v[R_t]}. \tag{A37}$$

The term $\boldsymbol{R}_{\tilde{R},h}$ captures the component of the dynamic path of R that is predictable from the included control R_t. More generally, when additional controls are included, $\boldsymbol{R}_{\tilde{R},h}$ captures the projection of R_{t+h} onto the space spanned by those controls (e.g., current and lagged values of R and other aggregate variables).

This expression highlights that unless $\boldsymbol{R}_{g,h}$ can be represented as a linear combination of the paths of R spanned by included controls, the CFA fails to address the HEDGE bias.

A.3.2. Dynamic Setting - Monte Carlo Simulations

Back to reference in main text.
In the following, I detail the DGPs used to produce Table A3 in the main body of the paper.

For reference, I reproduce the general system of equations:

(S2)

$$\begin{bmatrix} G_t \\ R_t \end{bmatrix} = \sum_{j=1}^{P} M_j \begin{bmatrix} G_{t-j} \\ R_{t-j} \end{bmatrix} + B \begin{bmatrix} u_{gt} \\ u_{rt} \end{bmatrix},$$

$$Y_{it+h} = \sum_{s=0}^{S} \beta_s s_{ig} G_{t+h-s} + \sum_{k=0}^{K} \gamma_k s_{ir} R_{t+h-k} + \sum_{l=1}^{L} \psi_l Y_{i,t+h-l} + u_{ih} + u_{th} + u^y_{it+h} \quad \forall h \geq 0.$$

Each DGP is evaluated at three different values for the correlation between the exposure shifters s_{ig} and s_{ir}. I fix the standard deviation of all idiosyncratic shocks to unity. The sample size is set to $N = 300$, $T = 1000$, and the number of repetitions is set to $J = 1000$.

TABLE A1. Monte Carlo DGPs — Dynamic Setting

DGP #	**Description**	m_{gg}	m_{rr}	m_{gr}	m_{rg}	β_s	γ_k	ψ_l
1	VAR(1), no lagged Y_i	0.8	0.8	0.5	0.0	$\{1.0\}$	$\{0.5\}$	$\{0.0\}$
2	VAR(1), AR(1) in Y_i	0.8	0.8	0.5	0.0	$\{1.0\}$	$\{0.5\}$	$\{0.8\}$
3	VAR(2), AR(1) in Y_i	$\{0.7, 0.2\}$	$\{0.4, 0.5\}$	$\{0.5, 0.0\}$	$\{0.0, 0.0\}$	$\{1.0\}$	$\{0.5\}$	$\{0.8\}$
4	VAR(1), lags in G and R in Y_i	0.7	0.4	0.5	0.0	$\{0.5, 0.2\}$	$\{0.5, 0.05\}$	$\{0.0\}$
5	VAR(1), unrestricted	0.8	0.8	0.5	−0.3	$\{1.0\}$	$\{0.5\}$	$\{0.0\}$

Each DGP is evaluated at $\phi \in \{0.1, 0.5, 1\}$. Parameters m_{xy} denote the effect of variable y on variable x in the VAR at different horizons. The β_s, γ_k, and ψ_l parameters correspond to the effects of G_{t-s}, R_{t-k}, and Y_{it-l} in the unit level outcome. Zero entries indicate either exclusion or no persistence.

A.3.3. Two-Way GE Feedback - Two Sources of Omitted Variable Bias

Back to reference in main text.

This appendix studies the cross-sectional identification of partial-equilibrium elasticities (PE) when there is two-way feedback (2W-GE) between aggregate variables. It shows that with generalized feedback between aggregate variables cross-sectional elasticities are generally contaminated by GE responses even when the covariance between exposure to the shock of interest ϵ_{gt} and exposure s_{ik} to some other aggregate variable K_t is zero.

The general setting is still characterized by (S3) - which I reproduce below for convenience:

$$\begin{bmatrix} G_t \\ R_t \end{bmatrix} = \sum_{j=1}^{P} M_j \begin{bmatrix} G_{t-j} \\ R_{t-j} \end{bmatrix} + B \begin{bmatrix} u_{gt} \\ u_{rt} \end{bmatrix},$$

(S2)

$$Y_{it} = \sum_{s=0}^{S} \beta_s s_{ig} G_{t-s} + \sum_{k=0}^{K} \gamma_k s_{ir} R_{t-k} + \sum_{l=1}^{L} \psi_l Y_{i,t-l} + u_i + u_t + u_{it}^{y}$$
$$s_{ir} = \phi s_{ig} + u_{is_r}.$$

Necessary Conditions to Difference-out GE using reduced form analysis. I study the necessary conditions for a reduced-form TWFE local projection to identify the partial equilibrium effect of a change in G_t on Y_{it+h}. Concretely, the regression under consideration is given by:

$$Y_{it+h} = b_h^{RF} s_{ig} \epsilon_{gt} + \lambda_{ih} + \lambda_{th} + e_{it+h}, \tag{A38}$$

where ϵ_{gt} is an observed innovation to G in period t. Let b_h^{RF} be the coefficient of interest, where the superscript RF refers to reduced form. To build intuition, I start with the following simplified DGP for G and R:

$$G_t = \rho_g G_{t-1} + \alpha R_t + u_{gt}, \qquad R_t = \rho_r R_{t-1} + \delta G_t + u_{rt} \tag{A39}$$

The VAR(1) representation can be written as

$$G_t = \frac{\rho_g}{1-\alpha\delta} G_{t-1} + \frac{\alpha\rho_r}{1-\alpha\delta} R_{t-1} + \frac{\alpha}{1-\alpha\delta} u_{rt} + \frac{1}{1-\alpha\delta} u_{gt} \tag{A40}$$

$$R_t = \frac{\rho_r}{1-\alpha\delta} R_{t-1} + \frac{\delta\rho_g}{1-\alpha\delta} G_{t-1} + \frac{1}{1-\alpha\delta} u_{rt} + \frac{\delta}{1-\alpha\delta} u_{gt}. \tag{A41}$$

Then, consider a simple DGP for Y_{it} given by:

$$Y_{it} = \beta s_{ig} G_t + \gamma s_{ir} R_t + u_{it}, \qquad s_{ir} = \phi s_{ig} + e_i^{s_r}. \tag{A42}$$

It can be shown that the coefficient of interest converges to the following expression for $h = 0$ and $h = 1$, respectively:

$$\hat{b}_0^{RF} \to \frac{\beta}{1-\alpha\delta} + \frac{\phi\gamma\delta}{1-\alpha\delta}, \qquad \hat{b}_1^{RF} \to \frac{\beta(\rho_g + \alpha\delta\rho_r)}{(1-\alpha\delta)^2} + \frac{\phi\gamma\delta(\rho_g + \rho_r)}{(1-\alpha\delta)^2}. \tag{A43}$$

Now, consider an economy in which exposure shifters to G and R are uncorrelated in the cross-section - that is, $\phi = 0$. This leaves us with the following.

$$\hat{b}_0 \to \frac{\beta}{1-\alpha\delta} \quad \neq \beta, \qquad \hat{b}_1 \to \frac{\beta(\rho_g + \alpha\delta\rho_r)}{(1-\alpha\delta)^2} \quad \neq \beta\rho_g. \tag{A44}$$

Both reduced form coefficients are contaminated by GE effects *even* in absence of HEDGE. A necessary condition to *fully* difference-out GE using a reduced-form cross-sectional regression is that there should be no two-way feedback between the aggregate variable of interest and any variable that responds to it in general equilibrium. Otherwise, the reduced-form coefficients identify a rescaled version of the true partial effect even when $cov[s_{ig}, s_{ir}) = 0$.

Is 2SLS the solution?. Suppose that instead of using the reduced form, we chose to instrument G_t with ϵ_{gt}. The first stage coefficient is given by:

$$\theta^{FS} = \frac{1}{1-\alpha\delta}. \tag{A45}$$

The 2SLS estimates converge to:

$$\hat{b}_0^{2S} \to \beta \quad = \beta, \qquad \hat{b}_1^{2S} \to \frac{\beta(\rho_g + \alpha\delta\rho_r)}{(1-\alpha\delta)} \quad \neq \beta\rho_g. \tag{A46}$$

For this DGP, using ϵ_{gt} as an instrument for G_t removes GE on impact but not at $h = 1$ (nor at any $h > 0$). What if instead we instrument G_{t+h} with ϵ_{gt}? In this case, the first stage coefficient is horizon-specific, and for $h = 1$, we get:

$$\theta_1^{FS} = \frac{\rho_g + \alpha\delta\rho_r}{(1-\alpha\delta)^2}. \tag{A47}$$

The first stage coefficient for $h = 0$ is the same as in the previous case. The 2SLS estimates converge to:

$$\hat{b}_0^{2S,H} \to \beta, \qquad \hat{b}_1^{2S,H} \to \beta, \tag{A48}$$

where I use the subscript H to denote that these estimates correspond to a 2SLS regression where G_{t+h} is being instrumented for, instead of G_t. In the context of this specific DGP, instrumenting G_{t+h} identifies the impact effect of a change in G and the local projection recovers the same coefficient at all horizons. Note that this does not identify the cumulative impulse response but does identify an object that is free of GE. Crucially, this worked because, for this DGP, G_{t+h} captures, or blocks, all paths between ϵ_{gt} and Y_{it+h}. Consider departing from this scenario by adding an autoregressive component in the unit-level outcome:

$$Y_{it} = \beta s_{ig} G_t + \psi Y_{it-1} + u_{it}^y, \tag{A49}$$

or, for lagged G to have an effect on the unit-level outcome:

$$Y_{it} = \beta s_{ig} G_t + \beta_1 s_{ig} G_{t-1} + u_{it}^y. \tag{A50}$$

Now, G_{t+h} no longer blocks all channels through which a change in ϵ_{gt} affects Y_{it+h}. Therefore, instrumenting G_{t+h} with ϵ_{gt} will not be enough to difference-out GE effects. The 2SLS estimates for these two alternative DGPs at $h = 1$ are:

$$\hat{b}_1^{2S} = \beta \frac{(\rho_g + \delta\alpha\rho_r)}{1 - \alpha\delta} + \beta\psi \neq \beta(\rho_g + \psi), \qquad \hat{b}_1^{2S} = \beta \frac{(\rho_g + \delta\alpha\rho_r)}{1 - \alpha\delta} + \beta_1 \neq \beta\rho_g + \beta_1, \tag{A51}$$

when instrumenting G_t, and

$$\hat{b}_1^{2S,H} = \beta + \beta\psi \frac{1 - \alpha\delta}{(\rho_g + \delta\alpha\rho_r)} \neq \beta(\rho_g + \psi), \quad \hat{b}_1^{2S,H} = \beta + \beta_1 \frac{1 - \alpha\delta}{(\rho_g + \delta\alpha\rho_r)} \neq \beta\rho_g + \beta_1, \tag{A52}$$

when instrumenting G_{t+h}. The expressions to the right of the $\neq$ sign correspond to the PE elasticities for each DGP and horizon.

The Decomposition Framework also addresses this OVB. Using the reduced-form responses estimated from:

$$Y_{it+h} = b_h \times s_{ig}\epsilon_{gt} + c_h \times s_{ig}\epsilon_{rt} +_{ih} + \lambda_{th} + e_{it+h} \tag{A53}$$

$$R_{t+h} = \nu_h \epsilon_{gt} + a_h \epsilon_{rt} + e_{rt+h}, \tag{A54}$$

one can implement the decomposition framework as outlined in 3.2 and consistently estimate the partial equilibrium elasticity at all horizons, even when there is two-way feedback between aggregates in a system like $S2$. Next, I present results from a Monte Carlo simulation exercise and the algebraic details for the first two horizons in a simplified DGP.

HEDGE Test. In environments with 2W-GE feedback, the interpretation of the HEDGE test requires additional care. When G_t responds endogenously to R_t, variation in ϵ_{rt} induces changes in both aggregate variables. As a result, even in the absence of comovement between exposure shifters (i.e., $\phi = 0$), the interaction $s_{ig}\epsilon_{rt}$ will be correlated with $s_{ig}G_t$.

In this case, a rejection of the null should not be interpreted as evidence of heterogeneous exposure to R per se (i.e., $\phi \neq 0$), but more generally as evidence of heterogeneous exposure to equilibrium responses of the economy. In particular, units that are more exposed to G will also be differentially exposed to the endogenous response of G to R. Therefore, the HEDGE test should be interpreted as a diagnostic for residual cross-sectional heterogeneity in exposure to aggregate responses, rather than a test exclusively of $\phi \neq 0$.

Monte Carlo Results. For illustration, I ran a Monte Carlo exercise in which I simulate data from alternative DGPs that fit into the system given by (S3) to estimate b_h. Throughout, I set $\phi = 0$ to focus on the omitted variable bias that arises independently of HEDGE. The sample size is set to $N = 300$ and $T = 500$, the number of repetitions to $J = 1000$, and the standard deviation of all idiosyncratic shocks to unity. Table A2 provides further details on each of the DGPs used in the Monte Carlo exercise. I consider four estimation strategies: (1) a TWFE reduced-form local projection, (2) the decomposition framework, (3) a TWFE local projection instrumenting G_t with ϵ_{gt}, and (4) a TWFE local projection instrumenting G_{t+h} with ϵ_{gt}. For each estimation strategy k, DGP z and horizon h, I compute the absolute difference between the true portable elasticity, β^P_{zh}, and the average estimate across all repetitions and then normalize it by β^P_{zh} to make it comparable across data-generating processes:

$$\theta^k_{zh} = \frac{|\, J^{-1}\sum_{j=1}^{J} \hat{b}^k_{jzh} - \beta^P_{zh} \,|}{|\, \beta^P_{zh} \,|}. \tag{A55}$$

Consistency requires $\theta^k_{zh} \to 0$. Table A3 shows the mean and standard deviation of θ^k_{zh} by horizon and estimation method. The decomposition approach outperforms all three estimation strategies at every horizon and yields unbiased estimates of the true portable elasticities, with only minor bias at longer horizons.

TABLE A2. Two-Way GE - Parameterizations of DGPs in Monte Carlo simulations

DGP #	Description	m_{gg}	m_{rr}	m_{gr}	m_{rg}	β_s	γ_k	ψ_l	ϕ
1	AR(1) in G and R, no lagged Y.	0.6	0.6	−0.3	0.5	$\{1.0\}$	$\{0.5\}$	$\{0\}$	0
2	Adds persistence in Y_{it} with one autoregressive lag.	0.6	0.6	0.1	0.5	$\{1.0\}$	$\{0.5\}$	$\{0.8\}$	0
3	Adds distributed lag response of Y to G via multiple β_s.	0.6	0.6	0.1	0.5	$\{1.0, 0.4, 0.2\}$	$\{0.5\}$	$\{0\}$	0
4	Combines lagged Y and lagged G.	0.6	0.6	0.1	0.5	$\{1.0, 0.4, 0.2\}$	$\{0.5\}$	$\{0.8\}$	0
5	Adds three autoregressive lags of Y.	0.6	0.6	0.1	0.5	$\{1.0\}$	$\{0.5\}$	$\{0.5, 0.3, -0.1\}$	0

Parameters m_{xy} denote the effect of variable y on variable x in the VAR at different horizons. The β_s, γ_k, and ψ_l parameters correspond to the effects of G_{t_s}, R_{t-k}, and lags of Y_{it-l} in the unit level outcome. Zero entries indicate either exclusion or no persistence.
Back to reference in main text.

TABLE A3. Absolute Mean Relative Bias by Estimation Method and Horizon

Estimation Strategy	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 5$
Decomposition Framework	0.002	0.005	0.009	0.015	0.019	0.027
	(0.001)	(0.002)	(0.006)	(0.010)	(0.010)	(0.016)
Reduced Form	0.002	0.078	0.211	0.391	0.587	0.788
	(0.002)	(0.098)	(0.273)	(0.469)	(0.636)	(0.744)
IV (t)	0.001	0.079	0.213	0.391	0.587	0.787
	(0.001)	(0.097)	(0.273)	(0.469)	(0.636)	(0.743)
IV (H)	0.001	0.624	1.556	3.494	5.159	8.680
	(0.001)	(0.026)	(0.191)	(1.662)	(1.123)	(2.353)

Standard errors shown in parentheses. Values are reported as fractions of the true portable elasticity (1 = 100 %). The sample size is set to $N = 300$ and $T = 500$ and the number of repetitions to $J = 1000$.
Back to reference in main text.

Decomposition Framework with 2W-GE - Detailed Algebra for $h = 0, 1$. The DGP is the same as that used before, which assumes a VAR(1) process for aggregate variables and only contemporaneous G and R to directly affect the local outcome.

$$
\begin{aligned}
G_t &= \rho_g G_{t-1} + \alpha R_t + u_{gt}, \\
R_t &= \rho_r R_{t-1} + \delta G_t + u_{rt}, \\
Y_{it} &= \beta s_{ig} G_t + \gamma s_{ir} R_t + u^y_{it}, \\
s_{ir} &= \phi s_{ig} + u^{Sr}_i, \quad s_{ig} \perp u^{Sr}_i.
\end{aligned} \tag{A56}
$$

Cross-sectional Estimates. The estimated coefficients of the cross-sectional step for $h = 0, 1$ converge to:

$$
E[\hat{b}_0] = \frac{\beta}{1-\alpha\delta} + \frac{\tilde{\gamma}\delta}{1-\alpha\delta}, \qquad E[\hat{b}_1] = \frac{\beta(\rho_g + \alpha\delta\rho_r)}{(1-\alpha\delta)^2} + \frac{\tilde{\gamma}\delta(\rho_g + \rho_r)}{(1-\alpha\delta)^2}, \tag{A57}
$$

$$
E[\hat{c}_0] = \frac{\beta\alpha}{1-\alpha\delta} + \frac{\tilde{\gamma}}{1-\alpha\delta}, \qquad E[\hat{c}_1] = \frac{\beta\alpha(\rho_g + \rho_r)}{(1-\alpha\delta)^2} + \frac{\tilde{\gamma}(\rho_r + \delta\alpha\rho_g)}{(1-\alpha\delta)^2}. \tag{A58}
$$

Time Series Estimates. The estimated coefficients from the time series step for $h = 0, 1$ converge to:

$$E[\hat{a}_0] = \frac{1}{1-\alpha\delta}, \qquad E[\hat{a}_1] = \frac{\rho_r + \delta\alpha\rho_g}{(1-\alpha\delta)^2}, \tag{A59}$$

$$E[\hat{\nu}_0] = \frac{\delta}{1-\alpha\delta}, \qquad E[\hat{\nu}_1] = \frac{\delta(\rho_r + \rho_g)}{(1-\alpha\delta)^2}. \tag{A60}$$

Decomposition. The counterfactual shock for $h = 0$ satisfies the following:

$$\nu_0 = a_0\tilde{\epsilon}_{r0} \rightarrow \frac{\delta}{1-\alpha\delta} = \frac{1}{1-\alpha\delta}\tilde{\epsilon}_{r0} \tag{A61}$$

$$\delta = \tilde{\epsilon}_{r0}. \tag{A62}$$

Next, we can evaluate the differential response to R at $\tilde{\epsilon}_{r0}$ and subtract it from the differential response to G:

$$\begin{aligned} E[\hat{b}_0 - \hat{c}_0 \times \hat{\tilde{\epsilon}}_{r0}] &= \frac{\beta}{1-\alpha\delta} + \frac{\tilde{\gamma}\delta}{1-\alpha\delta} - \left[\frac{\beta\alpha}{1-\alpha\delta} + \frac{\tilde{\gamma}}{1-\alpha\delta}\right] \times \delta \\ &= \frac{\beta(1-\alpha\delta)}{1-\alpha\delta} + \frac{\tilde{\gamma}\delta}{1-\alpha\delta} - \frac{\tilde{\gamma}\delta}{1-\alpha\delta} \\ &= \beta. \end{aligned}$$

In period $h = 1$, this shock $\tilde{\epsilon}_{r0} = \delta$ implies the following realization for R_{t+1}:

$$\boldsymbol{R}_{r,1} = \frac{\rho_r + \delta\alpha\rho_g}{(1-\alpha\delta)^2}\delta \quad \neq \quad \boldsymbol{R}_{g,1} = \frac{\delta(\rho_r + \rho_g)}{(1-\alpha\delta)^2}. \tag{A63}$$

We want to find a second date $h = 1$ shock $\tilde{\epsilon}_{r1}$ such that:

$$\begin{aligned} E[\hat{\nu}1] &= E[\hat{a}_1 \times \hat{\tilde{\epsilon}}_{r0} + \hat{a}_0] \times \hat{\tilde{\epsilon}}_{r1}] \\ \frac{\delta(\rho_r + \rho_g)}{(1-\alpha\delta)^2} &= \frac{\rho_r + \delta\alpha\rho_g}{(1-\alpha\delta)^2}\delta + \frac{1}{1-\alpha\delta} \times E[\hat{\tilde{\epsilon}}_{r1}] \\ &\rightarrow E[\hat{\tilde{\epsilon}}_{r1}] = \delta\rho_g. \end{aligned} \tag{A64}$$

Next, we evaluate the differential response of Y_i to these two R shocks, $\{\tilde{\epsilon}_{r0}, \tilde{\epsilon}_{r1}\}$:

$$E[\hat{c}_1 \times \hat{\tilde{\epsilon}}_{r0} + \hat{c}_0 \times \hat{\tilde{\epsilon}}_{r1}], \tag{A65}$$

and subtract it from the differential response of the cross-sectional outcome to a G shock in period $t+1$:

$$\begin{aligned}
E[\hat{b}_1 - \hat{c}_1 \times \hat{\hat{\epsilon}}_{r0} - \hat{c}_0 \times \hat{\hat{\epsilon}}_{r1}] &= \left[\frac{\beta(\rho_g + \alpha\delta\rho_r)}{(1-\alpha\delta)^2} + \frac{\tilde{\gamma}\delta(\rho_g + \rho_r)}{(1-\alpha\delta)^2}\right] \\
&\quad - \left[\frac{\beta\alpha(\rho_g + \rho_r)}{(1-\alpha\delta)^2} + \frac{\tilde{\gamma}(\rho_r + \delta\alpha\rho_g)}{(1-\alpha\delta)^2}\right]\delta - \frac{\beta\alpha + \tilde{\gamma}}{1-\alpha\delta}\rho_g\delta \\
&= \beta\rho_g = \beta_1^P.
\end{aligned} \tag{A66}$$

This process can be extended to further horizons.

A.3.4. Robustness to $cov[\epsilon_{gt}, A_t] \neq 0$

Back to reference in main text.

This appendix lifts the assumption that the researcher observes a purely structural innovation to G_t. I consider settings in which ϵ_{gt}, the observed shock of interest, is correlated with the realization of a third aggregate variable A_t. This nests cases in which the shock is not a source of exogenous variation at the aggregate level but satisfies exogeneity in a cross-sectional setting, conditional on the exposure shifters and time fixed effects. For illustration purposes, consider the following minimal setting.

$$G_t = \rho_g G_{t-1} + u_{gt}, \tag{A67}$$

$$R_t = \rho_r R_{t-1} + \delta G_t + u_{rt}, \tag{A68}$$

$$A_t = \kappa_a G_t + u_{at}, \tag{A69}$$

$$Y_{it} = \beta s_{ig} G_t + \gamma s_{ir} R_t + \chi s_{ia} A_t + u_i + u_t + u_{it}^y, \tag{A70}$$

$$s_{ir} = \phi s_{ig} + u_i^{s_r}, \tag{A71}$$

$$cov[s_{ig}, s_{ia}] = 0. \tag{A72}$$

Here, A_t is a third aggregate variable that contemporaneously responds to changes in G_t and to its own structural innovation. In addition, units are differentially exposed to A according to the exposure shifter s_{ia} and the parameter χ. In a cross-sectional TWFE regression, the shock ϵ_{gt} satisfies exogeneity with respect to A_t if $cov[s_{ig}, s_{ia}) = 0$. That is, the cross section identifies a causal effect of ϵ_{gt} if there is no HEDGE in relation to A_t.

Next, I show that the decomposition framework consistently identifies the portable elasticity in settings that use variation correlated to a third aggregate variable. I consider DGPs that fit into the following system of equations:

$$\begin{bmatrix} G_t \\ R_t \\ A_t \end{bmatrix} = \sum_{j=1}^{P} M_j \begin{bmatrix} G_{t-j} \\ R_{t-j} \\ A_{t-j} \end{bmatrix} + B \begin{bmatrix} u_{gt} \\ u_{rt} \\ u_{at} \end{bmatrix},$$

$$\begin{aligned} Y_{it} = &\sum_{s=0}^{S} \beta_s s_{ig} G_{t-s} + \sum_{k=0}^{K} \gamma_k s_{ir} R_{t-k} + \sum_{w=0}^{W} \chi_w s_{ia} A_{t-w}, \\ &+ \sum_{l=1}^{L} \psi_l Y_{i,t-l} + u_i + u_t + u_{it}^y, \\ s_{ir} = \phi s_{ig} + u_i^{s_r}, &\qquad s_{ia} = \xi s_{ig} + u_i^{s_a}, \qquad s_{ig} \sim N(1, u_{s_g}). \end{aligned} \tag{S3}$$

For simulations, I study DGPs that (i) satisfy the identification assumption of the researcher - namely $cov[s_{ig}, s_{ia}] = 0$ ($\xi = 0$); and (ii) fail to satisfy it because they display HEDGE in A_t, on top of R_t. In both cases, the decomposition identifies its object of interest. That is, a sequence of elasticities that are clean of the effects operating through changes in R. The difference is that in the first case, the estimated elasticities are also independent of changes in A_t, whereas in the second case they are a combination of the true portable elasticity and the HEDGE channel associated with A_t. Table A4 provides details on each of the DGPs used in the Monte Carlo exercise. In all cases, the sample size is set to $N = 100$ and $T = 500$, the number of repetitions to $J = 500$, and the standard deviation of all idiosyncratic shocks to unity. For each estimation strategy k, DGP z and horizon h, I compute the absolute difference between the true portable elasticity, β^P_{zh} and the average estimate across all repetitions and then normalize it by β^P_{zh} to make it comparable across data-generating processes:

$$\theta^k_{zh} = \frac{| J^{-1} \sum_{j=1}^{J} \hat{b}^k_{jzh} - \beta^P_{zh} |}{| \beta^P_{zh} |}. \tag{A73}$$

Consistency requires $\theta^k_{zh} \to 0$. Table A5 shows the mean and standard deviation of θ^k_{zh} by horizon and estimation method. The decomposition approach consistently recovers the true portable effects as opposed to the control function approach.

TABLE A4. Three-Way GE — Parameterizations of DGPs in Monte Carlo simulations

DGP - Description	m_{gg}	m_{gr}	m_{ga}	m_{rr}	m_{rg}	m_{ra}	m_{aa}	m_{ag}	m_{ar}	β_s	γ_k	χ_w	ψ_l	ξ
A is i.i.d..	0.9	$\{0, 0\}$	$\{0, 0\}$	0.7	$\{0.5, 0\}$	$\{0.3, 0\}$	0.0	$\{0.2\}$	$\{0, 0\}$	$\{1.0\}$	$\{0.5\}$	$\{0.2\}$	0.8	0
Adds persistence in A_t.	0.9	$\{0, 0\}$	$\{0, 0\}$	0.7	$\{0.5, 0\}$	$\{0.3, 0\}$	0.4	$\{0.2\}$	$\{0, 0\}$	$\{1.0\}$	$\{0.5\}$	$\{0.2\}$	0.8	0
Adds $R \to A$ channel via m_{ar}.	0.9	$\{0, 0\}$	$\{0, 0\}$	0.7	$\{0.5, 0\}$	$\{0.3, 0\}$	0.4	$\{0.2\}$	$\{0.4, 0\}$	$\{1.0\}$	$\{0.5\}$	$\{0.2\}$	0.8	0
Adds $A \to G$ channel via m_{ga}.	0.8	$\{0, 0\}$	$\{0.1, 0\}$	0.7	$\{0.5, 0\}$	$\{0.3, 0\}$	0.4	$\{0.2\}$	$\{0.2, 0\}$	$\{1.0\}$	$\{0.5\}$	$\{0.2\}$	0.8	0
3W-GE.	0.9	$\{-0.3, 0\}$	$\{0.1, 0\}$	0.7	$\{0.5, 0\}$	$\{0.3, 0\}$	0.4	$\{0.2\}$	$\{0.2, 0\}$	$\{1.0\}$	$\{0.5\}$	$\{0.2\}$	0.8	0
Adds lags of G, R, and A in Y_{it}.	0.9	$\{-0.3, 0\}$	$\{0.1, 0\}$	0.7	$\{0.5, 0\}$	$\{0.3, 0\}$	0.4	$\{0.2\}$	$\{0.2, 0\}$	$\{1.0, 0.2\}$	$\{0.5, 0.2\}$	$\{0.2, 0.3\}$	0.8	0
Adds HEDGE-A via $\xi > 0$.	0.9	$\{-0.3, 0\}$	$\{0.1, 0\}$	0.7	$\{0.5, 0\}$	$\{0.3, 0\}$	0.4	$\{0.2\}$	$\{0.2, 0\}$	$\{1.0\}$	$\{0.5\}$	$\{0.2\}$	0.8	0.5
Combines lagged inputs with HEDGE-A.	0.9	$\{-0.3, 0\}$	$\{0.1, 0\}$	0.7	$\{0.5, 0\}$	$\{0.3, 0\}$	0.4	$\{0.2\}$	$\{0.2, 0\}$	$\{1.0, 0.2\}$	$\{0.5, 0.2\}$	$\{0.2, 0.3\}$	0.8	0.5

Each DGP is evaluated at $\phi \in \{0.1, 0.5, 1\}$. Parameters m_{xy} denote the effect of variable y on variable x in the VAR at different horizons. The β_s, γ_k, χ_w and ψ_l parameters correspond to the effects of G_{t-s}, R_{t-k}, A_{t-j} and Y_{it-l} in the unit level outcome. Zero entries indicate either exclusion or no persistence.

TABLE A5. DGPs with $cov[A_t, \epsilon_{gt}] \neq 0$ - Absolute Mean Relative Bias by Estimation Method

Estimation Method	$h = 0$	$h = 1$	$h = 2$	$h = 3$	$h = 4$	$h = 5$	$h = 6$
Decomposition	0.004	0.005	0.007	0.009	0.010	0.011	0.012
	(0.003)	(0.003)	(0.005)	(0.006)	(0.007)	(0.008)	(0.010)
Control Function	1.266	0.560	0.331	0.236	0.246	0.343	0.453
	(0.414)	(0.167)	(0.073)	(0.129)	(0.189)	(0.202)	(0.210)

Standard errors in parenthesis. Values are reported as fractions of the true portable elasticity (1 = 100 %).
Based on $J = 500$ repetitions with $N = 100$ and $T = 500$.
Back to reference in main text.

A.4. Comparison with Other Estimation Approaches

Back to reference in main text.

This appendix discusses how the decomposition approach outlined in Section 3.2 compares to three alternative approaches: (i) Interactive fixed effects (Bai 2009), (ii) control function approach assuming s_{ir} is observed, (iii) 2SLS control function approach.

A.4.1. Interactive Fixed Effects

The cross-sectional outcome Y_{it} evolves according to

$$Y_{it} = \beta \times s_{ig} G_t + \gamma s_{ir} R_t + u_i + u_t + u^y_{it} \tag{A74}$$

The Interactive Fixed Effects (IFE) estimator (Bai 2009) augments the TWFE regression with L latent factors to absorb unobserved heterogeneity:

$$Y_{it+h} = \beta^{\text{IFE}}_h \times s_{ig} \epsilon_{gt} + \sum_{\ell=1}^{r} \lambda^{\ell}_i f^{\ell}_t + \lambda_{ih} + \lambda_{th} + e_{it+h}, \tag{A75}$$

where $\{\lambda^{\ell}_i, f^{\ell}_t\}$ are estimated jointly with β^{IFE}_h via alternating least squares (ALS). The rationale is that if the residual error contains a factor structure orthogonal to $s_{ig}\epsilon_{gt}$, the ALS procedure recovers both the factor and the causal coefficient without bias. After removing two-way fixed effects, the residual contains the GE term:

$$\eta_{it+h} = \gamma s_{ir} R_{t+h} + u^y_{it+h}. \tag{A76}$$

From (S3), the residual satisfies:

$$\text{(A77)} \qquad \eta_{it+h} = \gamma s_{ir} \underbrace{\left[\psi_h^R(\epsilon_{gt}) + \psi_h^R(\epsilon_{rt})\right]}_{R_{t+h}} + u_{it+h}^y = \underbrace{\gamma s_{ir}}_{\lambda_i} \underbrace{\left[\psi_h^R(\epsilon_{gt}) + \psi_h^R(\epsilon_{rt})\right]}_{f_t} + u_{it+h}^y$$

where $\psi_h^R(\epsilon_{kt})$ denotes the component of R_t driven by contemporaneous and past ϵ_k shocks for $k = g, r$. The residual admits a rank-1 factor structure, as all components of R_{t+h} load on the same cross-sectional vector s_{ir}. However, this factor is a function of both ϵ_{gt} and ϵ_{rt}, and is therefore correlated with the regressor $s_{ig}\epsilon_{gt}$. As a consequence, the ALS algorithm cannot separate the variation in Y_{it+h} generated by $s_{ig}\epsilon_{gt}$ from that generated by $\gamma s_{ir}\psi_h^R(\epsilon_{kt})$, because the two share the same cross-sectional loading direction and their time-series components are correlated.

With a single factor, the IFE model is correctly specified in terms of rank. However, the factor $f_t = \psi_h^R(\epsilon_{gt}) + \psi_h^R(\epsilon_{rt})$ is itself a function of ϵ_{gt} and therefore correlated with the regressor $s_{ig}\epsilon_{gt}$ in both time-series and cross-sectional dimensions. As a result, the orthogonality condition required for IFE to recover the causal coefficient fails. In the ALS procedure, the regressor and the factor compete to explain the same variation in Y_{it+h}, leading the estimator to attribute part of the treatment effect to the factor and producing an overcorrection relative to TWFE.

Figure A2 presents results from a Monte Carlo simulation to illustrate the previous discussion.

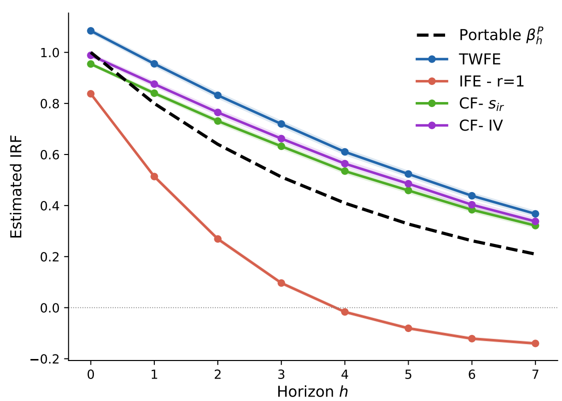


FIGURE A2. Monte Carlo Simulations - Alternative Estimation Approaches

Back to reference in main text.
Simulation results based on 1000 repetitions with $N = 100$ and $T = 300$. 90% confidence bands. The DGP is generated using the following parameter values: $\rho_g = \rho_r = .8$, $\delta = .5$, $\beta = 1$, $\gamma = .5$, $\phi = .5$. The variance of all innovations is set to unity.

A.4.2. Controlling for $s_{ir}R_t$

Suppose that we had knowledge of the exposure shifters s_{ir} and extended the baseline specification with an interaction between these and R_t:

$$Y_{it+h} = \beta_h^{CF-s_{ir}} \times s_{ig}\epsilon_{gt} + \gamma_h \times s_{ir}R_t + \lambda_{ih} + \lambda_{th} + e_{it+h} \tag{A78}$$

This strategy fails to recover the true portable elasticity under the same conditions as the control function approach. The interaction $s_{ir}R_t$ exploits the same time-series variation as the control function and therefore cannot separately identify the component of R_t driven by ϵ_{gt}. As a result, it does not control for the endogenous response of R_t to ϵ_{gt} unless the corresponding time paths are linearly dependent.

This extension highlights that the difficulty in addressing the bias associated with HEDGE stems from the dynamic response of macroeconomic aggregates. Figure A2 presents results from a Monte Carlo simulation to illustrate this point.

A.4.3. Controlling for R_t instrumented with ϵ_{rt}

This strategy differs from the control function approach by instrumenting R_t rather than using its realized value. By a similar argument, it also fails. While the control function uses the variation in R_t orthogonal to ϵ_{gt}, this approach isolates the component of R_t driven by ϵ_{rt}. Unless the path of R_t in response to ϵ_{rt} spans its response to ϵ_{gt}, this variation is insufficient to account for the endogenous response of R_t to ϵ_{gt}.

Figure A2 presents results from a Monte Carlo simulation to illustrate this point.

A.5. Forward-Looking Settings - Further Details

Figure A3 shows the response of R to ϵ_{gt}, the fiscal shock, and to each of the two scenarios for the persistence of ϵ_{it}, the monetary shock.

Figure A4 presents additional simulation results. The upper row is an extension of Figure 3 in the main text that considers additional news shocks. It can be seen that, both in the *Misaligned* and *Aligned*, incorporating further news shocks improves the performance of the decomposition until it finally recovers the true portable elasticities. In the *Aligned* case, using news beyond $\tilde{F} = 4$ yields negligible gains.

The bottom row considers an economy with weaker forward-looking dynamics by introducing a discount term in the Euler equation[39] of households (McKay, Nakamura, and Steinsson 2017). Discounting weakens the importance of expected output gaps for current

[39] The Euler equation is given by:

$$C_t^{-\frac{1}{\sigma}} = \vartheta\beta E_t[C_{t+1}{}^{-\frac{\alpha}{\sigma}}(1 + r_{t+1})], \tag{A79}$$

where α is the discounting parameter and ϑ is a constant calibrated to match steady-state consumption.

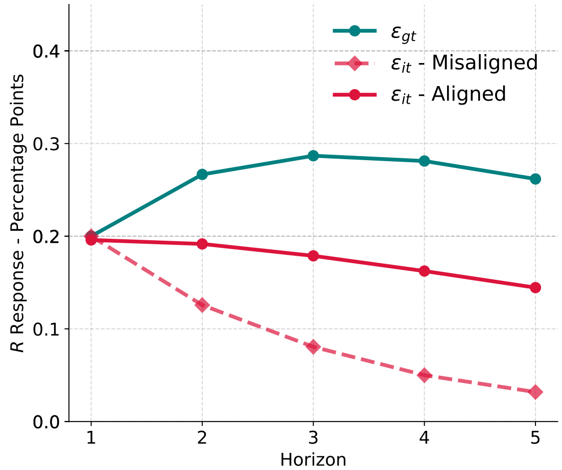
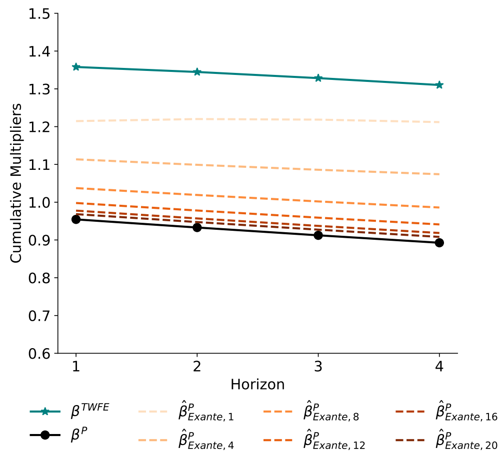


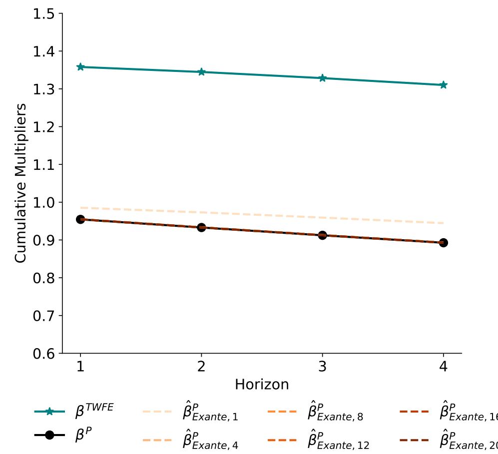
FIGURE A3. Path of R Conditional on Aggregate Shocks

Back to reference in main text.

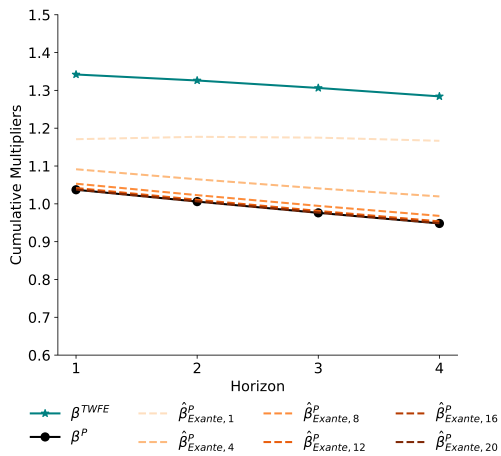
consumption responses and, overall, moves the economy closer to the Markov setting. I set the discounting parameter to $\alpha = .95$ (relative to a baseline without discounting, $\alpha = 1$). As expected, decreasing the strength of forward-looking dynamics improves the performance of a decomposition that uses a limited number of news shocks.



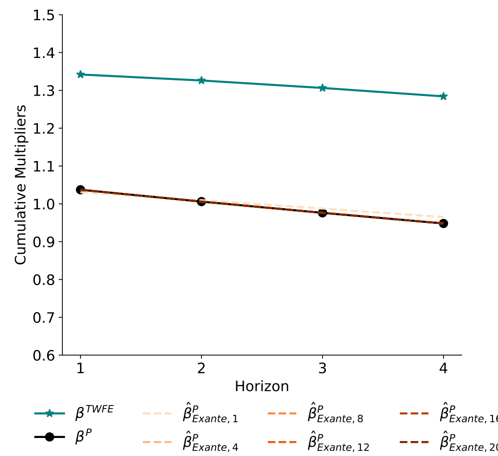
A. *Misaligned* R path - No Discounting



B. *Aligned* R path - No Discounting



C. *Misaligned* R path - Discounting



D. *Aligned* R path - Discounting

FIGURE A4. Forward-looking Settings - Additional Simulation Results

Misaligned refers to the case where the trajectory of R differs across shocks, while *Aligned* corresponds to the case where these trajectories are relatively similar. The top row shows the estimation results using different numbers of news shocks in an economy without discounting (i.e., standard Euler equation), while the bottom row corresponds to an economy with a discounted Euler equation (McKay, Nakamura, and Steinsson 2017). The discounting parameter is set to $\alpha = .95$.

Back to reference in main text.

Back to reference in main text.

A.6. Asymptotic Properties

Back to reference in main text.

This section establishes the consistency of the estimators for β_h^P. The argument proceeds in three steps. First, I establish consistency of the panel local projection estimators under large-N, fixed-T asymptotics. Second, I establish consistency of the time-series local projection estimators under large-T asymptotics. Third, I show that $\hat{\beta}_h^P$ is a continuous mapping of these estimators and therefore consistent by the Continuous Mapping Theorem under suitable non-degeneracy conditions.

Assumptions for Panel Local Projections

Throughout, I consider asymptotics as $N \to \infty$ for T fixed. Aggregate shocks and time fixed effects are treated as fixed in the cross-sectional limit.

ASSUMPTION A1.

$$E_i\left[s_{ig}e_{it+h} \mid \boldsymbol{X}\right] = 0,$$

where $\boldsymbol{X} = \epsilon_{gt}, \epsilon_{rt}, \lambda_{ih}, \lambda_{th}$. This is the key identifying condition for a panel TWFE local projection (that relies on cross-sectional variation for identification). It holds if, conditional on the realization of aggregate shocks, the fixed effects and included controls, high and low exposure units do not experience systematically different unobserved shocks.

ASSUMPTION A2.

$$\sup_{i,t} E\left[e_{it+h}^2 \mid \boldsymbol{X}\right] \leq \sigma^2 < \infty, \quad \sup_i s_{ig}^2 < \infty$$

ASSUMPTION A3. *For each t and h,*

$$\frac{1}{N^2}\sum_{i \neq j} s_{ig}s_{jg}cov[e_{it+h}, e_{jt+h} \mid \boldsymbol{X}] \to 0 \quad as\, N \to \infty.$$

This condition allows for common factors with heterogeneous loadings and/or cluster components. It requires that exposure shares are sufficiently dispersed within and across clusters, and asymptotically orthogonal to heterogeneous loadings (i.e. no HEDGE). This assumption does not preclude arbitrary serial correlation within units i (i.e. $cov[e_{it}, e_{is} \mid \boldsymbol{X}]$ is unrestricted for $s \neq t$).

ASSUMPTION A4.

$$\frac{1}{N}\sum_{i=1}^{N} \tilde{s}_{ig}^2 \to \mu_{s^2} > 0 \quad as\, N \to \infty,$$

where $\tilde{s}_{ig}$ denotes the within-transformed exposure. This ensures that cross-sectional variation in exposures does not degenerate and is an identification condition that allows b_h and c_{hf} to be separately identified from λ_{th}.

ASSUMPTION A5. *The aggregate shocks satisfy:*

$$\sum_{t=1}^{T} \tilde{\epsilon}_{gt}\,\tilde{\epsilon}_{rt+f} = 0 \quad \forall f = 0, \ldots, F, \qquad \textit{and} \qquad \sum_{t=1}^{T} \tilde{\epsilon}_{rt+f}\,\tilde{\epsilon}_{rt+f'} = 0 \quad \forall f \neq f'.$$

This holds exactly when all shock series are mutually orthogonal in the time-series dimension, which can be imposed without loss of generality by orthogonalizing the shocks prior to estimation. Assumption A5 is made to simplify the proof exposition because it allows the multivariate regression to be analyzed via scalar OLS formulas.

After within transformation:

$$\tilde{y}_{it+h} = b_h \tilde{s}_{ig} \tilde{\epsilon}_{gt} + \sum_{f=0}^{F} c_{h,f} \tilde{s}_{ig} \tilde{\epsilon}_{rt+f} + \tilde{e}_{it+h},$$

where $\tilde{z}$ denotes the within transformation of variable z.

Under Assumption A5, the regressors $\tilde{s}_{ig}\tilde{\epsilon}_{gt}$ and $\{\tilde{s}_{ig}\tilde{\epsilon}_{rt+f}\}_{f=0}^{F}$ are mutually orthogonal in the sample, so the TWFE estimators reduce to:[40]

$$\hat{b}_h = b_h + \frac{\sum_{t,i} \tilde{s}_{ig}\tilde{\epsilon}_{gt}\tilde{e}_{it+h}}{\sum_{t,i}(\tilde{s}_{ig}\tilde{\epsilon}_{gt})^2},$$

$$\hat{c}_{hf} = c_{hf} + \frac{\sum_{t,i} \tilde{s}_{ig}\tilde{\epsilon}_{rt+f}\tilde{e}_{it+h}}{\sum_{t,i}(\tilde{s}_{ig}\tilde{\epsilon}_{rt+f})^2}.$$

For fixed T and $N \to \infty$, it suffices to show that[41]

$$\frac{1}{NT} \sum_{t,i} (\tilde{s}_{ig}\tilde{\epsilon}_{kt})^2 \to C > 0 \;\; \forall\, k = g, r,$$

$$A_{Nth} = \frac{1}{N} \sum_{i} \tilde{s}_{ig}\tilde{e}_{it+h} \xrightarrow{p} 0 \;\text{ for each } t \text{ and } h.$$

[40]The general multivariate case without Assumption A5 requires positive definiteness of the matrix $Z'Z/NT$. Because all regressors share the same cross-sectional loading $\tilde{s}_{ig}$, this matrix converges to $\mu_{s^2} \cdot \Sigma_\epsilon$, where Σ_ϵ is the variance-covariance matrix of the shock series. Positive definiteness therefore holds if and only if Assumption A4 holds and the shocks are linearly independent over the T observed periods. The latter is assumed in the time-series step of the proof.

[41]The condition for the numerator follows from the fact that the numerator can be factored as:

$$\frac{1}{NT} \sum_{t,i} \tilde{s}_{ig}\tilde{\epsilon}_{kt}\tilde{e}_{it+h} = \frac{1}{T} \sum_{t} \tilde{\epsilon}_{kt} \underbrace{\left(\frac{1}{N} \sum_{i} \tilde{s}_{ig}\tilde{e}_{it+h} \right)}_{= A_{Nth}},$$

where the factoring uses the fact that $\tilde{\epsilon}_{kt}$ is constant across i and fixed conditional on $\boldsymbol{X}$. Since T is fixed, it suffices to show $A_{Nth} \xrightarrow{p} 0$ for each t and h; the full numerator is then a finite weighted sum of $o_p(1)$ terms and hence $o_p(1)$.

For the first condition, since T is fixed and $\tilde{\epsilon}_{kt}$ is constant across i,

$$\frac{1}{NT}\sum_{t,i}(\tilde{s}_{ig}\tilde{\epsilon}_{kt})^2 = \underbrace{\frac{1}{T}\sum_t \tilde{\epsilon}_{kt}^2}_{>0 \text{ a.s.}} \cdot \frac{1}{N}\sum_i \tilde{s}_{ig}^2 \longrightarrow \frac{1}{T}\sum_t \tilde{\epsilon}_{kt}^2 \cdot \mu_{s^2} > 0,$$

where the first factor is strictly positive almost surely provided the shocks are non-degenerate, and the second factor converges to a strictly positive constant by Assumption A4.

For the second condition, it suffices to show that:

$$E[A_{Nth} \mid \boldsymbol{X}] = 0 \quad \text{and} \quad V[A_{Nth} \mid \boldsymbol{X}] \to 0,$$

The first condition follows from Assumptions A1 and A2. Since the within transformation is a bounded linear operation in T, it preserves conditional mean independence.[42] The conditional variance of A_{Nth} is given by:

$$V[A_{Nth} \mid \boldsymbol{X}] = \frac{1}{N^2}\left(\sum_i \tilde{s}_{ig}^2 V[\tilde{e}_{it+h} \mid \boldsymbol{X}] + \sum_{i \neq j} \tilde{s}_{ig}\tilde{s}_{jg} cov[\tilde{e}_{it+h}, \tilde{e}_{jt+h} \mid \boldsymbol{X}]\right)$$

Weak cross-sectional dependence, as stated in A3, implies that the second term in brackets goes to zero asymptotically. Assumptions A4 and A2 imply[43] that the first term goes to zero too, hence:

$$V[A_{Nth} \mid \boldsymbol{X}] \leq \frac{C\sigma^2}{N}\frac{1}{N}\sum_i \tilde{s}_{ig}^2 = O\left(\frac{1}{N}\right) \to 0 \; as \; N \to \infty$$

Then, by Chebyshev's inequality:

$$P(|A_{Nth} - E[A_{Nth} \mid \boldsymbol{X}]| > \epsilon \mid \boldsymbol{X}) \leq \frac{V[A_{Nth} \mid \boldsymbol{X}]}{\epsilon^2} \to 0.$$

Since $E[A_{Nth} \mid \boldsymbol{X}] = 0$, we have $|A_{Nth} - E[A_{Nth} \mid \boldsymbol{X}]| = |A_{Nth}|$, and by the law of total probability:

$$P(|A_{Nth}| > \epsilon) = E[P(|A_{Nth}| > \epsilon \mid \boldsymbol{X})] \to 0.$$

[42]The within transformation writes $\tilde{e}_{it+h} = e_{it+h} - \bar{e}_{i\cdot} - \bar{e}_{\cdot t+h} + \bar{e}_{\cdot\cdot}$, where each cross-sectional average $\bar{e}_{\cdot t+h} = \frac{1}{N}\sum_j e_{jt+h}$ satisfies $E[\tilde{s}_{ig}\bar{e}_{\cdot t+h} \mid \boldsymbol{X}] = 0$ by Assumption A1 applied uniformly across j.

[43]The within transformation preserves the boundedness of second moments.

Therefore:

$$\hat{b}_h - b_h \xrightarrow{p} 0, \qquad \hat{c}_{hf} - c_{hf} \xrightarrow{p} 0 \quad \forall\ h \text{ and } \forall\ f.$$

Assumptions for Time-series Local projections

ASSUMPTION A6. *$\{\{\epsilon_{rt+f}\}_{f=0}^{F}, \epsilon_{gt}, e_{t+h}\}$ is covariance-stationary and ergodic with:*

$$E[\epsilon_{rt+f}] = E[\epsilon_{gt}] = E[e_{t+h}] = 0$$

ASSUMPTION A7.

$$E[\epsilon_{rt+f}^2] = \sigma_{r,f}^2 < \infty, E[\epsilon_{gt}{}^2] = \sigma_g^2 < \infty, \quad E[e_{t+h}^2] = \sigma_e^2 < \infty$$

ASSUMPTION A8.

$$E[\epsilon_{rt+f}^2] = \sigma_{r,f}^2 > 0, \quad E[\epsilon_{gt}^2] = \sigma_g^2 > 0$$

This condition guarantees that the shocks have variation over time.

ASSUMPTION A9.

$$E[\epsilon_{rt+f} e_{t+h}] = 0, \quad E[\epsilon_{gt} e_{t+h}] = 0$$

The OLS estimator $(\hat{\nu}_h, \{\hat{a}_{h,f}\}_{f=0}^{F})$ satisfies:

$$\begin{pmatrix} \hat{\nu}_h - \nu_h \\ \{\hat{a}_{hf} - a_{hf}\}_{f=0}^{F} \end{pmatrix} = \left(\frac{1}{T} \sum_t z_t z_t^\top \right)^{-1} \frac{1}{T} \sum_t z_t\, e_{t+h}$$

where $z_t = (\epsilon_{gt}, \{\epsilon_{rt+f}\}_{f=0}^{F})^\top$. By Assumption A6 and A8, as $T \to \infty$, $\frac{1}{T} \sum_t z_t z_t^\top \xrightarrow{p} \Sigma_{zz}$, which is positive definite under Assumption A8. By Assumptions A9 and A7 and the ergodic LLN, $\frac{1}{T} \sum_t z_t e_{t+h} \xrightarrow{p} E[z_t e_{t+h}] = 0$. Slutsky's theorem gives $(\hat{\nu}_h - \nu_h, \{\hat{a}_{hf} - a_{hf}\}_{f=0}^{F})^\top \xrightarrow{p} 0$.

Consistency of $\hat{\beta}_h^P$

For the Markov setting, the portable elasticity at horizon h is defined as:

$$\beta_h^P = g^M\Big(b_h, \{c_s, \nu_s, a_s\}_{s=0}^{H}\Big)$$

where g^M is a continuous mapping on the domain $\Theta_h^M = \Big\{(b_h, \{c_s, \nu_s, a_s\}_{s=0}^{H}) : \nu_0 \neq 0\Big\}$. In particular, the only parameter entering the denominator of g^M is a_0.

ASSUMPTION A10. *Non-zero Impact effect of ϵ_{rt} in the Markov setting:*

$$a_0 \neq 0$$

This assumption requires that the instrument for the GE variable has a non-zero effect upon impact and implies that g^M is continuous at the true parameter vector.

For forward-looking settings, the portable elasticity at horizon h is defined as:

$$\beta_h^P = g^{FL}\Big(b_h, \{c_{hf}\}_{f=0}^F, \{\nu_s, \{a_{sf}\}_{f=0}^F\}_{s=0}^H\Big)$$

where g^{FL} is a continuous mapping on the domain $\Theta_h^{FL} = \Big\{(b_h, \{c_{hf}\}_{f=0}^F, \{\nu_s, \{a_{sf}\}_{f=0}^F\}_{s=0}^H) :$ $\boldsymbol{A}$ has full column rank$\}$. $\boldsymbol{A}$ is a matrix that collects the impulse responses of R to the news shocks across horizons.

ASSUMPTION A11. *Rank Condition in the Forward-looking setting.*
The matrix $\boldsymbol{A}$ has full column rank.

This condition ensures the uniqueness of the solution to the least squares problem and guarantees that the mapping from time-series coefficients to the vector of counterfactual shocks, and hence to β_h^P, is continuous in a neighborhood of the true parameter vector. This is a sufficient condition for consistency.[44]

Since the parameter vector is finite-dimensional, marginal convergence in probability of each component implies joint convergence in probability. Given consistency of $(\hat{b}_h, \hat{c}_{hf}, \hat{\nu}_h, \hat{a}_h)$ for all $h \leq H$ and for all $f \in F$, it follows that

$$\hat{\Theta}_h^M = (\hat{b}_h, \{\hat{c}_s\}_{s=0}^H, \{\hat{\nu}_s, \hat{a}_s\}_{s=0}^H) \xrightarrow{p} (b_h, \{c_s\}_{s=0}^H, \{\nu_s, a_s\}_{s=0}^H) = \Theta_h^M,$$
$$\hat{\Theta}_h^{FL} = (\hat{b}_h, \{\hat{c}_{hf}\}_{f=0}^F, \{\hat{\nu}_s, \{\hat{a}_{sf}\}_{f=0}^F\}_{s=0}^H) \xrightarrow{p} (b_h, \{c_{hf}\}_{f=0}^F, \{\nu_s, \{a_{sf}\}_{f=0}^F\}_{s=0}^H) = \Theta_h^{FL}.$$

By the Continuous Mapping Theorem,

$$\hat{\beta}_h^P = g^k(\hat{\Theta}_h^k) \xrightarrow{p} g^k(\Theta_h^k) = \beta_h^P \quad \text{for} \;\; k = M, FL.$$

A.7. Inference

Back to reference in main text.

This appendix describes the procedure to compute standard errors for the estimated portable elasticities.

[44]Full column rank of A is a sufficient but not necessary condition for consistency. If A is rank deficient, the least squares problem admits multiple solutions for the vector of counterfactual shocks. However, β_h^P remains uniquely defined provided that the response of the unit-level outcome to the news shocks lies in the row space of A. Equivalently, for any vector d in the null space of A, it must hold that $\sum_{f=0}^F c_{hf} d_f = 0$ for all h. Economically, this condition requires that any linear combination of news shocks that leaves the path of R unchanged must also leave the path of the unit-level outcome unchanged. This is implied by the assumption of instrument validity, which requires that the news shocks affect the unit-level outcome only through their equilibrium effect on the path of R. Under this restriction, β_h^P depends only on the projection of $\{c_{hf}\}_f$ onto the row space of A, and full rank is not necessary for consistency.

Let θ^h denote the $(K_p + K_t) \times 1$ vector of parameters at horizon h, where K_p is the number of panel parameters and K_t is the number of time-series parameters. The set of parameters $\Theta = [\theta^{h'}]_{h=1}^{H}$ can be estimated via just identified GMM. For panel regressions, assume the data are already two-way demeaned; let $\tilde{x}_{it}$ represent the two-way demeaned version of the variable x_{it}. The moment conditions for each horizon h are given by:

$$m_k^{P,h}(\theta^h) = \frac{1}{NT}\sum_{i,t}\tilde{x}_{it}^k\tilde{e}_{i,t+h}, \quad k = 1,\dots,K_p, \tag{A80}$$

$$m_{k'}^{T,h}(\theta^h) = \frac{1}{T}\sum_{t}x_t^{k'}e_{t+h}, \quad k' = 1,\dots,K_t, \tag{A81}$$

where $\tilde{e}_{i,t+h}$ and e_{t+h} represent residuals from the panel and time-series regressions, respectively. Throughout, I refer to the first set as *panel moment conditions* and the second set as *time-series moment conditions*. Define the full vector of moment conditions for horizon h as:

$$m^h(\theta^h) = \left(m^{P,h}(\theta^h)\ m^{T,h}(\theta^h)\right), \tag{A82}$$

where $m^{P,h}(\theta^h)$ and $m^{T,h}(\theta^h)$ stack the panel and time-series moments, respectively. Let the stacked moment vector across all horizons be denoted as:

$$m(\Theta) = \left(m^1(\theta^1)\ m^2(\theta^2) \vdots m^H(\theta^H).\right) \tag{A83}$$

The vector of parameters Θ can be estimated in one step by minimizing the quadratic form:

$$\hat{\Theta} = \underset{\Theta}{\operatorname{argmin}}\ m(\Theta)'Wm(\Theta), \tag{A84}$$

where $W = I$, the identity matrix.

A.7.1. Standard Errors

The $(K_p+K_t)H \times (K_p+K_t)H$ covariance matrix of the moment conditions, $\hat{S}$, is calculated as follows. First, I assume zero covariance between moment conditions on different horizons h, implying that $\hat{S}$ is a block diagonal with blocks $\hat{S}^h$.[45] Within each horizon h:

[45] As an alternative I allow for covariance of moment conditions across horizons within each block (i.e., panel and time series moment conidtions). In that case, the covariance matrix is no longer block diagonal, and the off-diagonal blocks are estimated as

$$\hat{S}^{hh'} = \frac{1}{T}\sum_{t=1}^{T} m_t^h(\hat{\theta}^h)\, m_t^{h'}(\hat{\theta}^{h'})', \quad h \neq h'.$$

a. For panel moments, the approach can flexibly accommodate three clustering approaches: (i) time, (ii) unit, and (iii) two-way clustering, denoted respectively by $\hat{S}^h_t$, $\hat{S}^h_i$, and $\hat{S}^h_{2w}$.

b. For time-series moments, I employ a Newey–West heteroskedasticity-and-autocorrelation consistent (HAC) estimator:

$$\hat{S}^h_{HAC} = \hat{\Gamma}(0) + \sum_{\ell=1}^{L} w(\ell)\left[\hat{\Gamma}(\ell) + \hat{\Gamma}(\ell)'\right],$$

where

$$\hat{\Gamma}(\ell) = \frac{1}{T}\sum_{t=\ell+1}^{T} m_t^{T,h}(\hat{\theta}^h) m_{t-\ell}^{T,h}(\hat{\theta}^h)', \quad w(\ell) = 1 - \frac{\ell}{L+1}.$$

c. To capture the covariance between panel and time-series moments within the same time period, I define the aggregated moment vector at time t as

(A85) $$m_t^h(\theta^h) = \begin{pmatrix} m_t^{P,h}(\theta^h) \\ m_t^{T,h}(\theta^h) \end{pmatrix},$$

where

$$m_t^{P,h}(\theta^h) = \begin{pmatrix} \frac{1}{N}\sum_{i=1}^{N} \tilde{x}^1_{i,t}\tilde{e}_{i,t+h} \\ \vdots \\ \frac{1}{N}\sum_{i=1}^{N} \tilde{x}^{K_P}_{i,t}\tilde{e}_{i,t+h} \end{pmatrix}, \qquad m_t^{T,h}(\theta^h) = \begin{pmatrix} x^1_t e_{t+h} \\ \vdots \\ x^{K_t}_t e_{t+h} \end{pmatrix}.$$

The time-clustered covariance between panel and time-series moments is then

(A86) $$\hat{S}^h_{PT} = \frac{1}{T}\sum_{t=1}^{T} m_t^{P,h}(\hat{\theta}^h) m_t^{T,h}(\hat{\theta}^h)'.$$

The covariance matrix for horizon h is given by

(A87) $$\hat{S}^h = \begin{pmatrix} \hat{S}^h_t & \hat{S}^h_{PT} \\ (\hat{S}^h_{PT})' & \hat{S}^h_{HAC} \end{pmatrix}.$$

The full covariance matrix is block diagonal across horizons:

(A88) $$\hat{S} = \mathrm{diag}(\hat{S}^1, \hat{S}^2, \ldots, \hat{S}^H).$$

In Monte Carlo simulations, this approach leads to systematically larger standard errors and overcoverage when ignoring the cross-terms between panel and time series blocks. This may reflect the fact that moment conditions at different horizons load on the same underlying shocks, inducing strong cross-horizon dependence. For this reason, I adopt the block-diagonal specification as the baseline.

For completeness, I also consider allowing for covariance across horizons. For $h \neq h'$, define the off-diagonal block of the covariance matrix as

$$\hat{S}^{hh'} = \begin{pmatrix} \hat{S}^{hh'}_{PP} & \hat{S}^{hh'}_{PT} \\ \hat{S}^{hh'}_{TP} & \hat{S}^{hh'}_{TT} \end{pmatrix}, \tag{A89}$$

where

$$\begin{aligned}
\hat{S}^{hh'}_{PP} &= \frac{1}{T}\sum_{t=1}^{T} m_t^{P,h}(\hat{\theta}^h)\, m_t^{P,h'}(\hat{\theta}^{h'})', \\
\hat{S}^{hh'}_{TT} &= \frac{1}{T}\sum_{t=1}^{T} m_t^{T,h}(\hat{\theta}^h)\, m_t^{T,h'}(\hat{\theta}^{h'})', \\
\hat{S}^{hh'}_{PT} &= \frac{1}{T}\sum_{t=1}^{T} m_t^{P,h}(\hat{\theta}^h)\, m_t^{T,h'}(\hat{\theta}^{h'})', \\
\hat{S}^{hh'}_{TP} &= \frac{1}{T}\sum_{t=1}^{T} m_t^{T,h}(\hat{\theta}^h)\, m_t^{P,h'}(\hat{\theta}^{h'})'.
\end{aligned}$$

Under this specification, the full covariance matrix $\hat{S}$ is no longer block diagonal across horizons. I also consider a restricted version that rules out covariance between panel and time-series moments, both within and across horizons. In that case, I set

$$\hat{S}^{h}_{PT} = 0 \quad \text{for all } h, \qquad \hat{S}^{hh'}_{PT} = \hat{S}^{hh'}_{TP} = 0 \quad \text{for all } h \neq h',$$

while allowing $\hat{S}^{hh'}_{PP}$ and $\hat{S}^{hh'}_{TT}$ to be nonzero. Thus, the off-diagonal blocks reduce to

$$\hat{S}^{hh'} = \begin{pmatrix} \hat{S}^{hh'}_{PP} & 0 \\ 0 & \hat{S}^{hh'}_{TT} \end{pmatrix}, \qquad h \neq h'. \tag{A90}$$

The asymptotic variance-covariance matrix of $\hat{\Theta}$ can be computed as:

$$V(\hat{\Theta}) = \frac{1}{T}(G^{\top}\hat{S}^{-1}G)^{-1},$$

where G is the Jacobian of the moment conditions evaluated at $\hat{\Theta}$. The standard errors for the vector $\hat{\beta}^P = \{\hat{\beta}^P_h\}_{h=0}^{H}$ can be computed using the Delta method as follows:

$$\mathrm{SE}(\hat{\beta}^P) = \sqrt{\nabla f^{\top} V(\hat{\Theta}) \nabla f},$$

where ∇f is the Jacobian of β^P with respect to the vector of estimated parameters.

To evaluate the properties of the proposed inference procedure, I conduct Monte Carlo simulations in both static and dynamic settings.

In the static case, standard errors obtained from the joint GMM system coincide with those from the control function approach only when the covariance between panel and time-series moment conditions is properly accounted for. Ignoring this covariance leads to substantial overcoverage. Thus, the equivalence result for point estimates between the decomposition and control function approaches extends to inference when the shared variation across panel and time-series moments is incorporated.

In the dynamic setting, I consider four alternative estimators of the covariance matrix $\hat{S}$, which differ in whether they allow for (i) covariance between panel and time-series moments within a horizon ("cross-moments") and (ii) covariance across horizons ("cross-h"). The four cases are: (i) allowing cross-moments but imposing independence across horizons; (ii) imposing independence both across blocks and across horizons; (iii) allowing covariance both across blocks and across horizons; and (iv) allowing covariance across horizons within blocks while imposing zero covariance between panel and time-series moments.

The simulation results in FIgure A6 show that accounting for covariance across blocks, as in the static setting, improves coverage for the impact elasticity, but this does not extend to longer horizons. Beyond impact, allowing for cross-moment covariance leads to systematic undercoverage, regardless of whether cross-horizon dependence is included. In contrast, allowing for cross-horizon covariance alone leads to overcoverage. For this reason, I adopt as a baseline the specification that sets both cross-moment and cross-horizon covariances to zero, which displays the right coverage.

In addition, I conduct a Monte Carlo to explore coverage rates in small samples. Concretely, I repeat the exercise setting $N = 50$ and $T = 40$, similar to the sample size of the empirical application in Section 4. Figure A7 presents the results. The baseline specification (ii) exhibits some undercoverage at longer horizons, but still delivers the most stable coverage profile among the estimators considered.

Lastly, Figure A8 presents simulation results using the baseline specification (ii) replacing time-clustering by unit-clustering for panel moments. This exercise sets the sample back to $N = 100$ and $T = 300$. The results are in line with the findings in Almuzara and Sancibrián (2024), who point out that clustering at the unit level with common aggregate shocks can lead to substantial undercoverage.

A.7.2. Standard Error for Estimated Exante Counterfactual Shock

I briefly outline the construction of the standard error of ϵ_r^t for the single-shock case in the ex ante decomposition. The standard error of the ex ante shock is computed using the Delta method, treating the estimated impulse responses as jointly estimated. The ex ante shock size ϵ_r^t is defined as the scalar that minimizes the weighted distance between $\hat{\boldsymbol{R}}_g$

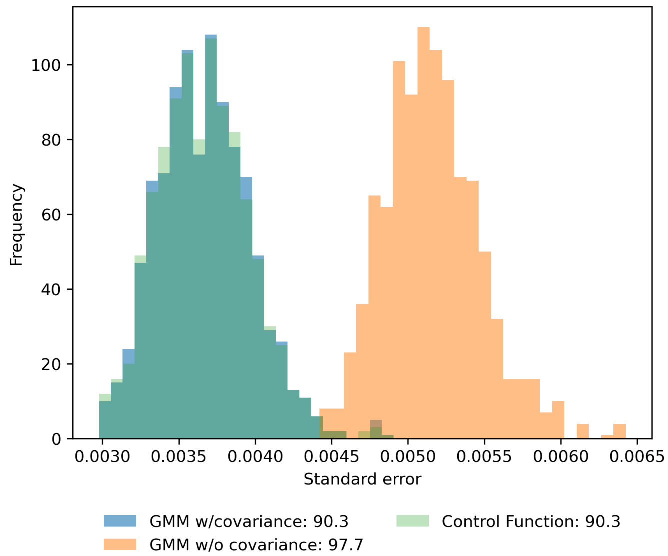


FIGURE A5.
Inference - Monte Carlo Simulation - Static Setting

The plot shows the distribution of standard errors from the control function approach (light green), GMM with time-cluster covariance between panel and time series moments (blue) and GMM with zero covariance across panel a time series moments. The coverage rate is shown in the legend next to the label for each case. Based on 1000 repetitions.

and $\hat{\boldsymbol{R}}_r$,

$$\epsilon_r^t = \arg\min_{\epsilon} (\hat{\boldsymbol{R}}_g - \epsilon\hat{\boldsymbol{R}}_r)' W (\hat{\boldsymbol{R}}_g - \epsilon\hat{\boldsymbol{R}}_r) = \frac{\hat{\boldsymbol{R}}_r' W \hat{\boldsymbol{R}}_g}{\hat{\boldsymbol{R}}_r' W \hat{\boldsymbol{R}}_r}. \tag{A91}$$

As a baseline, I set $W = 1/v(\hat{\boldsymbol{R}}_g)$. Let $z = (\hat{\boldsymbol{R}}_g', \hat{\boldsymbol{R}}_r')'$ collect the estimated impulse responses, with covariance matrix $\hat{S}_z$ obtained from the joint GMM system described above. The asymptotic variance of $\hat{\epsilon_r}^t$ is then given by

$$S(\hat{\epsilon}_r^t) = \nabla_z \epsilon_r^{t\prime} \hat{S}_z \nabla_z \epsilon_r^t, \tag{A92}$$

where $\nabla_z \epsilon_r^t$ denotes the gradient of ϵ_r^t with respect to z. This approach accounts for estimation uncertainty in the estimated impulse responses of R to each aggregate shock.

Back to reference in main text.

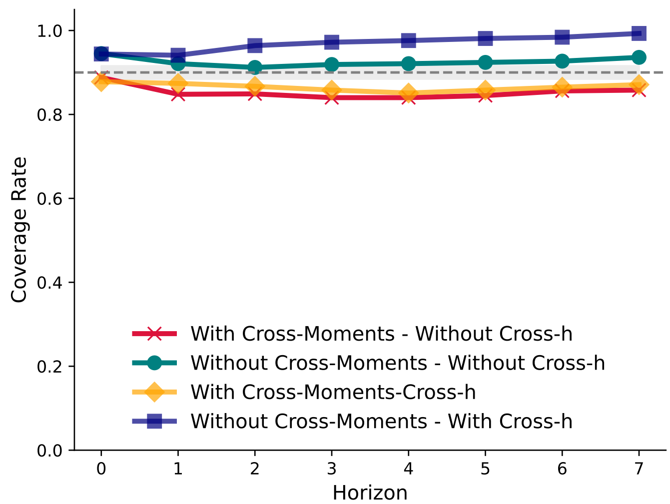


FIGURE A6.
Inference - Monte Carlo Simulation - Dynamic Setting

The figure reports coverage rates across horizons h for four alternative estimators of the covariance matrix $\hat{S}$. *With Cross-Moments – Without Cross-h*: $\hat{S}$ allows for covariance between panel and time-series moment conditions within each horizon (time clustering), but remains block diagonal across horizons. *Without Cross-Moments – Without Cross-h*: $\hat{S}$ is block diagonal both across horizons and between panel and time-series moments. *With Cross-Moments – With Cross-h*: $\hat{S}$ allows for covariance both across horizons and between panel and time-series moment conditions. *Without Cross-Moments – With Cross-h*: $\hat{S}$ allows for covariance across horizons within panel and within time-series moments, while imposing zero covariance between panel and time-series blocks. All results are based on 1000 Monte Carlo repetitions with N=100 and T=300. The DGP is parametrized as $\rho_g = \rho_r = .8$, $\delta = .5$, $\beta = 1$, $\gamma = .5$, $\phi = .5$. The variance of all innovations is set to unity.

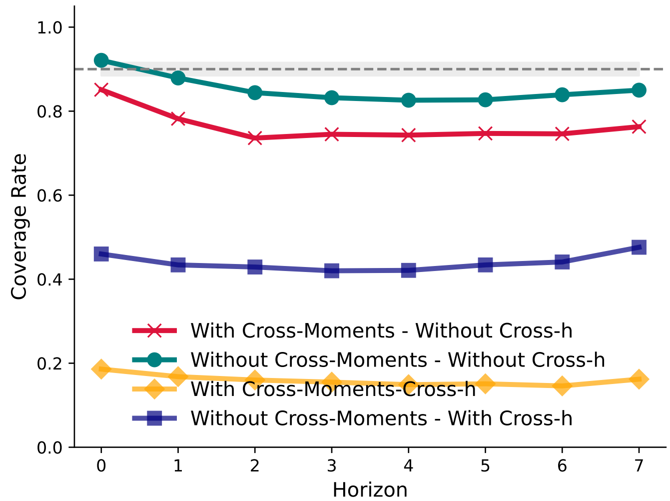


FIGURE A7.
Inference - Monte Carlo Simulation - Dynamic Setting with Small Sample

The figure reports coverage rates across horizons h for four alternative estimators of the covariance matrix $\hat{S}$. *With Cross-Moments – Without Cross-h*: $\hat{S}$ allows for covariance between panel and time-series moment conditions within each horizon (time clustering), but remains block diagonal across horizons. *Without Cross-Moments – Without Cross-h*: $\hat{S}$ is block diagonal both across horizons and between panel and time-series moments. *With Cross-Moments – With Cross-h*: $\hat{S}$ allows for covariance both across horizons and between panel and time-series moment conditions. *Without Cross-Moments – With Cross-h*: $\hat{S}$ allows for covariance across horizons within panel and within time-series moments, while imposing zero covariance between panel and time-series blocks. All results are based on 1000 Monte Carlo repetitions with $N = 50$ and $T = 40$. The DGP is parametrized as $\rho_g = \rho_r = .8$, $\delta = .5$, $\beta = 1$, $\gamma = .5$, $\phi = .5$. The variance of all innovations is set to unity.

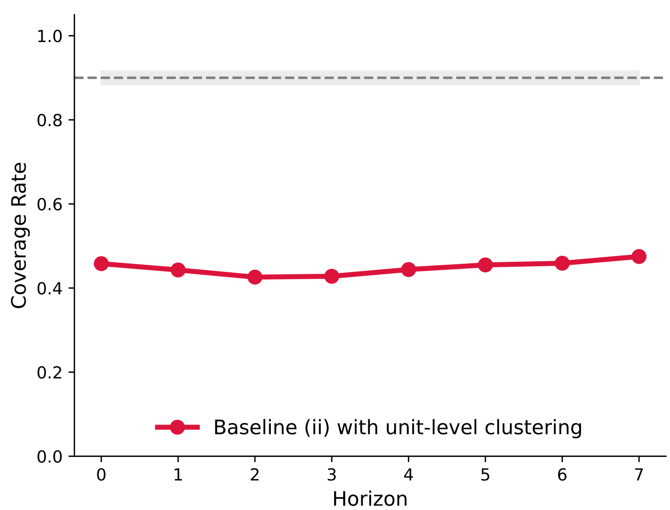


FIGURE A8.
Inference - Monte Carlo Simulation Dynamic Setting with Unit-level Clustering

The figure reports coverage rates across horizons h for the *Without Cross-Moments – Without Cross-h* estimator of the covariance matrix $\hat{S}$ using unit-level clustering for panel moment conditions. All results are based on $J = 1000$ Monte Carlo repetitions with $N = 100$ and $T = 300$. The DGP is parametrized as $\rho_g = \rho_r = .8$, $\delta = .5$, $\beta = 1$, $\gamma = .5$, $\phi = .5$. The variance of all innovations is set to unity.

Appendix B. Data

This appendix presents additional details and descriptive statistics on the data used for the estimation of cross-sectional fiscal multipliers in Section 4.

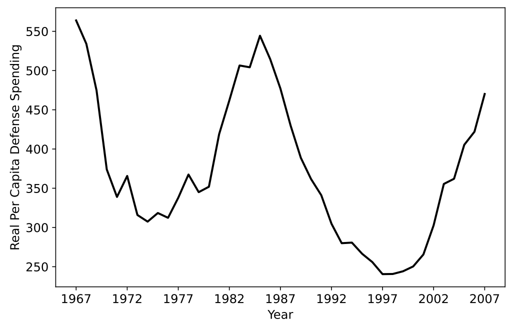
In Table A6 I present descriptive statistics for the main time series and cross-sectional variables used in the regression analysis. The sample period is 1967-2007. G_t^C measures the one-year change in real per-capita defense contracts relative to lagged GDP. The average change in G_t^C is close to zero with a standard deviation of around .3 percent of output. The standard deviation of the Romer-Romer monetary policy shock is close to 1 percentage point. The average exposure to defense spending is .85 with substantial heterogeneity across space. The participation of defense spending in output for the most exposed region is more than three times that of the aggregate economy, whereas for the least exposed one the participation is close to one tenth.

TABLE A6. Descriptive Statistics for Time Series and Cross-Sectional Variables

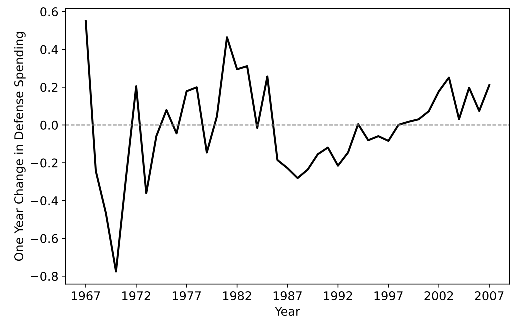
Variable	Mean	Median	Std Dev	Min	Max	Units
Aggregate Variables						
G_t^C	-0.01	-0.04	0.28	-0.78	1.03	% of Y_{t-1}^{pc}
T-bill Rate (3M)	4.20	4.25	3.02	0.03	14.03	%
Real T-bill Rate (3M)	1.19	1.39	1.93	-3.37	5.45	%
Romer-Romer MP shock	0.00	0.21	0.93	-2.18	1.89	%
Cross-sectional Variables						
Y_{it}^C	1.68	1.65	3.82	-29.52	34.39	% of Y_{it-1}^{pc}
G_{it}^C	-0.01	-0.01	0.88	-9.04	7.58	% of Y_{it-1}^{pc}
s_{ig}^{pre}	0.85	0.79	0.54	0.15	3.19	
$\overline{Y}_{it}^{pc}$	16731	16067	3107	11581	30096	Real p.c.

TABLE A7. Top 5 States by Exposure to Defense Spending - 1966-70 Average

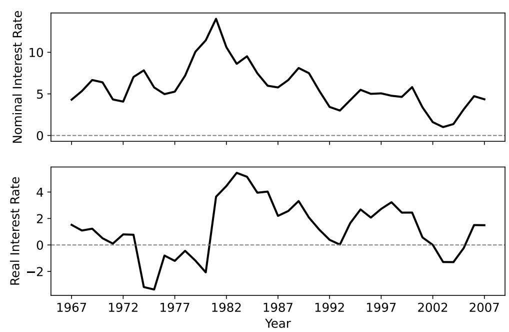
Top 5 States - 1966-70 Average	s_{ig}^{pre}
Connecticut	3.19
Texas	1.87
Missouri	1.70
California	1.68
Massachusetts	1.48



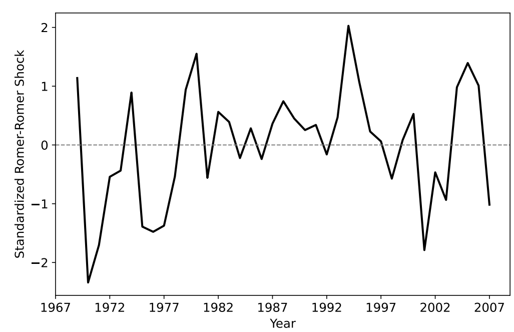
A. National Defense Spending (p.c.)



B. Δ_1 National Defense Spending (% of GDP)



C. 3-month T-bill Rate (%)



D. Romer-Romer Monetary Shock Series

FIGURE A9. Time Series Data

Nominal variables are deflated using the consumer price index. The Romer-Romer monetary shock series is expressed in standard deviations around a zero mean.

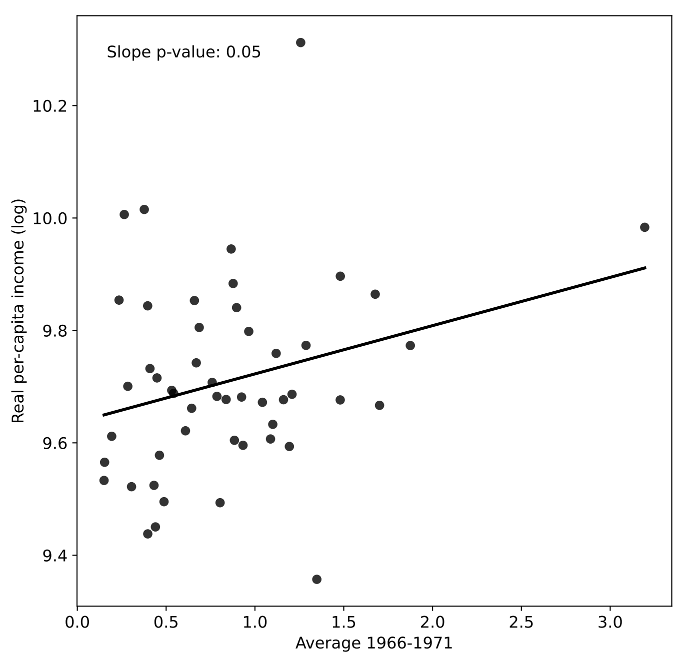


FIGURE A10. Exposure to Defense Spending and Per-Capita Income

Nominal variables are deflated using the consumer price index.

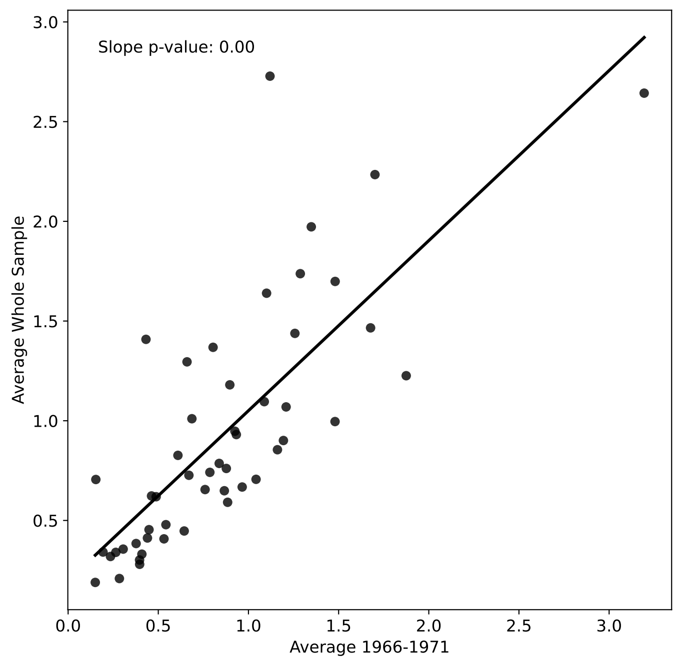


FIGURE A11. Pre-period and Period-Average Exposure to Defense Spending

Pre-period corresponds to the average for 1966-1970 while Period-average corresponds to the average for 1969-2007.

Appendix C. Estimation Results

C.1. HEDGE Test - Robustness

Figure A12 presents results when including the interaction between s_{ig} and the fiscal shock. Formally:

$$Y^C_{it+h} = b_h \times s_{ig} \times G^C_t c_h \times s_{ig} i_t + \lambda_{ih} + \lambda_{th} + \text{Controls} + e_{it+h} \qquad \text{for} \quad h = 1, 4. \tag{A93}$$

The estimated responses are quantitatively unchanged relative to the baseline estimates in Figure 4.

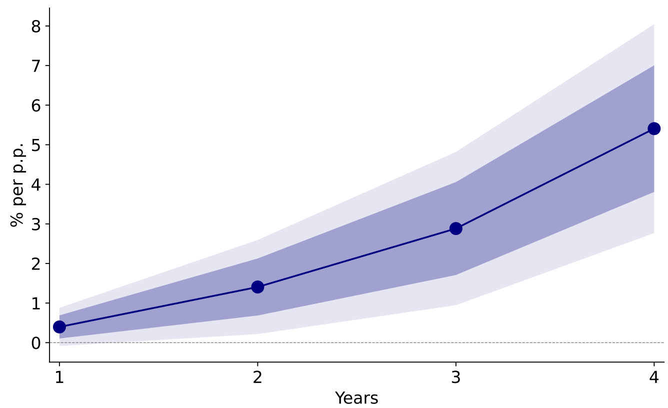


FIGURE A12. HEDGE Test - Robustness to including $s_{ig} \times G^C_t$

Shaded areas correspond to 68% (darker) and 90% (lighter) confidence bands. The sample covers 1970-2007. The point estimates measures the cumulative percent change in local output per percentage point increase in the Federal Funds rate.
Back to reference in main text.

Figure A13 plots the estimated responses when using the series of shocks of Aruoba and Drechsel (2024) instead of the Romer and Romer (2004) shocks. The sample is reduced to the years between 1980-2007 because of the shock series availability.

Lastly, I estimate state-specific output elasticities to changes in the interest rate. For each state j and horizon h, I estimate:

$$Y^C_{jt+h} = \sum_{s=0}^{S} c^j_h \times RR_{t-s} + \phi^j_h Y^C_{jt-1} + e_{jt+h}, \tag{A94}$$

where c^j_h measures the response of Y^C_{jt+h} to a one standard deviation Romer and Romer (2004) monetary shock. Figure A14 plots the estimated responses against s_{ig} - the exposure to defense spending used for identification. Each panel corresponds to a different horizon and the solid black line shows the estimated linear fit. Estimate state-specific elasticities

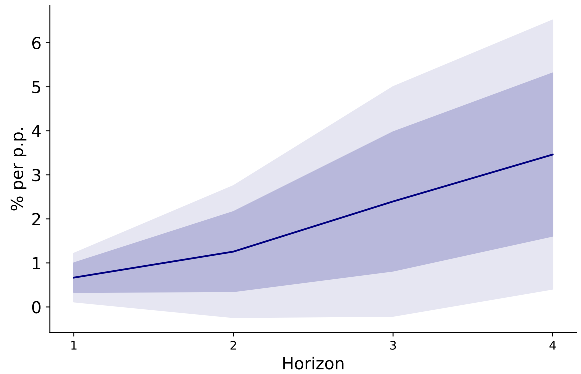
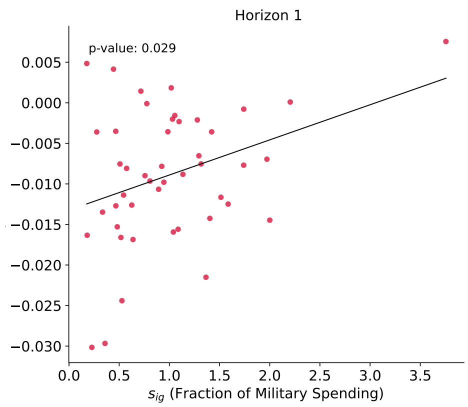


FIGURE A13. HEDGE Test- Differential Response to i_t - Aruoba and Drechsel (2024)

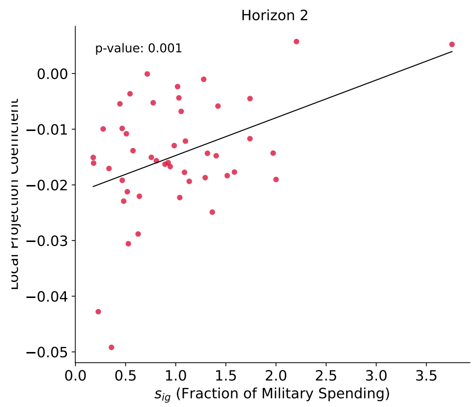
Shaded areas correspond to 68% (darker) and 90% (lighter) confidence bands. The sample covers 1980-2007. The point estimates measures the cumulative percent change in local output per percentage point increase in the Federal Funds rate.

Back to reference in main text.

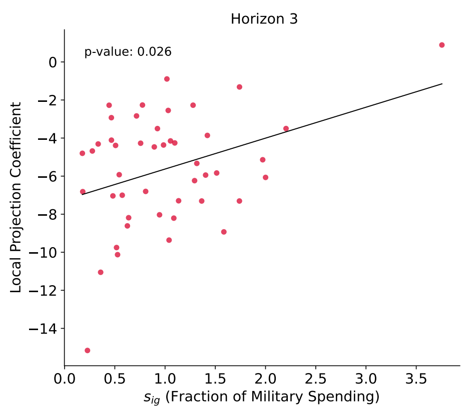
are positively correlated with s_{ig} at all four horizons. The estimated linear relationship is statistically significant at least at the 5% level for all horizons.



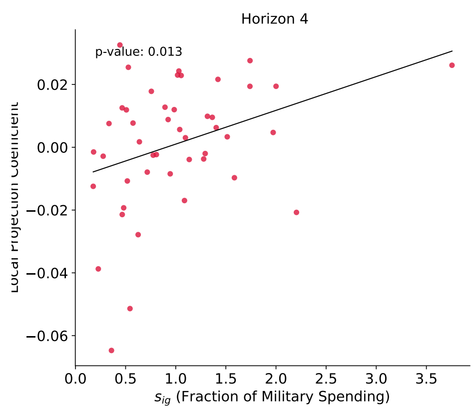
A. $h = 1$



B. $h = 2$



C. $h = 3$



D. $h = 4$

FIGURE A14. State-Specific Interest Rate Elasticities and s_{ig}

Back to reference in main text.

C.2. Decomposition of the US Cross Sectional Multiplier - Additional Results

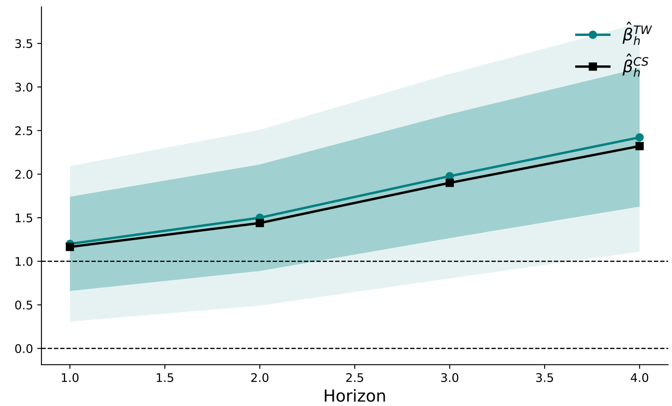


FIGURE A15. Comparison between $\hat{\beta}_h^{TW}$ and $\hat{\beta}_h^{CS}$

Shaded areas correspond to 68% (darker) and 90% (lighter) confidence bands. $\hat{\beta}_h^{TW}$ is the estimated coefficient on $s_{ig}G^C_{it+h}$ using a TWFE specification. $\hat{\beta}_h^{CS}$ is the estimated coefficient on $s_{ig}G^C_{it+h}$ from the cross-sectional step of the decomposition framework. Both specifications control for lagged output growth. Time-clustered standard errors.

Back to reference in main text.

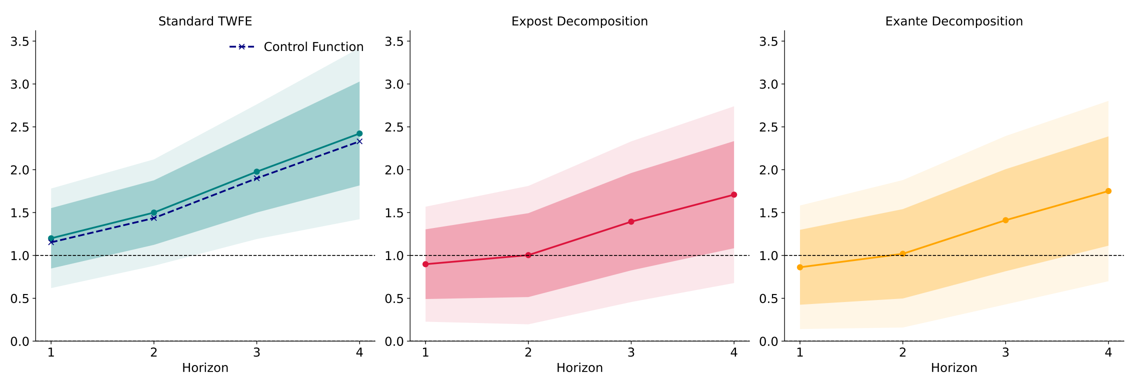


FIGURE A16. Decomposition Results for the US Multiplier - Unit-level Clustering

Shaded area represents 68% (darker) and 90% (lighter) confidence bands. Standard errors are computed as described in Appendix A.7 with unit-level clustering for panel moment conditions. Point estimates show the cumulative percentage change in local output when local defense spending increases by 1% of local output. *Ex-post Decomposition* shows results for the decomposition outlined in Subsection 3.2. *Ex-ante Decomposition* shows results for the decomposition outlined in Subsection 3.3 using a single shock. The blue dotted line shows the results from the control function approach.

Back to reference in main text.

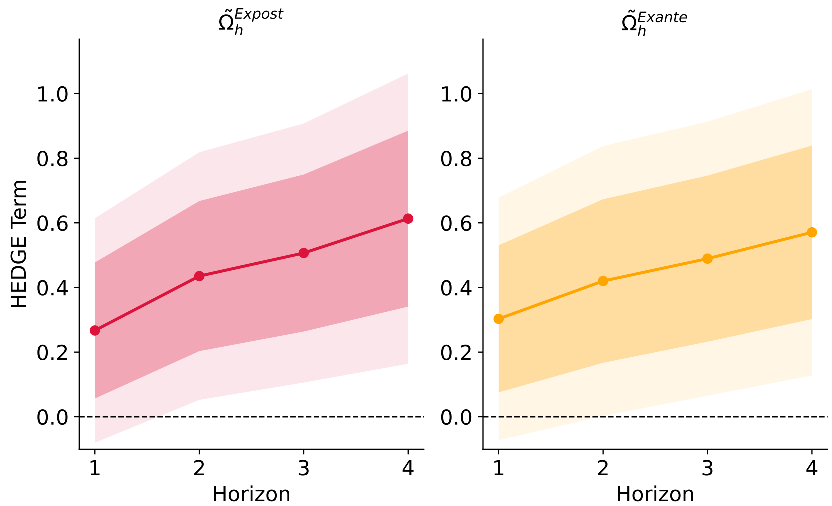


FIGURE A17. Estimated HEDGE Terms - Unit-level Clustering

The plot shows the point estimate and confidence bands for $\tilde{\Omega}_{h,Expost}$ on the left-hand side and for $\tilde{\Omega}_{h,Exante}$ on the right-hand side. Standard errors are computed as outlined in Subsection 3.4 with unit-level clustering for panel moment conditions.

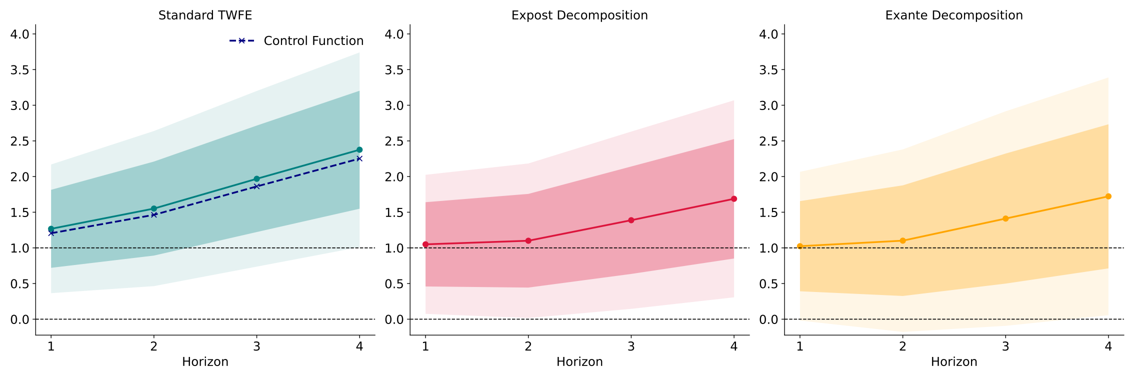
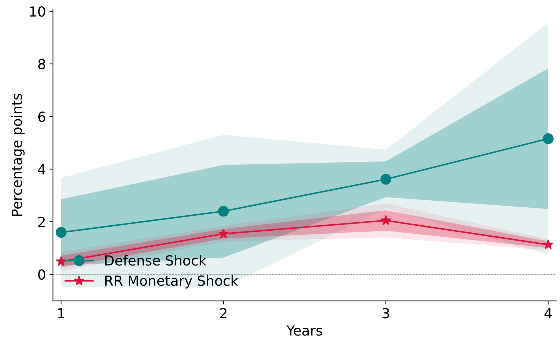


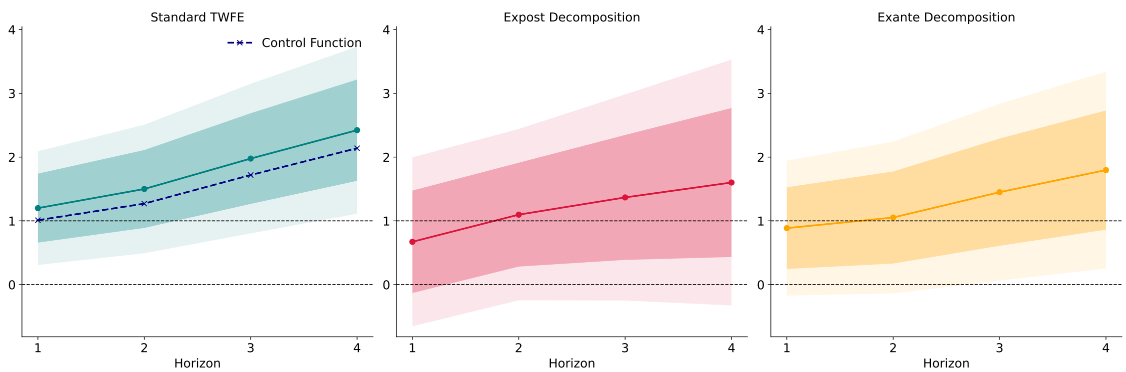
FIGURE A18. Decomposition Results for the US Multiplier - All States

Shaded area represents 68% (darker) and 90% (lighter) confidence bands. Standard errors are computed as described in Appendix A.7 with time-level clustering for panel moment conditions. Point estimates show the cumulative percentage change in local output when local defense spending increases by 1% of local output. *Ex-post Decomposition* shows results for the decomposition outlined in Subsection 3.2. *Ex-ante Decomposition* shows results for the decomposition outlined in Subsection 3.3 using a single shock. The blue dotted line shows the results from the control function approach.

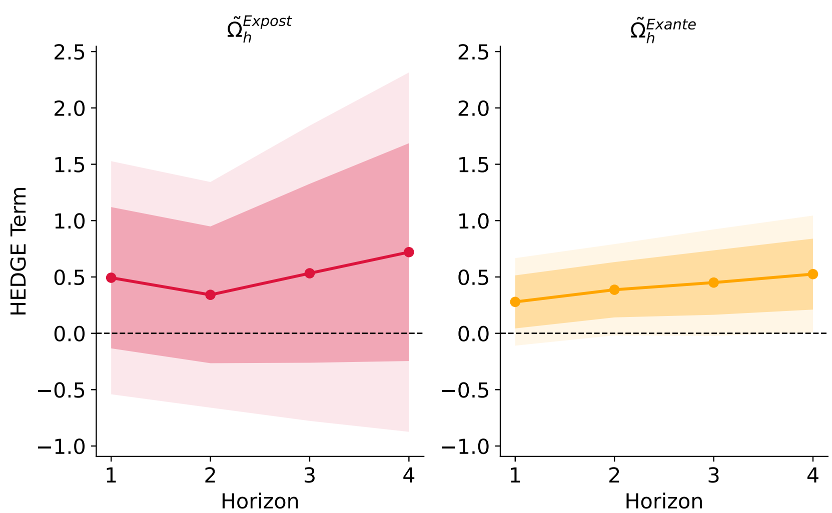
Back to reference in main text.



A. Response of Real Federal Funds Rate



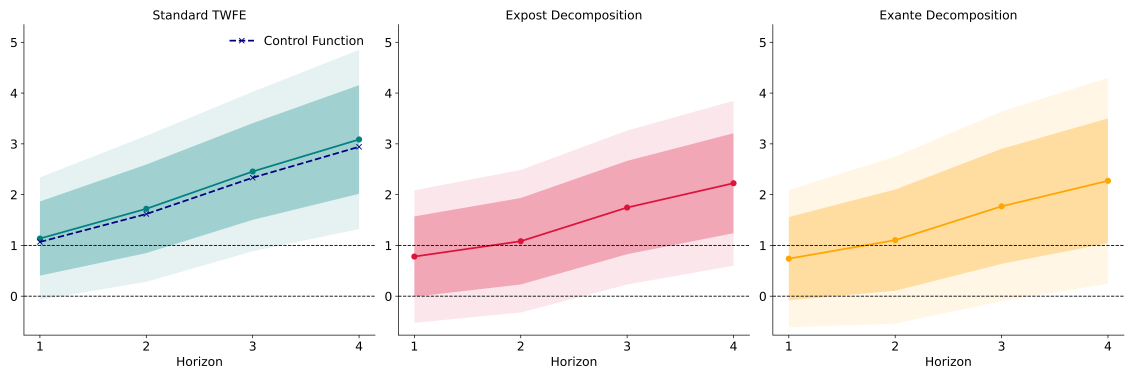
B. Cross-sectional Fiscal Multipliers



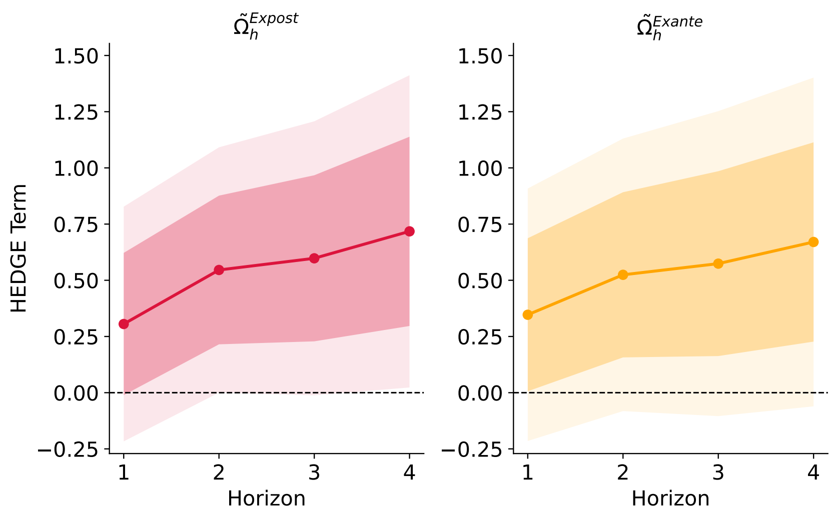
C. HEDGE Terms

FIGURE A19. Results using the Real Federal Funds Rate

Shaded areas correspond to 68% (darker) and 90% (lighter) confidence bands. Results from using the estimated path of the real Federal Funds rate to construct ex-post and ex-ante sequences of counterfactual shocks. All specifications include the same set of controls as in the baseline. Standard errors are computed as outlined in Subsection 3.4 with time-level clustering for panel moment conditions. *Back to reference in main text.*



A. Cross-sectional Fiscal Multipliers

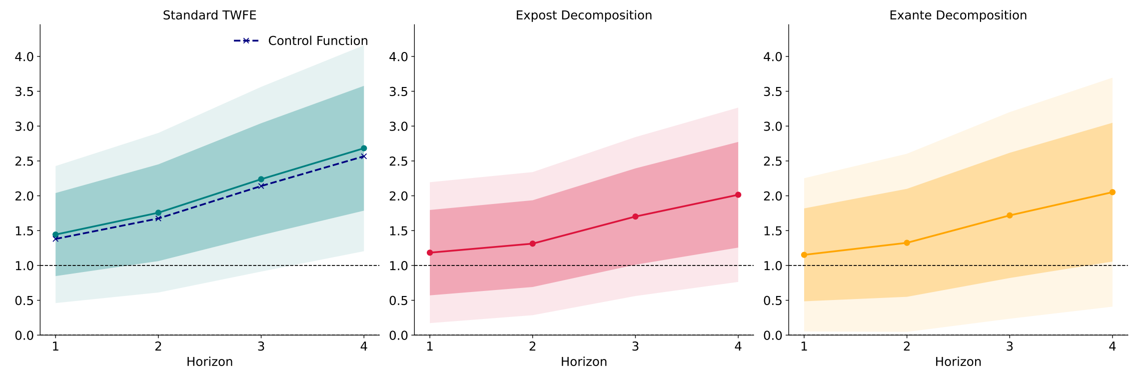


B. HEDGE Terms

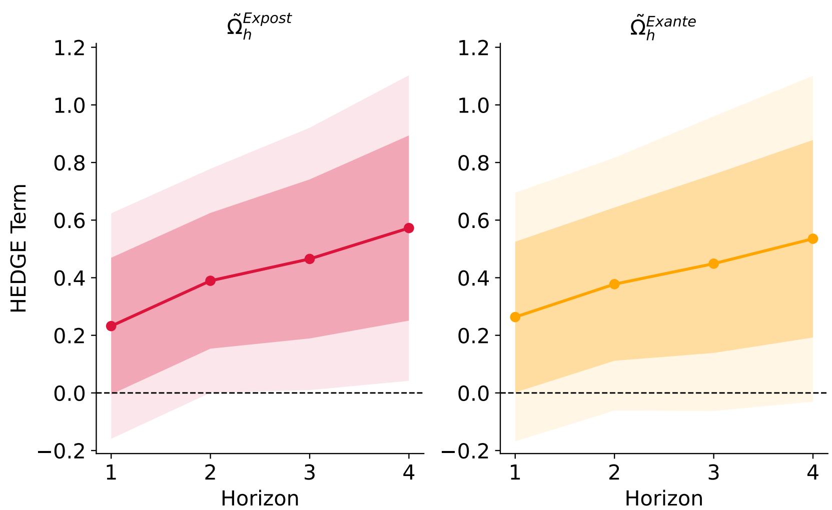
FIGURE A20. Average Exposure Share 1966-2007

Shaded areas correspond to 68% (darker) and 90% (lighter) confidence bands. Results of using the average exposure to defense spending between 1966-2007 instead of the average during the pre-period. Standard errors are computed as outlined in Subsection 3.4 with time-level clustering for panel moment conditions.

Back to reference in main text.



A. Cross-sectional Fiscal Multipliers

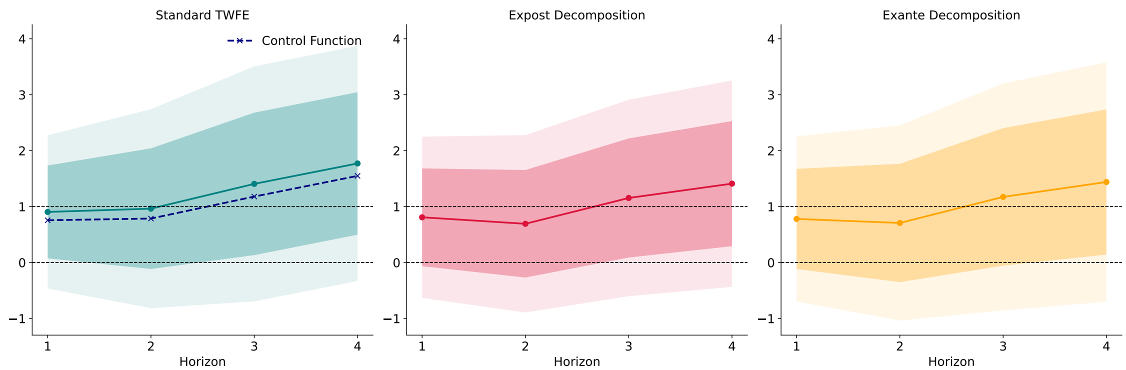


B. HEDGE Terms

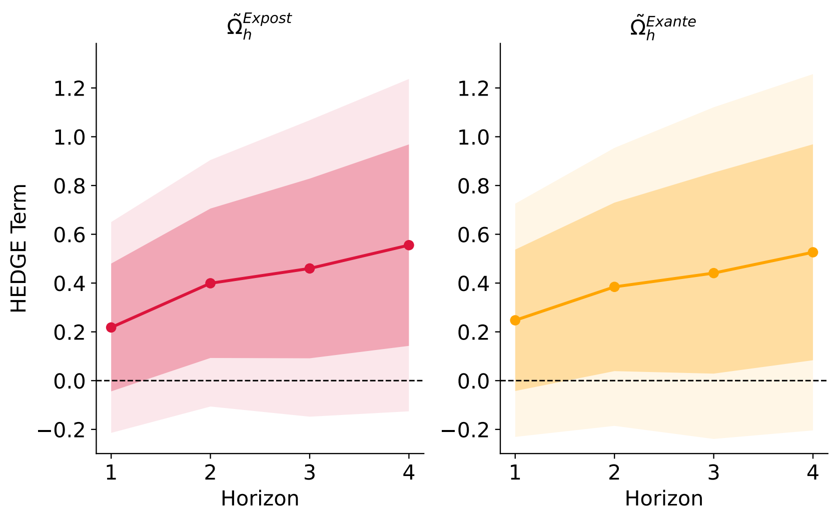
FIGURE A21. No Controls in Cross-sectional Regressions

Shaded areas correspond to 68% (darker) and 90% (lighter) confidence bands. Results from dropping the control for lagged output growth. Standard errors are computed as outlined in Subsection 3.4 with time-level clustering for panel moment conditions.

Back to reference in main text.



A. Cross-sectional Fiscal Multipliers



B. HEDGE Terms

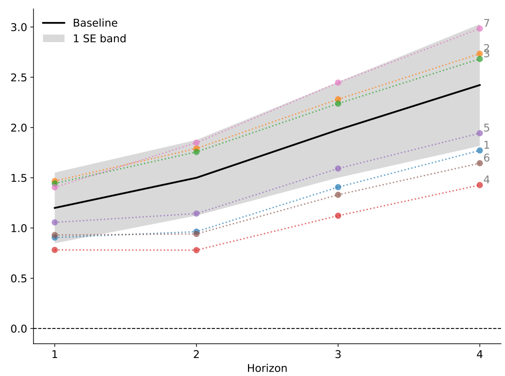
FIGURE A22. Adding Lagged Shocks as Controls

Shaded areas correspond to 68% (darker) and 90% (lighter) confidence bands. Results from adding $s_{ig}G^C_{t-1}$ as control in the TWFE regression and, both $s_{ig}G^C_{t-1}$ and $s_{ig}RR_{t-1}$ in the cross-sectional step regression. Standard errors are computed as outlined in Subsection 3.4 with time-level clustering for panel moment conditions.

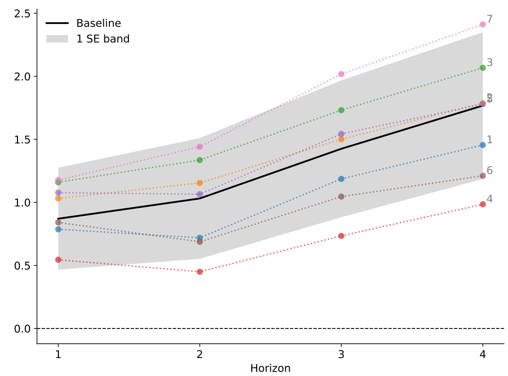
Back to reference in main text.

TABLE A8. Alternative Controls in Cross-sectional Regressions

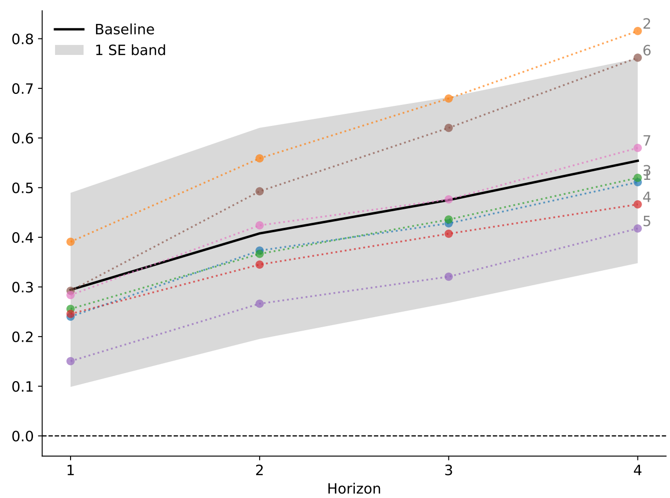
Spec #	**Included Controls**
Baseline	Y^C_{it-1}
1	Y^C_{it-1}, $s_{ig} \times G^C_{t-1}$, $s_{ig} \times RR_{t-1}$
2	$\ln(Y_{it-1})$
3	No controls
4	Y^C_{it-1}, $s_{ig} \times G^C_{t-1}$
5	$s_{ig} \times RR_{t-1}$, $s_{ig} \times G^C_{t-1}$
6	$\ln(Y_{it-1})$, $s_{ig} \times G^C_{t-1}$, $s_{ig} \times RR_{t-1}$
7	Y^C_{it-1}, $s_{ig} \times RR_{t-1}$



A. TWFE Multiplier



B. Portable Fiscal Multipliers - Ex-ante Decomposition



C. HEDGE Term - Ex-ante Decomposition

FIGURE A23. Alternative Controls

Shaded areas correspond to a one standard deviation. Baseline: Y^C_{it-1}; 1: Y^C_{it-1}, $s_{ig} \times G^C_{t-1}$, $s_{ig} \times RR_{t-1}$; 2: $\ln(Y_{it-1})$; 3: No controls; 4: Y^C_{it-1}, $s_{ig} \times G^C_{t-1}$; 5: $s_{ig} \times RR_{t-1}$, $s_{ig} \times G^C_{t-1}$; 6: $\ln(Y_{it-1})$, $s_{ig} \times G^C_{t-1}$, $s_{ig} \times RR_{t-1}$; 7: Y^C_{it-1}, $s_{ig} \times RR_{t-1}$.

Back to reference in main text.

Appendix D. Model Details

This Appendix presents a full model of a monetary union with two heterogeneous regions labeled Home and Foreign. The model is an extension of Nakamura and Steinsson (2014) to include two types of households - *Ricardians* and *Hand-to-Mouth* - as in Herreno and Pedemonte (2025). In addition, regions can differ in their (i) household composition, (ii) intertemporal elasticity of substitution, (iii) size, and (iv) exposure to government spending shocks.

D.1. Home Households

The total population of the monetary union is normalized to one. The size of the Home region is n, while the size of the Foreign region is $(1-n)$. The Home region is populated by two types of agents - *Ricardians* and *Hand-to-mouth* - indexed with subscripts R and C, respectively. The share of the Home population that is *Hand-to-mouth* is given by λ^H. Both households have the same preferences over labor and consumption, but differ in their access to asset markets. They maximize lifetime utility subject to their budget constraints.

$$\max_{C_{H,k,t+s}, L_{H,k,t+s}} E_t \sum_{s=0}^{\infty} \beta^{t+s} U\Big(C_{H,k,t+s}, L_{H,k,t+s}\Big) \quad \forall \;\; k = C, R.$$

The budget constraints for each type of household are the following:

$$P_t C_{H,C,t} \leq (1-\tau_t) W_{H,t} L_{H,C,t} + \frac{(1-z)}{\lambda^H} \Upsilon_{H,t} - T_{C,t}, \tag{A95}$$

$$P_t C_{H,R,t} + B_{H,t+1} \leq (1-\tau_t) W_{H,t} L_{H,R,t} + (1+i_t) B_{H,t} + \frac{z}{1-\lambda^H} \Upsilon_{H,t} - T_{R,t}. \tag{A96}$$

P_t is the Home CPI (i.e. the cost of acquiring one unit of the home consumption bundle). B_{t+1} are holdings of nominal bonds that pay the national interest rate i_{t+1}. $W_{H,t}$ is the nominal wage rate in the Home region. $L_{H,k,t}$ is per-capita labor supply of households of type k. $\Upsilon_{H,t}$ are aggregate nominal profits from firms in the Home region. A share z of the aggregate profits is transferred to the Ricardian households such that the profits per capita are given by $\frac{z}{1-\lambda}\Upsilon_{H,t}$, and similarly for the *hand-to-mouth* households. $T_{H,k,t}$ are lump-sum taxes levied by the national government on household type k.

The Home final consumption good, $C_{H,k,t}$, is a CES bundle of Home and Foreign composite goods, $C^H_{H,k,t}$ and $C^F_{H,k,t}$:

$$C_{H,k,t} = \left[\phi_H^{\frac{1}{\eta}} {C^H_{H,k,t}}^{\frac{\eta-1}{\eta}} + (1-\phi_H)^{\frac{1}{\eta}} {C^F_{H,k,t}}^{\frac{\eta-1}{\eta}} \right]^{\frac{\eta}{\eta-1}} \quad \forall \;\; k = C, R. \tag{A97}$$

I use superscripts to denote the region where production takes place and subscripts to denote the region where consumption takes place (i.e. $C^F_{H,k,t}$ is a composite of goods produced in region F consumed by household type k located in region H). The parameter η governs the elasticity of substitution between domestic and foreign goods. Parameter ϕ_H captures the share of Home goods in total consumption and, hence, the extent of home bias. Both parameters govern the strength of expenditure-switching in response to changes in relative regional prices. Each regional composite good is a CES bundle of individual varieties.

$$\text{(A98)} \qquad C^H_{H,k,t} = \left[\int_0^1 c^H_{H,k,t}(z)^{\frac{\theta}{\theta-1}} dz \right]^{\frac{\theta-1}{\theta}}, \qquad C^F_{H,k,t} = \left[\int_0^1 c^F_{H,k,t}(z)^{\frac{\theta}{\theta-1}} dz \right]^{\frac{\theta-1}{\theta}},$$

where $c^H_{H,k,t}(z)$ and $c^F_{H,k,t}(z)$ denote consumption of Home and Foreign variety z, respectively. The parameter θ governs the elasticity of substitution between varieties produced within a given region. Goods trade freely across regions; therefore, both regions face the same prices for each individual variety. I denote prices at the variety level by $p^H_t(z)$ and $p^F_t(z)$, respectively. Price indexes are defined as follows.

$$\text{(A99)} \qquad P_{H,t} = \left[\int_0^1 p_{H,t}(z)^{1-\theta} dz \right]^{\frac{1}{1-\theta}}, \qquad P_{F,t} = \left[\int_0^1 p_{F,t}(z)^{1-\theta} dz \right]^{\frac{1}{1-\theta}},$$

and

$$\text{(A100)} \qquad P_t = \left[\phi^H P^{1-\eta}_{H,t} + (1-\phi^H) P^{1-\eta}_{F,t} \right]^{\frac{1}{1-\eta}}.$$

$P_{H,t}$ and $P_{F,t}$ are the PPI prices indexes of goods produced in the Home and Foreign region, respectively. As mentioned above, P_t is the CPI price index of the Home region. I assume that preferences are separable in labor and consumption. Concretely,

$$U\left(C_{H,k,t}, L_{H,k,t}\right) = \frac{C^{1-\sigma_H{}^{-1}}_{H,k,t}}{1-\sigma_H{}^{-1}} - \psi \frac{L^{1+\frac{1}{\nu}}_{H,k,t}}{1+\frac{1}{\nu}},$$

where σ_H is the intertemporal elasticity of substitution and ν is the Frisch elasticity of labor supply. Utility maximization by Ricardian households yields a standard Euler equation:

$$\text{(A101)} \qquad C^{-\sigma_H^{-1}}_{H,R,t} = \beta E_t \left[C^{-\sigma_H^{-1}}_{H,R,t+1} (1+i_t) \frac{P_t}{P_{t+1}} \right].$$

The intratemporal trade-off between labor and consumption is dictated by:

$$\text{(A102)} \qquad \frac{\psi L_{H,k,t}^{\nu^{-1}}}{C_{H,k,t}^{-\sigma_H^{-1}}} = (1-\tau_t)\frac{W_{H,t}}{P_t} \quad k = C, R.$$

Note that this trade-off is the same for both Ricardians and *hand-to-mouth* households.

Lastly, minimization of costs to achieve a given consumption level $C_{H,k,t}$ implies the following demand schedules:

$$\text{(A103)} \qquad C_{H,k,t}^{H} = \phi^{H} C_{H,k,t}\left(\frac{P_{H,t}}{P_t}\right)^{-\eta}, \qquad C_{H,k,t}^{F} = (1-\phi^{H}) C_{H,k,t}\left(\frac{P_{F,t}}{P_t}\right)^{-\eta},$$

$$\text{(A104)} \qquad c_{H,k,t}^{H}(z) = C_{H,k,t}^{H}\left(\frac{p_{H,t}(z)}{P_{H,t}}\right)^{-\theta}, \qquad c_{H,k,t}^{F}(z) = C_{H,k,t}^{F}\left(\frac{p_{F,t}(z)}{P_{F,t}}\right)^{-\theta},$$

for $k = C, R$.

D.2. Foreign Households

The problem for the Foreign region is symmetric, so I only present some key equations here. The share of *Hand-to-mouth* households in the Foreign region is denoted by λ^F. The final good bundle of the Foreign region, $C_{F,k,t}$ is given by:

$$\text{(A105)} \qquad C_{F,k,t} = \left[\phi_F^{\frac{1}{\eta}} C_{F,k,t}^{F\;\frac{\eta-1}{\eta}} + (1-\phi_F)^{\frac{1}{\eta}} C_{F,k,t}^{H\;\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}} \quad \forall \;\; k = C, R,$$

where ϕ_F governs the degree of home-bias in the Foreign region. The Foreign Euler equation for Ricardian households is:

$$\text{(A106)} \qquad C_{F,R,t}^{-\sigma_F^{-1}} = \beta E_t\left[C_{F,R,t+1}^{-\sigma_F^{-1}}(1+i_t)\frac{P_t^*}{P_{t+1}^*}\right],$$

where P_t^* is the CPI price level in the Foreign region given by:

$$\text{(A107)} \qquad P_t^* = \left[\phi^F P_{F,t}^{1-\eta} + (1-\phi^F) P_{H,t}^{1-\eta}\right]^{\frac{1}{1-\eta}}.$$

D.3. Risk-sharing condition

Complete markets and the Euler equations of the Home and Foreign region imply the following risk-sharing condition:

$$\frac{C_{F,R,t}^{-\sigma_F^{-1}}}{C_{H,R,t}^{-\sigma_H^{-1}}} = \omega_t Q_t, \tag{A108}$$

where $\omega_t = 1$ on a balanced growth path and $Q_t = \frac{P_t^*}{P_t}$ is the real exchange rate between regions in the monetary union. Note that the risk-sharing condition is in terms of Ricardian consumption only.

D.4. Household Aggregates

Per capita consumption and labor supply in the region r satisfy:

$$C_{r,t} = \lambda^r C_{r,C,t} + (1-\lambda^r)C_{r,R,t}, \qquad L_{r,t} = \lambda^r L_{r,C,t} + (1-\lambda^r)L_{r,R,t}, \tag{A109}$$

for $r = H, F$. Aggregate consumption and labor supply satisfy:

$$C_t = nC_{H,t} + (1-n)C_{F,t}, \qquad L_t = nL_{H,t} + (1-n)L_{F,t}. \tag{A110}$$

D.5. Production

The problem of Home and Foreign producers is symmetric, so I focus on Home firms. The economy features a continuum of monopolistically competitive producers of Home varieties, indexed by z, each operating the technology $y_t^H(z) = L_{H,t}^a(z)$. They take regional wages as given and set prices subject to a la Calvo stickiness. The demand schedule for firm z is given by:

$$y_t^H(z) = \left[nC_{H,t}^H + (1-n)C_{F,t}^H + nG_t^H\right]\left(\frac{p_{H,t}(z)}{P_{H,t}}\right)^{-\theta}, \tag{A111}$$

where $C_{H,t}^H$ and $C_{F,t}^H$ are the per-capita demands for Home goods coming from Home and Foreign households, respectively. G_t^H is per-capita government demand of the Home good.[46] Firms take regional wages, $W_{H,t}$, as given such that nominal marginal costs are:

$$MC_t^H(z) = \frac{W_{H,t}}{aL_{H,t}^{a-1}(z)}. \tag{A112}$$

[46]This formulation assumes that the government has the same CES preferences over individual varieties as households.

Firms get to reset prices with probability $1-\alpha$ every period (Calvo, 1983). They choose prices to maximize the discounted value of future real profits subject to the demand schedule in (A111). Concretely, the problem of a firm that gets to reset prices at t is:

$$\max_{p^*_{H,t}(z)} \sum_{s=0}^{\infty} \alpha^s \Lambda_{t+s|t} y^H_{t+s|t}(z) \left[p^*_{H,t}(z) - \frac{W_{H,t+s}}{a L^{a-1}_{H,t+s}(z)} \right] \qquad \text{s.t.} \quad \text{(A111)}\,,$$

where $\Lambda_{t+s|t}$ is the stochastic discount factor of the Ricardian household, $p^*_{H,t}(z)$ is the optimal reset price and $y^H_{t+s|t}(z)$ is the demand faced in period $t+s$ conditional on resetting prices in period t. The solution to this problem yields the following New-Keynesian Regional Phillips Curve (NK-RPC) for Home PPI inflation:

$$\hat{\pi}_{H,t} = \kappa_H \hat{mc}^H_t + \beta E_t[\hat{\pi}_{H,t+1}], \tag{A113}$$

where $\hat{\pi}_{H,t}$ is the deviation of the Home PPI inflation from a steady state of zero inflation, $\kappa_H = \frac{(1-\alpha)(1-\alpha\beta)}{\alpha}$ governs the slope of the Home NK-RPC, $\hat{mc}^H_t$ is the deviation of Home real marginal costs from the steady state marginal costs. Due to the symmetry of the problem, we can express the Foreign NK-RPC as:

$$\hat{\pi}_{F,t} = \kappa^F \hat{mc}^F_t + \beta E_t[\hat{\pi}_{F,t+1}], \tag{A114}$$

where κ_F and κ_h differ whenever the degree of price stickiness is heterogeneous between regions.

D.6. Government

The government consumes Home and Foreign varieties and raises revenues through lump-sum and income taxes on households. Government consumption follows an AR(1) process and its within-region composition mimics that of private consumption.[47] Per-capita total demand for Home and Foreign goods is:

$$G^H_t = (1-\rho_g) G_{ss} + \rho_g G^H_{t-1} + e^{g^H}_t + e^{G_h}_t, \tag{A115}$$

$$G^F_t = (1-\rho_g) G_{ss} + \rho_g G^F_{t-1} + e^{g^F}_t + e^{G_f}_t, \tag{A116}$$

[47] This implies that per-capita government spending in varieties from the Home and Foreign region are given by:

$$g^H_t(z) = G^H_t \left(\frac{p_{H,t}(z)}{P_{H,t}} \right)^{-\theta}, \qquad g^F_t(z) = G^F_t \left(\frac{p_{F,t}(z)}{P_{F,t}} \right)^{-\theta}.$$

where G_{ss} is the steady-state participation of government spending in output, which I assume is common between regions. The parameter ρ_g governs the persistence of fiscal shocks. Regional government spending is subject to two different idiosyncratic shocks: $e_t^{g^k}$ and $e_t^{G_k}$ for $k = H, F$. The first one is a purely region-specific shock as in Nakamura and Steinsson (2014). The process for the second one, $e_t^{G_k}$, is as follows:

$$e_t^{G_H} = \frac{s_H}{n}\epsilon_t^G, \qquad e_t^{G_F} = \frac{1-s_H}{1-n}\epsilon_t^G, \tag{A117}$$

where ϵ_t^G is an aggregate fiscal shock that loads differently between regions. The per capita loadings are given by $\frac{s_H}{n}$ and $\frac{1-s_H}{1-n}$, respectively. The case where $s_h = n$ represents a government spending shock that hits all regions with equal intensity. The structure of this new shock mimics the Bartik-type instruments that are encountered in the local multipliers literature and results in a closer map between the model and the data.[48] Lastly, total government spending G_t satisfies:

$$G_t = nG_t^H + (1-n)G_t^F. \tag{A118}$$

The government follows a balanced budget where

$$P_tG_t = T_t + \tau_t\Big(nW_{H,t}L_{H,t} + (1-n)W_{F,t}L_{F,t}\Big), \tag{A119}$$

where T_t are the total lump sum taxes and P_t the price index of government purchases.

D.7. Monetary Policy

The monetary authority sets the nominal interest rate following a Taylor rule:

$$i_t = \rho_i i_{t-1} + (1-\rho_i)\Big(i_{ss} + \phi_\pi \Pi_t^{agg} + \phi_y \hat{y}_t^{agg}\Big) + u_{it}, \tag{A120}$$

where $\hat{x}_t$ denotes the deviation of variable x from the zero inflation steady state. Π_t^{agg} is aggregate inflation, defined as $\Pi_t^{agg} = n\Pi_t + (1-n)\Pi_t^*$. Π_t and Π_t^* are the CPI inflation rates in the Home and Foreign region, respectively. The national output gap is defined by $\hat{y}_t^{agg} = n\hat{y}_t^H + (1-n)\hat{y}_t^F$.

[48] The reason is that in a setting where aggregate policies are not completely *differenced-out*, the size of the aggregate response becomes relevant. It therefore makes a difference whether we study a fiscal shock to a marginal region versus an aggregate and sizable increase in government spending that loads differently across regions. The aggregate responses to these two types of shocks will be different. In the extreme, a shock to an infinitesimal region would trigger an infinitesimal aggregate response, so whether this is differenced out or not will likely not matter much.

D.8. Market Clearing

Per capita Home and Foreign output are

$$Y_t^H = \frac{1}{n}\int_0^1 y_t^H(z)dz, \qquad Y_t^F = \frac{1}{1-n}\int_0^1 y_t^F(z)dz. \tag{A121}$$

Then, the total output is

$$Y_t = nY_t^H + (1-n)Y_t^F. \tag{A122}$$

Clearing of the goods market requires the following.

$$nY_t^H = nC_{H,t}^H + (1-n)C_{F,t}^H + nG_t^H, \tag{A123}$$

$$(1-n)Y_t^F = nC_{H,t}^F + (1-n)C_{F,t}^F + (1-n)G_t^F, \tag{A124}$$

where all variables refer to amounts per capita.

D.9. Calibration

Table A9 summarizes the calibration. I stick to the calibration in Nakamura and Steinsson (2014) for most parameters. The first part of the table presents all parameters that are the same in both papers. In the second part of the table, I show parameter choices that differ from Nakamura and Steinsson (2014) or are newly introduced in this paper. The baseline calibration sets $\lambda_k = 0$ and generates heterogeneous interest rate sensitivities through differences in the IES of Ricardian households. These are set to $\sigma_H = 1.25$ and $\sigma_F = .75$. The parameter s_H, which governs the degree of Home bias in the aggregate government spending shock, ϵ_t^G, is set to .8. This implies that the Home region is relatively more exposed to government spending shocks.

TABLE A9. Baseline Calibration

A. First Set of Parameters		
β	Discount factor	.99
ν	Frisch Elasticity of Labor Supply	1
η	Elasticity of Substitution between Home and Foreign Goods	2
θ	Elasticity of Substitution between individual varieties	7
a	Labor Coefficient in Production Function	.63
i_{ss}	Steady-state Nominal Rate	.01
G_{ss}	Participation of Government Spending in Steady State	.2
τ	Income Tax Rate	0
ϕ_H	Home Bias in Home Region	.85
α	Degree of Price Stickiness	.75
ρ_i	Persistence of Monetary Policy Rule	.8
ϕ_π	Inflation Coefficient in Taylor Rule	1.5
ϕ_y	Output Coefficient in Taylor Rule	.1
B. Second set of Parameters		
n	Size of Home Region	.5
z	Ricardian Participation in Profits	1
λ_r	Share of *Hand-to-Mouth* in Region r	0
σ_H (σ_F)	IES of Ricardian Households in Home (Foreign) region	1.25 (.75)
s_H	Home Bias of Aggregate Government Spending Shock	.8
ρ_g	Persistence of Government Spending	.85

Back to DSGE Monte Carlo exercise in forward-looking decomposition framework.
Back to Subsection 5.2.
Back to Subsection 5.4 - matching model and data.