

# Identification and Estimation of Causal Effects in High-Frequency Event Studies\*

ALESSANDRO CASINI<sup>†</sup>

University of Rome Tor Vergata

ADAM McCLOSKEY<sup>‡</sup>

University of Colorado at Boulder

31st October 2024

## Abstract

We provide precise conditions for nonparametric identification of causal effects by high-frequency event study regressions, which have been used widely in the recent macroeconomics, financial economics and political economy literatures. The high-frequency event study method regresses changes in an outcome variable on a measure of unexpected changes in a policy variable in a narrow time window around an event or a policy announcement (e.g., a 30-minute window around an FOMC announcement). We show that, contrary to popular belief, the narrow size of the window is not sufficient for identification. Rather, the population regression coefficient identifies a causal estimand when (i) the effect of the policy shock on the outcome does not depend on the other variables (separability) and (ii) the surprise component of the news or event dominates all other variables that are present in the event window (*relative exogeneity*). Technically, the latter condition requires the ratio between the variance of the policy shock and that of the other variables to be infinite in the event window. Under these conditions, we establish the causal meaning of the event study estimand corresponding to the regression coefficient and the consistency and asymptotic normality of the event study estimator. Notably, this standard linear regression estimator is robust to general forms of nonlinearity. We apply our results to Nakamura and Steinsson's (2018a) analysis of the real economic effects of monetary policy, providing a simple empirical procedure to analyze the extent to which the standard event study estimator adequately estimates causal effects of interest.

**JEL Classification:** C32, C51

**Keywords:** Causal effects, Event study, High-frequency data, Identification.

---

\*We thank Emi Nakamura, Mikkel Plagborg-Møller and Jón Steinsson for useful comments. The replication code is available on our websites.

<sup>†</sup>Dep. of Economics and Finance, University of Rome Tor Vergata, Via Columbia 2, Rome, 00133, IT. Email: [alessandro.casini@uniroma2.it](mailto:alessandro.casini@uniroma2.it).

<sup>‡</sup>Dep. of Economics, University of Colorado at Boulder, 256 UCB, Boulder, CO 80309, US. Email: [adam.mccloskey@colorado.edu](mailto:adam.mccloskey@colorado.edu).

# 1 Introduction

Randomized controlled experiments offer an ideal framework for identifying causal effects. However, in macroeconomics and finance controlled experiments cannot be run in practice. Instead, economists often search for pseudo-experiments, that is situations in which one can extract plausible exogenous variation in policy and use this variation to estimate the effect of the policy on some economic outcome [cf. [Nakamura and Steinsson \(2018b\)](#)]. Recently, there has been a surge of interest in high-frequency event study regressions for estimating causal effects in applied work in macroeconomics, financial economics and political economy, among others. The idea behind the event study approach based on high-frequency data is that in a narrow time window around a policy announcement or data release [e.g., a Federal Open Market Committee (FOMC) announcement, a U.S. employment report released by the Bureau of Labor Statistics, a GDP release report by the Bureau of Economic Analysis, etc.], one can extract the unexpected change or surprise in the policy and regress the changes in an outcome variable within the narrow window on the policy surprises to estimate the causal effect of the policy.

Though this approach has become central to empirical work, there are no corresponding theoretical results establishing identification of causal effects via this method, and informal identification arguments used by different authors do not always coincide.<sup>1</sup> In this paper we establish precise conditions for nonparametric identification of causal effects by high-frequency event study regressions. We show that, contrary to popular belief, the narrow size of the time window the event study regression is run over is not sufficient for the identification of causal effects. Rather, the population regression coefficient identifies a causal estimand when (i) the effect of the policy shock on the outcome does not depend on the other variables (separability) and (ii) the surprise component of the news or event dominates all other variables that are present in the event window (*relative exogeneity*). Under these conditions, we establish the causal meaning of the event study estimand corresponding to the regression coefficient and the consistency and asymptotic normality of the event study regression estimator. Notably, this standard linear regression estimator is robust to general forms of nonlinearity (e.g., between the outcome and policy variables).

Separability holds, for example, when the model is linear, which is often assumed in applied work. The key condition deserving of careful scrutiny in practice, relative exogeneity, holds when the policy shock has infinite variance while the other variables have finite variance within the

---

<sup>1</sup>Recent examples of informal identification arguments include [Bauer and Swanson \(2023a\)](#) and [Nakamura and Steinsson \(2018a\)](#), who discussed the exogeneity and relevance of monetary policy surprises typically constructed around FOMC announcements for identifying the macroeconomic effects of monetary policy shocks.

event window. More precisely, relative exogeneity holds when the ratio between the variance of the policy shock and that of the other variables (i.e., background noise) is infinite in the event window. Thus, relative exogeneity also holds when the variance of the policy shock is finite while the variance of the background noise is vanishing. The latter variables correspond to factors that are not specific to the announcement and may also be present in non-announcement periods. In contrast, the policy shock occurs in a lumpy manner as the unexpected part of the news quickly spreads among economic agents. It is this lumpy manner in which a disproportionate amount of policy news is revealed that can justify relative exogeneity. Even when the policy shock does not have infinite variance, which can be difficult to verify in practice, we show that the event study estimator has low bias for a weighted average of causal effects when the variance of the policy shock is large enough relative to that of other variables in the window. In this sense, relative exogeneity can be seen as an idealized limiting case that can serve as a good approximation to the practically-relevant case of a very large variance ratio for the policy shock relative to the other variables in the window.

Relative exogeneity relates to the size of the event window. As the size of the window expands, it becomes less likely that the policy shock dominates all other variables within the window, making it more likely that relative exogeneity provides a poor approximation. Relative exogeneity can also fail if there is information leakage about the policy news or some market anticipation of the policy change. Information leakage or frictions result in a reduction in the variance of the policy shock relative to the other variables in the window. The finance literature has recently documented strong evidence of information leakage, informal communication and informed trading around policy announcements [see, e.g., [Cieslak, Morse, and Vissing-Jorgensen \(2019\)](#), [Cieslak and Schrimpf \(2019\)](#) and [Lucca and Moench \(2015\)](#)]. However, if the extent of the information leakage is small, then it will lead only to partial market anticipation and will not prevent the news from coming out in a lumpy manner at the release time, thereby allowing relative exogeneity to hold.

When separability and relative exogeneity hold we show that (i) any reverse causality from the outcome variable to the policy variable does not generate bias since the reverse causality is dominated by the policy shock and (ii) common unobserved factors correlated with the policy variable do not generate omitted variables bias since they are also dominated by the policy shock. Since the event study method regresses changes in an outcome on the unexpected changes in a policy variable at dates in the policy sample, the event study estimand is equal to the corresponding population regression coefficient. We establish that the latter identifies a weighted average

of standardized marginal causal effects (MCEs) of the policy on the outcome.<sup>2</sup> When relative exogeneity is violated, we show that the event study estimand can be decomposed into the same weighted average of MCEs and a selection bias factor. The selection bias factor is decreasing in the variance of the policy shock so that even when relative exogeneity fails, the event study estimator will not have substantial bias so long as the variance of the policy shock is relatively large.

It is difficult to use statistical tests to verify relative exogeneity. Existing tests for (in)finite variance [e.g., [Trapani \(2016\)](#)] require a large sample of the corresponding random variable to be observed or residuals from a correctly-specified regression. Given the potential for simultaneity and omitted variables bias in high-frequency event study regressions, these requirements are not satisfied in the current context. In addition, the OLS event study estimator can still perform well when relative exogeneity holds approximately. A test for infinite variance of the policy shocks alone, even when feasible, is unable to detect when relative exogeneity provides a good enough approximation for the event study estimator to perform well in practice. Instead of trying to test for infinite variance, we introduce a simple empirical procedure that can be used for a sensitivity analysis to diagnose whether relative exogeneity is “close enough” to holding that the OLS event study estimator should be expected to have mean-squared error at least as small as that of an oracle estimator in a corresponding regression with no endogeneity. This sensitivity analysis uses information available in the relevant control sample, e.g., days when there is no FOMC meeting, in order to obtain a proxy for the variance of the variables that are not specific to the policy within the event window. This procedure requires no additional data relative to that already used by existing high-frequency event studies and can thus be applied easily. We introduce our procedure in the context of an empirical example aimed at assessing the causal effects of monetary policy news on real interest rates. The empirical results show that relative exogeneity is likely to approximately hold in the analysis of [Nakamura and Steinsson \(2018a\)](#) based on a 30-minute or 1-day window for the outcome and policy variable.

Finally, we use our identification framework to shed light on the recent debate in the literature on some puzzling event study regression results involving the causal effect of monetary policy on Blue Chip forecasts of real GDP. We show that the estimates with signs opposite to the predictions from standard macroeconomic models likely arise because relative exogeneity fails in this type of regression. Since the Blue Chip forecast revision is constructed as a one-month change while the monetary policy surprise is constructed as a 30-minute change, it is highly unlikely that the

---

<sup>2</sup>The MCE is the derivative of the potential outcome with respect to the policy variable. This should not be confused with the marginal treatment effect (MTE) used in microeconometrics which refers to the treatment effect of a program or intervention for those at the margin of participation [cf. [Heckman and Vytlacil \(2001\)](#)].

variance of the policy shocks dominates the variance of the other variables that determine the Blue Chip forecast: the latter aggregate news over a much longer time frame. Thus, any endogeneity of the policy surprise is not likely to be drowned out by the variation in the policy shock.

The rest of the paper is structured as follows. In Section 2 we conduct a review of the high-frequency event study literature and detail a few empirical examples, present our identification results and discuss robustness to information leakage. Section 3 presents the asymptotic properties of the OLS event study estimator. Section 4 proposes a sensitivity analysis to verify the identification conditions in the context of an empirical example and discusses some identification issues in event study regressions that involve Blue Chip forecasts. We conclude in Section 5.

## 2 Identification in High-Frequency Event Studies

We begin with a brief review of the event study methodology in empirical work in Section 2.1. We then introduce the potential outcomes framework and provide formal conditions for nonparametric identification of causal effects via the event study regressions in Section 2.2. We present the identification results in Section 2.3 and relate them to the literature in Section 2.4. Robustness to information leakage is discussed in Section 2.5.

### 2.1 Event Study Design

Consider a system of dynamic simultaneous equations that relate an outcome variable  $Y_t$  and a (measure of) policy action  $D_t$  to each other:

$$\begin{aligned} Y_t &= \beta D_t + X_t' \theta + Z_t' \gamma_1 + u_t, \\ D_t &= \alpha Y_t + X_t' \phi + Z_t' \gamma_2 + e_t, \end{aligned} \tag{2.1}$$

where  $X_t$  represents observed macroeconomic variables that might influence both the outcome and policy variable,  $Z_t$  represents unobserved macroeconomic factors that also might affect the outcome and policy variable and  $u_t$  and  $e_t$  are serially uncorrelated shocks that are mutually uncorrelated. The parameters  $\beta$  and  $\alpha$  are scalars while  $\theta$ ,  $\phi$ ,  $\gamma_1$  and  $\gamma_2$  are finite-dimensional vectors.  $X_t$  is a vector that could include lags of  $Y_t$  and  $D_t$ . The parameter of interest is  $\beta$  which captures the response of the outcome variable to the policy action. The policy variable is endogenous, i.e.,  $D_t$  reacts to  $Y_t$ . Further, since  $Z_t$  is not observed there is an omitted variables problem as  $Z_t$  and  $D_t$  may be correlated. Hence, the identification of  $\beta$  requires one to overcome both simultaneity and

an omitted variables problem.

The event study approach based on high-frequency observations (e.g., weekly, daily and intradaily frequencies) of  $Y_t$  and  $D_t$  can be used to address these identification issues. The idea is that in a narrow time window around a policy announcement or data release (e.g., FOMC announcement, U.S. employment report, GDP data release, etc.) one can extract the unexpected change (or surprise) in the policy action to form  $D_t$ . The sample consists of observations of  $Y_t$  and  $D_t$  at the dates corresponding to the relevant policy announcement, data release or event. We call this sample the policy sample and denote it by  $\mathbf{P}$ . Under the identification conditions to be discussed below, a simple OLS regression of  $Y_t$  on  $D_t$  over the policy sample (i.e., over all  $t \in \mathbf{P}$ ) recovers causal effects of the policy action on the outcome.

We now present a few examples that use the model (2.1).

**Example 1.** [FOMC announcements: Bauer and Swanson (2023a, 2023b), Kuttner (2001), Nakamura and Steinsson (2018a) and Rigobon and Sack (2004)] Several authors investigated the impact of monetary policy on the real economy using the event study approach.<sup>3</sup> Kuttner (2001) explained how to use Federal funds futures contracts to separate changes in the Fed funds rate (i.e., the short-term interest rate) into anticipated and unanticipated monetary policy actions, the latter being  $D_t$ . In Kuttner (2001)  $D_t$  is the 1-day change in the spot-month Federal funds future rate and  $Y_t$  represents a yield on a zero-coupon Treasury bill (or bond) at some maturity or a change in an asset price. The policy sample  $\mathbf{P}$  collects the dates of the FOMC announcements and the dates when the Fed funds target rate was changed (if that did not coincide with an FOMC announcement).

This analysis was further elaborated by, among others, Nakamura and Steinsson (2018a) who used intradaily data and looked at a 30-minute window surrounding each FOMC announcement. The authors considered 30-minute changes in the zero-coupon yields and instantaneous forward rates constructed using Treasury Inflation Protected Security data at different maturities and changes in survey expectations on output and inflation, as the outcome variable  $Y_t$ . To construct the monetary policy news  $D_t$  the authors extracted the first principle component of the unanticipated change over 30-minute windows of five interest rates chosen among the Federal funds futures and eurodollar futures. The latter provide a direct measure of the unexpected component of the policy change. They estimated the causal effect of  $D_t$  on  $Y_t$  by running an event study OLS

---

<sup>3</sup>See Ai and Bansal (2018), Cochrane and Piazzesi (2002), Cook and Hahn (1989), Gürkaynak, Sack, and Swanson (2005), Lucca and Moench (2015), Bernile, Hu, and Tang (2016), Hu, Pan, Wang, and Zhu (2022), Caballero and Simsek (2022, 2023), Cieslak, Morse, and Vissing-Jorgensen (2019), Cieslak and McMahon (2023), Cieslak and Schrimpf (2019), Hansen, McMahon, and Prat (2018), Hanson and Stein (2015), Jarociński and Karadi (2020), Michelacci and Paciello (2020), Neuhierl and Weber (2019) and Swanson (2021).

regression of  $Y_t$  on  $D_t$  for dates in the policy sample:

$$Y_t = \beta D_t + \tilde{u}_t, \quad t \in \mathbf{P}, \quad (2.2)$$

where  $\tilde{u}_t$  is an error term. The control variables  $X_t$  that could be included in this analysis are monthly releases of major macroeconomic variables or other low-frequency variables. For example, the consumer price index (CPI), nonfarm payrolls, producer price index (PPI), retail sales, etc. However, in practice event study regressions do not typically involve control variables.<sup>4</sup>

It is not obvious how to justify the exclusion of either the observed or unobserved factors  $X_t$  and  $Z_t$  from (2.2). Asset prices likely react to  $X_t$  and  $Z_t$  even within the event window, no matter how small the window is. If not, a recursive argument soon would contradict asset pricing theory. Think about splitting the regular trading hours for the U.S. stock market, 9:30am-4:00pm, into non-overlapping small time windows of, for example, 30 minutes. If one assumes that asset prices do not respond to observed or unobserved macroeconomic factors over such a tight window, then applying this argument recursively to each trading day implies that asset prices never respond to such factors. That is, the choice of a very tight window bracketing a policy announcement is not a sufficient condition for precluding omitted variables bias from the regression (2.2) due to the correlation of  $D_t$  with  $X_t$  and/or  $Z_t$ . We will show that under our identification conditions, which do not refer explicitly to the size of the window, the OLS event study regression (2.2) that excludes both the observed and unobserved factors  $X_t$  and  $Z_t$  can still recover causal effects of interest.

The recent literature documented evidence pointing to simultaneous determination of  $D_t$  and  $Y_t$  in FOMC announcement applications. See Section 2.4 for details. We will also show that under our identification conditions, the OLS event study regression (2.2) recovers causal effects of interest in the presence of simultaneity.

**Example 2.** [Macroeconomic announcements: Faust, Rogers, Wang, and Wright (2007), Gürkaynak, Kısacikoğlu, and Wright (2020) and Gürkaynak, Sack, and Swanson (2005)] Several works studied the effects of macroeconomic announcements on changes in asset prices and exchange rates in a narrow window around news releases such as the U.S. employment report, U.S. GDP releases, Census reports on retail sales, and CPI and PPI data releases. Gürkaynak, Sack, and Swanson (2005) also estimated the event study regression (2.2) but with a vector-valued  $D_t$  and coefficient  $\beta$ , where  $Y_t$  is the log-return on an asset or a change in a bond yield and  $D_t$  is a vector of news, or unexpected, components of the considered macroeconomic announcements and  $u_t$  is a serially

---

<sup>4</sup>See Bauer and Swanson (2023a) and Rigobon and Sack (2003) for exceptions.

uncorrelated error term.<sup>5</sup> Unlike for the FOMC announcements, there are no traded instruments from which to infer market expectations for macroeconomic announcements. Thus, in order to identify surprise announcements the literature relies instead on economists' forecasts from surveys: each element of  $D_t$  is computed as the difference between the actual macroeconomic data release and its market expectation obtained from the most recent survey. These surveys are the Blue Chip Economic Indicators Survey or those run by Action Economics, or alternatively by Bloomberg. Thus,  $D_t$  captures the surprise component of each data release.

Reverse causality from  $Y_t$  to  $D_t$  is ruled out by the authors' identification argument that in a 20-minute window around news releases, changes in asset prices do not cause news. For example, the employment report that is released on the first Friday of each month pertains to the labor market data in the previous month. Thus, by construction changes in asset prices within the window affect neither the news release nor the survey expectations since the latter are collected earlier.<sup>6</sup>

The exclusion of the unobserved factors  $Z_t$  that can affect both asset prices and the unexpected component of the news is easier to justify than in the case of FOMC announcements. Since  $D_t$  here involves economic data releases about the previous month and survey expectations are collected in the days preceding the event window, correlation between  $Z_t$  and  $D_t$  seems unlikely. However, economic and business news and other events that occur in the hours or days before the macroeconomic data release may require time to incorporate into financial markets and may be subject to uncertainty with respect to how they are interpreted by other market participants. Therefore,  $Z_t$  may contain information about other news and events (e.g., company-specific news) that occur before the announcement window and therefore influence survey expectations, and in turn  $D_t$ , generating correlation between  $Z_t$  and  $D_t$ . This issue is independent from the size of the window used in the event study regression.

**Example 3.** [Political and war announcements: [Acemoglu, Hassan, and Tahoun \(2018\)](#), [Dube, Kaplan, and Naidu \(2011\)](#), [Garred, Stickland, and Warrinnier \(2023\)](#) and [Guidolin and La Ferrara \(2007\)](#)]. We briefly note that the high-frequency event study methodology is also applied in the field of political economy. [Guidolin and La Ferrara \(2007\)](#) provided evidence that violent conflict may be perceived by investors as beneficial to incumbent firms. They focused on the Angolan civil war and its effects on the industry of diamond production. They exploited the sudden ceasefire of the civil war after the announcement of the death of the rebels' leader, Jonas Savimbi, on February 22, 2002, showing that international stock markets perceived Savimbi's death as bad news for the

---

<sup>5</sup>Actually, [Gürkaynak, Sack, and Swanson \(2005\)](#) considered a vector of asset prices  $Y_t$ . But since the cross-sectional variation is not exploited for identification we simply take  $Y_t$  to be a scalar.

<sup>6</sup>The amount of time between when the survey is given and the release of the news is typically less than a week.



companies in the diamond industry operating in Angola.

## 2.2 Nonparametric Event Study Design and Identification Conditions

Although we discussed examples in the high-frequency event study literature in the context of the linear model (2.1) in the previous section, our identification results for event study regressions actually apply to a much more general nonparametric class of simultaneous equations models.<sup>7</sup> Unlike the linear model (2.1), this more general class of models allows for general forms of nonlinear relationships between the variables in the system. Specifically, let  $Z_t$  denote an unobserved random vector and  $u_t$  and  $e_t$  denote scalar-valued random shocks to the outcome  $Y_t$  and policy  $D_t$ , respectively. When the observation belongs to the policy sample  $\mathbf{P}$ , the outcome variable  $Y_t$  is an unknown nonparametric function of the policy variable  $D_t$ ,  $Z_t$ ,  $u_t$  and the time index  $t$  while the policy variable  $D_t$  is simultaneously an unknown nonparametric function of  $Y_t$ ,  $Z_t$ ,  $e_t$  and  $t$ :

$$Y_t = \varphi_Y(D_t, Z_t, u_t, t) \quad \text{and} \quad D_t = \varphi_D(Y_t, Z_t, e_t, t), \quad (2.3)$$

for all  $t \in \mathbf{P}$ . The unknown structural functions  $\varphi_Y$  and  $\varphi_D$  may be nonlinear in their arguments, allow for simultaneous causality of  $Y_t$  and  $D_t$  and, since  $Z_t$  is unobserved, allow for omitted variables in the system determining  $Y_t$  and  $D_t$ .

In Example 1,  $D_t$  depends on the Federal Reserve's best estimates of the strength of the economy in the near-term and of potential inflationary pressures. Those estimates will be influenced by the macroeconomic factors  $Z_t$  and by changes in asset prices  $Y_t$ . The monetary policy shock  $e_t$  represents shifts in the preferences of individual FOMC members or in the manner in which their views are aggregated. For example, it could include changes to policy makers' goals and beliefs about the economy, political factors, and the temporary pursuit of objectives other than changes in the outcomes of interest (e.g., targeting inflation rather than unemployment or exchange rates). The shock to the outcome equation,  $u_t$ , captures any change in  $Y_t$  not attributable to the common macroeconomic factors  $Z_t$  and policy action,  $D_t$ . In Example 1,  $u_t$  is referred to as the asset price shock and is primarily driven by shifts in investors' risk preferences.

To establish the identification of causal effects we impose some structure on the system, in particular, a partial additive separability of the structural function for the outcome variable.

---

<sup>7</sup>To simplify the exposition, we present the identification conditions and results for the case when there are no control or pre-treatment variables  $X_t$  included in the design since they do not play an important role for the identification analysis and are not typically included in the high-frequency event study regressions, in which case they can be considered elements of  $Z_t$ .

**Assumption 1.** (*Structural form separability*) For all  $t \in \mathbf{P}$ ,  $Y_t = \varphi_Y(D_t, Z_t, u_t, t) = \varphi_{Y,D}(D_t, t) + \varphi_{Y,u}(Z_t, u_t, t)$  for some functions  $\varphi_Y$ ,  $\varphi_{Y,D}$  and  $\varphi_{Y,u}$ .

If some smoothness on  $\varphi_{Y,D}(\cdot, \cdot)$  is assumed, then Assumption 1 implies the following restrictions on partial effects:

$$\frac{\partial^2 Y_t}{\partial D_t \partial u_t} = 0 \quad \text{and} \quad \frac{\partial^2 Y_t}{\partial D_t \partial Z_t} = 0.$$

This means that the marginal effect of the policy on the outcome variable does not depend on the shock to the outcome variable or the unobserved factors  $Z_t$ . Without this additional smoothness, Assumption 1 implies that the effect of the policy on the outcome does not vary with  $Z_t$  or  $u_t$ .

Next, we assume that the simultaneous structural equations imply a reduced-form for  $D_t$  that is analogously additively separable so that the effect of  $e_t$  on  $D_t$  is not influenced by  $Z_t$  or  $u_t$ .

**Assumption 2.** (*Reduced-form separability*) For all  $t \in \mathbf{P}$ ,  $D_t = g_{D,e}(e_t, t) + g_{D,u}(Z_t, u_t, t)$  for some functions  $g_{D,e}$  and  $g_{D,u}$ .

Assumption 2 is implied by Assumption 1, separability in  $\varphi_D(Y_t, Z_t, e_t, t)$  across  $(Y_t, Z_t)$  and  $e_t$  and standard invertibility requirements on the system of equations. This holds trivially if the true data-generating process is a system of linear simultaneous equations, like (2.1) under standard rank conditions. In structural VARs, invertibility is a standard condition that is imposed to obtain the reduced-form shocks from the structural shocks [see Plagborg-Møller (2019) for a discussion about invertibility of impulse responses]. We directly impose Assumption 2 on the reduced-form rather than imposing invertibility requirements to avoid introducing further notation for defining the partial inverse of a multivariate function. Assumptions 1-2 are satisfied in the event study applications considered in empirical work since the specification for the outcome and treatment variables is typically a simultaneous linear equations model as in (2.1).

**Assumption 3.** (*Structural shocks*) For all  $t \in \mathbf{P}$ ,  $u_t$  and  $e_t$  have zero mean and no serial correlation, are mutually independent, and are each independent from  $Z_t$ .

Under Assumption 3,  $u_t$  and  $e_t$  are interpreted as structural shocks, i.e., primitive, unanticipated impulses that are unforecastable and mutually uncorrelated.

In the assumed specification (2.3),  $Y_t$  and  $D_t$  are determined simultaneously and so are endogenous. Often endogeneity is overcome by using instrumental variables. This leads to the identification of the local average treatment effect via the instrumental variables estimand [cf. Imbens and Angrist (1994)]. In contrast, the event study approach can identify causal effects in the

presence of endogeneity without the need to find instrumental variables.<sup>8</sup> The key idea is that in a narrow time window around a particular event (or change in policy) the variation in the policy variable is dominated by the variation in the policy shock. The effect of the outcome variable on the policy variable is still present within the time window but it is negligible relative to the effect of the policy shock. Similarly, the effect of the omitted variables  $Z_t$  on  $D_t$  is negligible relative to that of  $e_t$ . This idea is formalized by the following assumption.

**Assumption 4.** (*Relative exogeneity*) For all  $t \in \mathbf{P}$ ,

- (i)  $\sigma_{e,t}^2 = \text{Var}(e_t) \rightarrow \infty$ ,
- (ii)  $\text{Var}(g_{D,e}(e_t, t))$  is increasing in  $\sigma_{e,t}^2$ ,
- (iii)  $\mathbb{E}(\varphi_{Y,u}(Z_t, u_t, t)^2)$  and  $\mathbb{E}(g_{D,u}(Z_t, u_t, t)^2)$  are finite.

Assumption 4 requires that within the event window the policy shock has infinite variance (condition (i,ii)) while the other variables have finite variance (condition (iii)) and so the policy shock dominates the changes in the policy variable in the window. Assumption 4(iii) is implied by a correspondence between the boundness of the second moments of  $e_t$ ,  $Z_t$  and  $u_t$  with those of  $D_t$  and  $Y_t$ . This easily holds for (2.1) since the second moments of  $Y_t$  and  $D_t$  depend on the second moments of  $e_t$ ,  $Z_t$  and  $u_t$  and on the moments of their products. Altogether, Assumption 4 implies that at periods immediately surrounding a policy announcement (i.e.,  $t \in \mathbf{P}$ ) the policy shock  $e_t$  dominates the other shock  $u_t$  and the omitted factors  $Z_t$ . The latter involve shocks and factors that are not related to the announcement and are also present in non-announcement periods. Hence, it is reasonable to expect these variables to have finite variance. In contrast, the policy news shock is much more pronounced in the announcement window and occurs in a lumpy manner as it is completely unexpected. Consequently, the market reacts leading to realized volatility and trading volume to significantly decline before the announcement and then jump at the announcement [see, e.g., Lucca and Moench (2015) and Hu, Pan, Wang, and Zhu (2022)]. It is this lumpy manner in which a disproportionate amount of policy news is revealed that can make the policy variable  $D_t$  *relatively exogenous*. When this occurs we show that within the event window: (i) the reverse causality problem disappears (changes in the outcome variable do not affect changes in the policy variable since the latter are entirely driven by the policy shock);<sup>9</sup> (ii) the common unobserved

---

<sup>8</sup>It is well-known that it is hard to find valid instruments in macroeconomics applications. Considering Example 1, it is difficult to find any instrument that would affect asset price returns ( $Y_t$ ) without changing short-term interest rates ( $D_t$ ) as any variable related to the macroeconomic outlook would not satisfy this criterion. Neither would variables related to corporate revenues and profits since they would likely contain information about the economic outlook and be correlated with interest rate changes.

<sup>9</sup>The same conclusion holds when observable factors are included in the analysis since these are typically low-frequency variables that are dominated by the policy shock  $e_t$ .

factors  $Z_t$  do not generate omitted variables bias. Note that the endogeneity problem is overcome only in the policy sample  $\mathbf{P}$ . In the control sample  $\mathbf{C}$ , defined as the collection of all  $t$  such that  $t \notin \mathbf{P}$  (e.g., days with no FOMC announcement), the endogeneity remains.

Based on Assumption 4, we establish our identification results by taking the limits as

$$\sigma_{e,t}^2 \rightarrow \infty, \quad (2.4)$$

and refer to this limiting case as *relative exogeneity*.<sup>10</sup> The condition (2.4) makes clear one point that has been overlooked by the empirical literature. Namely, it requires that the “large” variance condition for the policy shock has to hold for all  $t \in \mathbf{P}$ . For example, suppose that it is satisfied only for a few announcements in the policy sample. Since the variance cannot be negative, an estimate of the average variance computed in the policy sample could still be very large. This could be misleadingly interpreted as support for relative exogeneity. However, the endogeneity in most of the policy sample would not be overcome and relative exogeneity would fail. Hence, verifying that the sample variance of  $D_t$  is large does not allow one to conclude that relative exogeneity holds.

### 2.2.1 Potential Outcomes Framework

To conduct our identification analysis, we introduce the relevant potential outcomes framework, which is useful for determining nonparametric conditions under which the OLS event study estimands have a causal interpretation. We define causal effects using the notion of potential outcomes introduced by Rubin (1974) and extended to time series settings by Angrist and Kuersteiner (2011) and Rambachan and Shephard (2021). Potential outcomes are defined as the counterfactuals of  $Y_t$  that would arise in response to a hypothetical value of the policy variable  $D_t$ .

**Definition 1.** The potential outcome,  $Y_t(d)$ , is defined as the value taken by  $Y_t$  if  $D_t = d$ .

We assume that  $d \in \mathbf{D}$  for an appropriate set  $\mathbf{D}$ . A potential outcome  $Y_t(d)$  describes which value the outcome would have taken at time  $t$  under treatment value  $d$ . The definition implies that the potential outcome  $Y_t(d)$  does not depend on future treatments. Rambachan and Shephard (2021) described this property as “non-anticipating potential outcomes”. The definition also implies that the potential outcome  $Y_t(d)$  does not depend on past treatments (unless they are elements of  $Z_t$ ). This is realistic in the high-frequency event study setting for two reasons. First,  $Y_t$

---

<sup>10</sup>More precisely, relative exogeneity holds when the ratios between the variance of the policy shock and the variances of the other variables in the event window diverge. Thus, one may instead frame relative exogeneity as the condition for which the variance of the policy shock is finite and the variance of the background noise vanishes in the event window. The same identification results can also be shown to hold under this latter condition.

typically measures a change in an asset price or a survey forecast within a narrow window around a policy announcement. Thus, market efficiency or rational expectations, respectively, imply that market participants or professional forecasters use all public information available at the start of the window and so  $Y_t$  is constructed by conditioning on that information. Second, for  $t \in \mathbf{P}$  the latest past treatment is  $D_{t-1}$  ( $\{t-1\} \in \mathbf{C}$ ) which captures the surprise change in a policy within a 30-minute window that does not involve any policy announcement and so the change is very small or equal zero. The potential outcome should not be confused with the outcome  $Y_t = Y_t(D_t)$ . Finally, note that in terms of the structural function in (2.3),  $Y_t(d) = \varphi_Y(d, Z_t, u_t, t)$ .

The notation  $Y_t(d)$  focuses on the effect of the current treatment  $d$  on the current outcome. In our context, the hypothesis of no causal effects of the policy means that  $Y_t(d) = Y_t(d')$  for all  $d, d' \in \mathbf{D}$ . To analyze the causal effect of the policy variable, it is useful to define the effect of a marginal change in the policy variable on the potential outcome, the MCE. Define the normalized variables  $\tilde{Y}_t(d) = \sigma_{D,t}^{-1} Y_t(d)$ ,  $\tilde{D}_t = \sigma_{D,t}^{-1} D_t$  and  $\tilde{e}_t = \sigma_{D,t}^{-1} e_t$  for all  $t$ , where  $\sigma_{D,t}^2 = \text{Var}(D_t)$ . We impose two technical assumptions to enable MCEs to be well-defined under relative exogeneity (2.4). The first is on the support of the normalized policy variable and the second is on the smoothness of the normalized potential outcome process.

**Assumption 5.** For all  $t \in \mathbf{P}$ ,  $\tilde{D}_t \in \mathbf{D} = [\underline{d}, \bar{d}]$  with  $\underline{d} < \bar{d}$ .

**Assumption 6.** (Differentiability) For all  $t \in \mathbf{P}$ ,  $\tilde{Y}_t(d)$  is continuously differentiable in  $d \in (\underline{d}, \bar{d})$ .

Under these assumptions, the MCE of the time  $t$  policy on the time  $t$  normalized potential outcome is defined as  $\partial \tilde{Y}_t(d) / \partial d$ . Assumptions 5-6 involve the normalized quantities instead of the actual quantities since we analyze estimands and estimators under relative exogeneity (2.4), for which the support of  $e_t$  and  $D_t$  necessarily become unbounded. Following [Rambachan and Shephard \(2021\)](#), to apply basic tools such as the fundamental theorem of calculus we need the argument of the relevant function to have support on a closed interval.

## 2.3 Identification Results

The simple event study regression approach regresses the outcome variable  $Y_t$  on the policy variable  $D_t$  at dates in the policy sample  $t \in \mathbf{P}$ . Thus, the corresponding event study estimand at time  $t \in \mathbf{P}$  is the linear projection estimand:

$$\beta_{\text{ES},t} = \frac{\text{Cov}(Y_t, D_t)}{\text{Var}(D_t)}. \quad (2.5)$$

We begin by establishing the value of this estimand in the general nonparametric model (2.3) without imposing relative exogeneity (cf. Assumption 4) and successively obtain corresponding results under relative exogeneity and then linear homogeneous treatment effects implied by (2.1), the linear model commonly assumed in practice. To establish the first result we use the following assumption in place of Assumption 4.

**Assumption 7.** For all  $t \in \mathbf{P}$ ,  $\mathbb{E}(g_{D,e}(e_t, t)^2)$ ,  $\mathbb{E}(\varphi_{Y,u}(Z_t, u_t, t)^2)$  and  $\mathbb{E}(g_{D,u}(Z_t, u_t, t)^2)$  are finite, and  $\mathbb{E}(g_{D,e}(e_t, t)^2)$  is decreasing in  $\sigma_{e,t}^2$ .

The event study estimand can be decomposed into a weighted average of MCEs and a selection bias factor.

**Theorem 1.** Let Assumptions 1-3 and 5-7 hold. Then for  $t \in \mathbf{P}$ ,

$$\beta_{\text{ES},t} = \int_{\mathbf{D}} \frac{\partial \tilde{Y}_t(d)}{\partial d} \mathbb{E}(H_t(d)) dd + \Delta_t,$$

where  $H_t(d) = \mathbf{1}\{d \leq \tilde{D}_t\}(\tilde{D}_t - \mathbb{E}(\tilde{D}_t))$  with  $\mathbb{E}(H_t(d)) \geq 0$ , and  $\int_{\mathbf{D}} \mathbb{E}(H_t(d)) dd = 1$ ,

$$\Delta_t = \mathbb{E} \left[ \tilde{Y}_t(\underline{d}) \left( \tilde{D}_t - \mathbb{E}(\tilde{D}_t) \right) \right] = \text{Cov} \left( \tilde{Y}_t(\underline{d}), \tilde{D}_t \right),$$

and  $|\Delta_t|$  is decreasing in  $\sigma_{e,t}^2$ .

The bias  $\Delta_t$  depends on the covariance between the policy variable and the potential outcome  $Y_t(\underline{d})$ . This bias would be zero if the policy  $D_t$  were randomly assigned, i.e., when there is no reverse causality and no omitted variables. Unfortunately, this is quite unrealistic in practice. However in Theorem 2, we show that as  $\sigma_{e,t}^2$  grows large, this bias term disappears. This implies that the event study estimand and the event study regression estimates can have small enough bias to remain meaningful so long as  $\sigma_{e,t}^2$  is large enough.

The following theorem establishes that  $\beta_{\text{ES},t}$  identifies a weighted average of MCEs of the time  $t$  policy on the time  $t$  outcome when relative exogeneity holds.

**Theorem 2.** Let Assumptions 1-6 hold. Then for  $t \in \mathbf{P}$ , as  $\sigma_{e,t}^2 \rightarrow \infty$

$$\beta_{\text{ES},t} \rightarrow \lim_{\sigma_{e,t}^2 \rightarrow \infty} \int_{\mathbf{D}} \frac{\partial \tilde{Y}_t(d)}{\partial d} \mathbb{E}(H_t(d)) dd,$$

where  $H_t(d) = \mathbf{1}\{d \leq \tilde{D}_t\}(\tilde{D}_t - \mathbb{E}(\tilde{D}_t))$  with  $\mathbb{E}(H_t(d)) \geq 0$ ,  $\int_{\mathbf{D}} \mathbb{E}(H_t(d)) dd = 1$ .

For the weighted average of the MCEs in Theorem 2, a higher weight

$$\mathbb{E}(H_t(d)) = \mathbb{E}\left(\widetilde{D}_t - \mathbb{E}(\widetilde{D}_t) \mid d \leq \widetilde{D}_t\right) \times \mathbb{P}\left(d \leq \widetilde{D}_t\right)$$

is not necessarily given to large values of  $d$  because large values of  $d$  may be associated with small tail probabilities of the distribution of  $\widetilde{D}_t$ ,  $\mathbb{P}(d \leq \widetilde{D}_t)$ . The intuition on how  $\beta_{\text{ES},t}$  is able to recover a weighted average of MCEs under relative exogeneity is as follows. Relative exogeneity, in combination with Assumptions 1-3, imply that when appropriately normalized, the potential outcome  $Y_t(d)$  and the policy variable  $D_t$  behave as if they are uncorrelated. Intuitively, the effect of the policy shock on  $D_t$  is not influenced by  $Y_t$  (separability) and so changes in the policy are entirely determined by changes in the policy shock (relative exogeneity). In addition, as  $\sigma_{e,t}^2 \rightarrow \infty$  the relative bias generated by the omitted variables  $Z_t$  becomes negligible since the correlation between  $Z_t$  and  $D_t$  is an order of magnitude smaller than the variation in  $D_t$  generated by  $e_t$ .

It is noteworthy that the event study estimand obtained from a standard linear regression is able to recover a nonparametric causal effect in this case. The event study empirical applications in the literature most often assume a linear simultaneous equations model. Our assumptions on the structural and reduced-form as well as the moment conditions [cf. Assumptions 1-4(ii,iii)] are easily satisfied in those contexts. Under the linear model (2.1) and relative exogeneity, Theorem 2 implies that  $\beta_{\text{ES},t} \rightarrow \beta$  since  $\partial \widetilde{Y}_t(d) / \partial d = \beta$ .

The key identification condition is relative exogeneity, which also relates to the size of the time window surrounding an announcement or event. As the size of the window expands, it becomes less likely that the policy shock dominates all other variables within the window. Given that the narrow size of the window is not a sufficient condition for identification, our results suggest that empirical work using the high-frequency event study regression approach should be very careful to isolate the surprise component of the policy news so that the policy shock dominates the other variables in the event window. Hence, our theoretical results support the concerns expressed recently by Bauer and Swanson (2023a, 2023b) on the credibility of some high-frequency event study estimates when the outcome variable is the Blue Chip forecast revision that is based on a one-month window.

Note that the high-frequency event study method is different from heteroskedasticity-based identification [cf. ?], sometimes referred to as Rigobon's method. The latter also uses information from the control sample, which includes windows (either 30-minute or 1-day windows) that do not bracket an announcement or event. The main heteroskedasticity-based identification condition is that the volatility of the policy shock is larger in the event windows than in the control windows. This is different from relative exogeneity for two reasons. First, the required increase in volatility

of the policy shock is not relative to the other variables in the event window but relative to the policy shock in the control windows. Second, the increase in volatility need not be infinite. Thus, heteroskedasticity-based identification requires neither stronger nor weaker conditions than the relative exogeneity condition. In addition, Rigobon’s estimator is an instrumental variables estimator and therefore different from the OLS high-frequency event study estimator. Since these two estimators correspond to different estimands, they identify different causal effect in general.

## 2.4 Relation of the Theoretical Results with the Empirical Literature

There are essentially no formal identification results about high-frequency event studies in the literature. Empirical works have often mentioned that changes in the policy variable in the event windows are dominated by the information about future monetary policy contained in the FOMC announcements [see, e.g., Nakamura and Steinsson (2018a)]. However, precise conditions have not been provided. Rigobon and Sack (2004) noted that the bias of the OLS event study estimator disappears in a stylized linear model when relative exogeneity holds. However, the model they analyzed is admittedly a “clear oversimplification”. As such, this stylized model does not enable practically-relevant identification analysis or econometric results. In addition, showing that the bias disappears in this context does not imply identification of a causal estimand. Nevertheless, we credit Rigobon and Sack’s heuristic analysis as an early insight into understanding the validity of the OLS event study estimator. Gürkaynak and Wright (2013) surveyed the empirical literature and presented useful discussions. They also argued that the event study regression works even when the variance of the policy shock is not large relative to the variance of the other variables because in a narrow window around an FOMC announcement the policy news can depend on lagged changes in asset returns but not on contemporaneous changes. That is, they wrote  $D_t = \alpha Y_{t-j} + Z'_{t-j} \gamma_2 + e_t$ , where  $j \geq 1$ . However, as we explain in Example 1, the choice of a narrow window is not sufficient for ruling out omitted variables bias. Further, when  $Y_t$  is the change in an asset price and  $D_t$  is, for example, the price change in the Federal funds or eurodollar futures, assuming that  $Y_t$  does not contemporaneously affect  $D_t$  constitutes a strong empirical restriction.

In recent work, Bauer and Swanson (2023a, 2023b), Cieslak (2018) and Miranda-Agrippino and Ricco (2021) provided empirical evidence for some predictability of  $D_t$  with publicly available macroeconomic or financial market information that predates the FOMC announcement. This implies that  $D_t$  does not correctly isolate the unexpected component of the policy surprise and may be in fact simultaneously determined with the outcome variable  $Y_t$ .<sup>11</sup> Bauer and Swanson

---

<sup>11</sup>This simultaneity follows from the plausible correlation between the publicly available macroeconomic and



(2023a) proposed to take the residuals from a regression of those surprises on the economic and financial variables that predate the announcements. So in Bauer and Swanson (2023a)  $D_t$  is actually the orthogonalized monetary policy surprise rather than the surprise itself. However, Bauer and Swanson (2023a) showed that the orthogonalized policy variables yield the same results as the unadjusted policy variables when  $Y_t$  is measured as the change in a 30-minute window, as for the case of asset prices or Treasury yields. This corroborates our result that some endogeneity of  $D_t$  does not preclude the validity of the approach if the variance of the unadjusted policy shocks is much larger than that of the other variables in the system.

Evidence of nonlinearities is often documented in the empirical literature [cf. Bauer and Swanson (2023a)]. Our results establish the causal meaning of the event study estimand when the relationship between the outcome and the policy variable is potentially nonlinear as long as the additive separability conditions hold.

As we explain more in detail below, the finance literature has recently documented strong evidence of information leakage, informal communication and informed trading around policy announcements [Bernile, Hu, and Tang (2016), Cieslak, Morse, and Vissing-Jorgensen (2019), Cieslak and McMahon (2023), Cieslak and Schrimpf (2019) and Lucca and Moench (2015)]. For example, government agencies routinely allow pre-release access to information to accredited news agencies under embargo agreements. Bernile, Hu, and Tang (2016) found evidence consistent with informed trading during embargoes of the FOMC announcements. They documented significant abnormal order imbalances that are in the direction of the subsequent policy surprises and showed that the information contained in lockup-related trading activity (i.e., the window immediately before the scheduled release) predicts the market reaction to the actual FOMC announcement. Here the information leakage may arise from the news media with pre-release access or from other FOMC insiders with incentives to mimic such behavior.

Lucca and Moench (2015) documented large average excess returns on U.S. equities in anticipation of monetary policy decisions made at scheduled FOMC meetings.<sup>12</sup> This pre-FOMC drift is not found for fixed-income assets. They noted that the pre-FOMC drift cannot be explained by changes in the public information set in the twenty-four hours ahead of the FOMC meeting as FOMC members refrain from providing monetary policy information through speeches and interviews in the week before FOMC meetings. They were more inclined to attribute the

---

financial market information prior to the announcement and the macroeconomic factors at the time of the announcement that may affect  $Y_t$ .

<sup>12</sup>More specifically, Lucca and Moench (2015) looked at unconditional excess returns in the twenty-four hours before scheduled FOMC announcements while Hu, Pan, Wang, and Zhu (2022) looked at the overnight excess returns before the same announcements.

pre-FOMC drift to informational frictions. This is empirically supported by Cieslak, Morse, and Vissing-Jorgensen (2019) who showed that large pre-FOMC drift is the result of news leakage prior to the announcement of unexpectedly accommodating monetary policy. They provided evidence of systematic informal communication, including both outright leaks emerging in the media and private newsletters and systematic preferential access to the Fed enjoyed by some private financial institutions. The subsequent literature [see, e.g., Hu, Pan, Wang, and Zhu (2022)] found that other major U.S. macroeconomic news announcements give rise to pre-announcement drifts in excess returns. Here the sources of the leakage depend on the specific context. Another implication of leakage, serial dependence, is also documented empirically [see, e.g., Bernile, Hu, and Tang (2016), Cieslak, Morse, and Vissing-Jorgensen (2019) and Lucca and Moench (2015)]. Overall, the literature contains substantive evidence of information leakage and informal public communication.

As we discuss below, the leakage documented in the literature does not necessarily imply that relative exogeneity is violated. Intuitively, as long as the key news is revealed with the actual announcements, the high-frequency event study is still characterized by the lumpy manner with which a disproportionate amount of information is unveiled to the public.<sup>13</sup>

With regards to the choice of the length of the event window, the early literature commonly used a 1-day window while the more recent literature recommended to use narrower windows with the goal of reducing the background noise. In addition, it is common in the literature to use the same window for both  $Y_t$  and  $D_t$ . Nakamura and Steinsson (2018a) is an exception as they used a 1-day window for  $Y_t$  and a 30-minute window for  $D_t$ .

Although a narrower window is associated with a smaller probability of including news about events other than the policy announcement, it also has disadvantages relative to using a longer window for  $Y_t$ . First, with some information leakage, asset prices  $Y_t$  may respond before the announcement is actually made. Second, with learning or sluggish market adjustments, asset prices may take some time to incorporate the news and so changes in  $Y_t$  may occur also in the hours after the announcement.<sup>14</sup> Our analysis shows that the key identification condition (relative

---

<sup>13</sup>This argument applies to the effects of President Trump’s tweets that criticize the Federal Reserve on financial markets documented by Bianchi, Kind, and Kung (2020). They showed that those tweets had a negative effect on the expected Fed funds rate with the magnitude growing by horizon. If the tweet occurs in between the time the survey is collected and the scheduled macroeconomic announcement, then  $D_t$  does not isolate the expected component of the news correctly. This challenges relative exogeneity. However, if the key macroeconomic data information is revealed in the announcement, then  $D_t$  is primarily driven by the infinite variance policy shock  $e_t$  and so missing the effect of the tweets on the updates of the expectations is negligible.

<sup>14</sup>Note that the complications arising from the effect of information leakage and learning on  $Y_t$  are different from those we discuss in Section 2.5 for proper construction of  $D_t$ . Intuitively, the impact of information leakage on  $Y_t$  depends on the potential causal effect of the policy news on  $Y_t$ . In the linear model (2.1) this is captured by  $\beta$ . If  $\beta = 0$ , then information leakage has no effect on  $Y_t$  while it does complicate the proper construction of the surprise

exogeneity) does not explicitly refer to the size of the window, it only requires that, whatever window length is chosen, the policy shock dominates any other shock within that window.

Theorem 1 allows us to provide a formal explanation for a recent debate on some puzzling event study regression results documented in the literature. It has been shown that regressions of private-sector macroeconomic forecast revisions on monetary policy surprises often produce coefficients with signs opposite to those of standard macroeconomic models.<sup>15</sup> Campbell, Evans, Fisher, and Justiniano (2012), Nakamura and Steinsson (2018a) and Romer and Romer (2000) argued in favor of the “Fed information effect” for which these puzzling results are due to monetary policy surprises revealing private information held by the Federal Reserve. Bauer and Swanson (2023a) challenged these views, arguing that these event study estimates suffer from omitted variables bias. In Section 4, we analyze this puzzle in detail and show that in the event study regressions involving the Blue Chip forecasts, it is unlikely that relative exogeneity holds because the forecast revisions are evaluated at a much lower frequency than the policy variable. Intuitively, while the policy surprise  $D_t$  is constructed as a 30-minute change,  $Y_t$  is the one-month change in the Blue Chip forecasts and so the latter likely has a large variance relative to the former as it aggregates all news and factors that are relevant over the month. Thus, it becomes important to control for macroeconomic and financial variables that predate the announcements in this context. This explains why using the orthogonalized shocks indeed allowed Bauer and Swanson (2023a) to overturn the documented puzzling estimates. In contrast, when  $Y_t$  is a 30-minute or 1-day change in an asset price or Treasury yield, the variance of  $D_t$  is much larger in relative terms and can eliminate the endogeneity arising from omitted variables that predate the FOMC announcement.

## 2.5 Robustness to Information Leakage

It is interesting to analyze when relative exogeneity may or may not fail due to information leakage about the policy news or some market anticipation of the policy change. Leakage is highly relevant in applications as documented recently in the finance literature discussed above. Leakage has the following empirical features. First, it leads to market anticipation of the news that is to be revealed at the event time. This may reduce the variance of  $e_t$  substantially for  $t \in \mathbf{P}$ , possibly making the policy shock on the same order of magnitude as the other random variables in the system, failing to dominate them in the event window. Second, leakage can be associated with

---

$D_t$  irrespective of the value of  $\beta$ .

<sup>15</sup>For example, a surprise monetary policy tightening is associated with a statistically significant upward revision in the Blue Chip consensus forecasts for real GDP growth. This is inconsistent with the standard macroeconomic view that a monetary policy tightening should cause future GDP to fall.

learning and sluggish market adjustments since the news may initially reach a small number of market participants and, through their reactions, may slowly spread into the market. This is likely to generate serial dependence in the policy shock over the hours or days prior to the scheduled announcement. We introduce these features into the model and discuss how this can alter the identification results. The introduction of leakage into the model depends on the specific context of the event under consideration. Here we focus on FOMC announcements.

Consider two successive FOMC announcement dates  $T_0, T_1 \in \mathbf{P}$  where  $T_0 < T_1$ . There are usually six weeks between any two successive FOMC announcements.<sup>16</sup> Consider splitting the regular trading hours within these six weeks into non-overlapping 30-minute windows indexed by  $t = T_0 + 1, \dots, T_1$ . The following model for the policy shock is useful for describing some of the empirical features of leakage:

$$e_t = \phi_t \sigma_t v_t + (1 - \phi_t) (\vartheta_1 v_{t-1} + \dots + \vartheta_q v_{t-q}), \quad (2.6)$$

where  $v_t$  is a white noise process with zero mean and unit variance,  $|\vartheta_j| < \infty$  for all  $j = 1, \dots, q$  and  $q > 0$  is a finite integer. Assume that  $\sigma_t^2 \rightarrow \infty$  if  $t \in \mathbf{P}$  and  $|\sigma_t^2| < \infty$  if  $t \in \mathbf{C}$ . The parameter  $\phi_t$  is the leakage parameter:

$$\phi_t = \begin{cases} 1, & \text{no leakage} \\ \sigma_t^{-1}, & \text{leakage} \end{cases}.$$

When  $\phi_{T_1} = 1$  there is no leakage as we have  $e_{T_1} = \sigma_{T_1} v_{T_1}$ ,  $\mathbb{E}(e_{T_1}) = 0$ ,  $\text{Var}(e_{T_1}) = \sigma_{e, T_1}^2 = \sigma_{T_1}^2$  and  $\mathbb{E}(e_t e_{t-j}) = 0$  for  $j > 0$ . Then, relative exogeneity holds ( $\sigma_{e, T_1}^2 \rightarrow \infty$  since  $T_1 \in \mathbf{P}$ ) and the identification results of Section 2.3 apply. On the other hand, when  $\phi_{T_1} = \sigma_{T_1}^{-1}$  there is leakage and relative exogeneity fails. To see this, note that

$$e_{T_1} = v_{T_1} + (1 - \sigma_{T_1}^{-1}) (\vartheta_1 v_{T_1-1} + \dots + \vartheta_q v_{T_1-q}),$$

$\mathbb{E}(e_{T_1}) = 0$ ,  $\text{Var}(e_{T_1}) < \infty$  since  $1 - \sigma_{T_1}^{-1} \rightarrow 1$ ,  $|\vartheta_j| < \infty$  for all  $j = 1, \dots, q$ , and  $\mathbb{E}(e_t e_{t-j}) \neq 0$  for  $j = 1, \dots, q$  and  $\mathbb{E}(e_t e_{t-j}) = 0$  for  $j > q$ . Thus, when  $\phi_{T_1} = \sigma_{T_1}^{-1}$  the variance of the policy shock  $e_{T_1}$  is not an order of magnitude larger than the variances of  $Z_t$  and  $u_t$ . Intuitively, the market has anticipated the content of the FOMC announcement. Further,  $e_{T_1}$  exhibits serial dependence up to  $q$  lags. This captures the idea that information leakage is associated with learning and sluggish

---

<sup>16</sup>To be precise, there are eight regularly scheduled FOMC announcements per year that are spaced roughly six to eight weeks apart.

market adjustments so that the information content of the announcement can be predicted.<sup>17</sup> The model also implies that the leakage cannot start in the  $q$  periods following an FOMC announcement, i.e.,  $\phi_t = 1$  for  $t = T_0 + 1, \dots, T_0 + q$ . Otherwise, the information leakage would be correlated with the news from the previous announcement, contradicting the idea behind leakage.

The model (2.6) suggests several testable empirical implications. Relative exogeneity implies that the variance of the policy variable is unbounded at each announcement window. Statistically, this corresponds to a jump in  $D_t$  at the time of the announcement. Preliminary inspection of the high-frequency time series data on  $D_t$  can be useful. A formal test involves testing for infinite variance or testing for jumps [see, e.g., [Trapani \(2016\)](#) and [Li, Todorov, and Tauchen \(2017\)](#)]. Apart from the assumptions involved, a limitation of these tests for our purposes is that they do not provide information on whether a particular application may be characterized by a value of  $\sigma_{e,t}^2$  that, although possibly finite, is relatively large enough to imply low bias in the event study estimand (see [Theorem 1](#)). We introduce a simple procedure in [Section 4](#) to diagnose whether relative exogeneity is “close enough” to holding that the event study estimator should be expected to perform well in practice.

A second implication of leakage is the serial dependence in  $e_t$  prior to the announcement. Unfortunately,  $e_t$  is not observable and is not recoverable in the presence of endogeneity. However, serial dependence in  $e_t$  implies serial dependence in  $D_t$ , which is observed. A test for leakage can thus be obtained from testing for autocorrelation in  $D_t$  in sub-samples close to the announcement.

If the extent of the information leakage is small, then it will lead only to partial market anticipation and will not prevent the news from coming out in a lumpy manner at the release time. Thus, relative exogeneity can hold in this case. However,  $Y_t$  will not capture the overall effect of the policy news as asset prices respond also during the lockup window (i.e., the window immediately before the scheduled release), leading to attenuation bias in the event study estimator. One way to address this issue would be to take  $Y_t$  to be the change in the relevant asset price in a wider window (e.g., a 1-hour or 1-day window) around the FOMC announcement than in the 30-minute window used for  $D_t$ .

Finally, information leakage could also have negative consequences for obtaining good measures of market expectations about the policy news thereby making it difficult to accurately construct the surprise component of the policy action  $D_t$ . Failure to isolate the expected component of the news leads to attenuation bias as asset prices have already responded to the expected part of the policy news. Whether this attenuation bias is important or not depends on how large the

---

<sup>17</sup>We use a moving-average specification for  $e_t$  here because the dependence that stems from leakage is limited to a few periods prior to the FOMC announcement [see, e.g., [Lucca and Moench \(2015\)](#)].

information leakage is relative to the unexpected part of the policy news that is driven by  $e_t$  in the event window. When the leakage is small, one does not need to extend the event window to bracket the pre-release embargo when computing  $D_t$  since the policy shock during the 30-minute window is still an order of magnitude larger than the other variables.

### 3 Properties of the OLS Event Study Estimator

In this section we establish the asymptotic properties of the OLS event study estimator. Theorem 2 shows that a weighted average of MCEs can be identified by the ratio of the covariance between  $Y_t$  and  $D_t$  and the variance of  $D_t$ . This is exactly what the OLS event study estimator estimates under homogeneous treatment effects and covariance-stationarity of the normalized processes.<sup>18</sup> However, when  $\beta_{\text{ES},t}$  varies with  $t$ , it is infeasible to estimate each  $\beta_{\text{ES},t}$  separately. In this section, we seek to obtain the limiting behavior of the OLS event study estimator generally, without imposing homogeneous treatment effects or covariance-stationarity. We show that the OLS event study estimator estimates a time-average of the  $\beta_{\text{ES},t}$ 's under relative exogeneity. Let  $T_P$  denote the number of observations in  $\mathbf{P}$ . The event study estimator is defined as

$$\hat{\beta}_{\text{ES}} = \frac{\sum_{t=1}^{T_P} (D_t - \bar{D}) (Y_t - \bar{Y})}{\sum_{t=1}^{T_P} (D_t - \bar{D})^2},$$

where  $\bar{D} = T_P^{-1} \sum_{t=1}^{T_P} D_t$  and  $\bar{Y} = T_P^{-1} \sum_{t=1}^{T_P} Y_t$ , and an intercept is added to the regression.

Let  $\sigma_D^2 = \lim_{T_P \rightarrow \infty} T_P^{-1} \sum_{t=1}^{T_P} \text{Var}(D_t)$ ,  $D_t^* = \sigma_D^{-1} D_t$ ,  $Y_t^* = \sigma_D^{-1} Y_t$  and  $Y_t^*(d) = \sigma_D^{-1} Y_t(d)$  for all  $t \in \mathbf{P}$ . We make the following assumption in order to study the asymptotic properties of  $\hat{\beta}_{\text{ES}}$ .

**Assumption 8.** *As  $T_P \rightarrow \infty$  we have*

(i)  $T_P^{-1} \sum_{t=1}^{T_P} (D_t^* - \bar{D}^*) (Y_t^* - \bar{Y}^*) \xrightarrow{\mathbb{P}} \int_0^1 c(D^*, Y^*, s) ds,$

(ii)  $T_P^{-1} \sum_{t=1}^{T_P} (D_t^* - \bar{D}^*)^2 \xrightarrow{\mathbb{P}} 1,$

where  $\bar{D}^* = T_P^{-1} \sum_{t=1}^{T_P} D_t^*$ ,  $\bar{Y}^* = T_P^{-1} \sum_{t=1}^{T_P} Y_t^*$ , and  $c(D^*, Y^*, s) = \lim_{T_P \rightarrow \infty} \text{Cov}(D_{[T_P s]}^*, Y_{[T_P s]}^*)$ .

Assumption 8 requires that a law of large numbers holds in an infill asymptotic embedding where the observations originally defined on the time span  $t = 1, \dots, T_P$  are mapped into the unit interval  $[0, 1]$  through  $s = t/T_P$ . We refer to the index  $s \in [0, 1]$  as the rescaled time index. This is a mild

---

<sup>18</sup>Technically speaking,  $D_t$  and  $Y_t$  cannot be said to be covariance-stationary because their second moments may diverge under relative exogeneity. In contrast, it is meaningful to say that the normalized processes are covariance-stationary.

assumption. It allows the observations to be heterogeneous, i.e., to have time-varying moments. If one assumes covariance-stationarity, then  $c(D^*, Y^*, s) = c(D^*, Y^*)$  and the infill asymptotic embedding is no longer required. Additionally,  $D_t^* = \widetilde{D}_t$ ,  $Y_t^* = \widetilde{Y}_t$  and  $c(D^*, Y^*) = c(\widetilde{D}, \widetilde{Y})$  since  $\sigma_{D,t}^2 = \sigma_D^2$  for all  $t \in \mathbf{P}$  by covariance-stationarity of the normalized processes.

**Theorem 3.** *Let Assumptions 1-3 and 5-8 hold. As  $T_P \rightarrow \infty$  we have  $\widehat{\beta}_{\text{ES}} \xrightarrow{\mathbb{P}} \beta_{\text{ES}}$  where*

$$\beta_{\text{ES}} = \lim_{T_P \rightarrow \infty} \int_0^1 \int_{\mathbf{D}} \frac{\partial Y_{[T_P s]}^*(d)}{\partial d} h(d, s) \, d \, d \, s + \Delta, \quad \text{and}$$

$$h(d, s) = \lim_{T_P \rightarrow \infty} \mathbb{E} \left( \mathbf{1}\{d \leq D_{[T_P s]}^*\} (D_{[T_P s]}^* - \mathbb{E}(D_{[T_P s]}^*)) \right)$$

with  $h(d, s) \geq 0$ ,  $\int_{\mathbf{D}} h(d, s) \, d \, d = 1$  and  $|\Delta| = \left| \lim_{T_P \rightarrow \infty} \int_0^1 \mathbb{E}(Y_{[T_P s]}^*(\underline{d}) (D_{[T_P s]}^* - \mathbb{E}(D_{[T_P s]}^*))) \, ds \right|$  is decreasing in  $\lim_{T_P \rightarrow \infty} \sigma_{e, [T_P s]}^2$  for all  $s \in [0, 1]$ .

Theorem 3 shows that the OLS event study estimator is consistent for a weighted average of standardized MCEs, plus the bias term  $\Delta$ . The former is characterized by two types of averaging. First, there is averaging over time for a given treatment  $d$ . Second, there is averaging over different treatments for a given rescaled time  $s$ . These treatment effects are said to be standardized because they involve the standardized outcome  $Y_{[T_P s]}^*$  rather than the original outcome  $Y_{[T_P s]}$ .

When  $Y_t$  is the change in an asset price, it is interesting to note that in the special case for which relative exogeneity fails,  $Z_t$  is absent and  $u_t$  is driven solely by financial microstructure noise, the OLS event study estimator suffers from attenuation bias so that it can be used to bound a true causal effect from below. Microstructure noise typically appears in ultra high-frequency data (e.g., 5 minutes and less). Since the common size of the window in event study regressions is 20 or 30 minutes, microstructure noise may be less relevant in this context, though its presence ultimately depends on the liquidity of the asset under consideration.

The following theorem establishes the consistency of the OLS event study estimator under relative exogeneity (cf. Assumption 4). Moreover, since  $|\Delta|$  is decreasing in  $\sigma_{e,t}^2$ , the large-sample bias of  $\widehat{\beta}_{\text{ES}}$  for estimating the weighted average of standardized MCEs is small when  $\sigma_{e,t}^2$  is large. For technical reasons inherent to the proof, to establish this result we take the limits as  $T_P \rightarrow \infty$  and  $\min_{t \in \mathbf{P}} \sigma_{e,t}^2 \rightarrow \infty$  sequentially.

**Theorem 4.** *Let Assumptions 1-6 and 8 hold. Then as  $T_P \rightarrow \infty$ , then  $\min_{t \in \mathbf{P}} \sigma_{e,t}^2 \rightarrow \infty$ ,*

$$\widehat{\beta}_{\text{ES}} \xrightarrow{\mathbb{P}} \lim_{\min_{t \in \mathbf{P}} \sigma_{e,t}^2 \rightarrow \infty} \lim_{T_P \rightarrow \infty} \int_0^1 \int_{\mathbf{D}} \frac{\partial Y_{[T_P s]}^*(d)}{\partial d} h(d, s) \, d \, d \, s. \quad (3.1)$$

Under covariance-stationarity the event study estimand on the right-hand side of (3.1) reduces to the estimand in Theorem 2:

$$\lim_{\min_t \in \mathbf{P}} \lim_{\sigma_{e,t}^2 \rightarrow \infty} \lim_{T_P \rightarrow \infty} \int_{\mathbf{D}} h(d) \int_0^1 \frac{\partial \tilde{Y}_{[T_P s]}(d)}{\partial d} ds dd = \lim_{\sigma_{e,t}^2 \rightarrow \infty} \int_{\mathbf{D}} \frac{\partial \tilde{Y}_t(d)}{\partial d} \mathbb{E}(H_t(d)) dd,$$

since  $h(d, s) = \mathbb{E}(H_t(d))$ ,  $Y_t^* = \tilde{Y}_t$  and  $\partial \tilde{Y}_t(d) / \partial d$  is invariant to  $t$  by stationarity.

Finally, we analyze the asymptotic distribution of the event study estimator.

**Assumption 9.** Let  $\varepsilon_t = Y_t^* - \bar{Y}^* - \beta_{\text{ES}}(D_t^* - \bar{D}^*)$ . As  $T_P \rightarrow \infty$  we have

$$\frac{1}{\sqrt{T_P}} \sum_{t=1}^{T_P} (D_t^* - \bar{D}^*) \varepsilon_t \xrightarrow{d} \mathcal{N}(0, J),$$

where

$$J = \int_0^1 J(s) ds, \quad \text{and} \quad J(s) = \lim_{T_P \rightarrow \infty} \mathbb{E} \left[ \varepsilon_{[T_P s]}^2 \left( D_{[T_P s]}^* - \mathbb{E}(D_{[T_P s]}^*) \right)^2 \right].$$

Assumption 9 requires a central limit theorem to hold in an infill asymptotic embedding. The asymptotic variance  $J$  allows for heteroskedasticity but not serial correlation. This is implied by Assumption 3 and no serial correlation in  $Z_t$ . The assumption of no serial correlation in high-frequency event studies is standard since the observations are in first-differences and the original series are often asset prices. When there is serial correlation in either  $u_t$ ,  $e_t$  or  $Z_t$  the asymptotic variance in the assumption should be modified to

$$\begin{aligned} J &= \lim_{T_P \rightarrow \infty} \frac{1}{T_P} \sum_{t=1}^{T_P} \sum_{l=1}^{T_P} \mathbb{E} [\varepsilon_t \varepsilon_l (D_t^* - \mathbb{E}(D_t^*)) (D_l^* - \mathbb{E}(D_l^*))] \\ &= \int_0^1 \int_0^1 \lim_{T_P \rightarrow \infty} \mathbb{E} \left[ \varepsilon_{[T_P s]} \varepsilon_{[T_P r]} \left( D_{[T_P s]}^* - \mathbb{E}(D_{[T_P s]}^*) \right) \left( D_{[T_P r]}^* - \mathbb{E}(D_{[T_P r]}^*) \right) \right] ds dr. \end{aligned}$$

**Theorem 5.** Let Assumptions 1-3 and 5-9 hold. As  $T_P \rightarrow \infty$  we have

$$\sqrt{T_P} (\hat{\beta}_{\text{ES}} - \beta_{\text{ES}}) \xrightarrow{d} \mathcal{N}(0, J).$$

Theorem 5 shows that heteroskedasticity-robust standard errors suffice for valid inference when there is no serial correlation in  $u_t$ ,  $e_t$  and  $Z_t$ . Under covariance-stationarity  $J(s)$  does not depend on  $s$ , and so  $J = \mathbb{E}[\varepsilon_t^2 (D_t^* - \mathbb{E}(D_t^*))^2]$ . However, heteroskedasticity-robust standard errors



are still required as  $\varepsilon_t$  may not be homoskedastic. Finally, the following corollary establishes the unbiased asymptotic normality of  $\widehat{\beta}_{\text{ES}}$  under relative exogeneity.

**Corollary 1.** *Let Assumptions 1-6 and 8-9 hold. Then as  $T_P \rightarrow \infty$ , then  $\min_{t \in \mathbf{P}} \sigma_{e,t}^2 \rightarrow \infty$ ,*

$$\sqrt{T_P} \left( \widehat{\beta}_{\text{ES}} - \lim_{\min_{t \in \mathbf{P}} \sigma_{e,t}^2 \rightarrow \infty} \lim_{T_P \rightarrow \infty} \int_0^1 \int_{\mathbf{D}} \frac{\partial Y_{[T_P s]}^*(d)}{\partial d} h(d, s) \text{d}d \text{d}s \right) \xrightarrow{d} \mathcal{N}(0, J).$$

## 4 Empirical Analysis of Identification

In Section 4.1 we present a simple procedure that can be used as a sensitivity analysis to determine whether relative exogeneity provides a good approximation in practice and apply it to the analysis of interest rate responses to monetary policy shocks. In Section 4.2 we discuss some identification issues in the event study regressions that involve Blue Chip forecasts about real GDP growth.

### 4.1 Response of Interest Rates to Monetary Policy News

We consider the regression of Nakamura and Steinsson (2018a) which is a special case of (2.1),

$$Y_t = \beta D_t + \tilde{u}_t, \quad t \in \mathbf{P}, \quad (4.1)$$

where  $Y_t$  is the 1-day change in the 2-Year or 5-Year U.S. Treasury instantaneous real forward rate,  $D_t$  is the policy news surprise constructed by the authors as a change over a 30-minute window (see Example 1 above) and  $\tilde{u}_t$  is an error term. The policy variable can be endogenous, i.e.,  $\mathbb{E}(\tilde{u}_t | D_t) \neq 0$ . However, Theorem 4 implies that if relative exogeneity (2.4) holds then the OLS event study estimator  $\widehat{\beta}_{\text{ES}}$  in (4.1) is consistent for  $\beta$ . Moreover, even if  $\sigma_{e,t}^2$  is finite but much larger than the variance of  $\tilde{u}_t$ , Theorem 3 implies that  $\widehat{\beta}_{\text{ES}}$  has low bias.

We propose a simple empirical procedure to check if  $\sigma_{e,t}^2$  is relatively large enough that  $\widehat{\beta}_{\text{ES}}$  has low bias. We apply our analysis to the setting of Nakamura and Steinsson (2018a) for concreteness, noting that it can be straightforwardly applied to any other event study regression. While it is possible to estimate  $\text{Var}(D_t)$  using the sample variance of  $D_t$  in the policy sample, this could be very large even when relative exogeneity is satisfied only for a few announcements since the variance cannot be negative, as discussed above. Additionally, it is more difficult to estimate  $\sigma_{u,t}^2$  since  $\tilde{u}_t$  is not observed and the OLS residuals may not be close in probability to the corresponding true errors  $\tilde{u}_t$  given that  $\mathbb{E}(\tilde{u}_t | D_t) \neq 0$ . The key idea to our procedure is that since  $\tilde{u}_t$  includes macroeconomic news or factors that are present even when there is no announcement, the order of

magnitude of its variance can be retrieved from the order of magnitude of the variance of  $Y_t$  in the control sample  $\mathbf{C}$ . In fact, in the control sample the variance of  $D_t$  and  $\tilde{u}_t$  have the same order of magnitude and given that  $|\beta| < \infty$ , the order of magnitude of the variance of  $Y_t$  for  $t \in \mathbf{C}$  is the same as that of  $\tilde{u}_t$  so long as (4.1) also holds for  $t \in \mathbf{C}$ . Thus, we can proxy the average variance of  $\tilde{u}_t$  for  $t \in \mathbf{P}$  by the average variance of  $Y_t$  for  $t \in \mathbf{C}$ .

We initially simulate the regression (4.1) calibrated to the corresponding regression in Nakamura and Steinsson (2018a) for the control sample, making the draws of  $D_t$  and  $\tilde{u}_t$  independent, and in each draw we estimate  $\beta$  by OLS. We repeat this many times and compute bias, mean-absolute error (MAE), and mean-squared error (MSE) for this idealized OLS estimator. We name it the “oracle” estimator since it is obtained in the absence of endogeneity and it is efficient by the Gauss-Markov Theorem.

The performance of the oracle estimator is compared to that of the corresponding OLS event study estimator when we allow for correlation between  $D_t$  and  $\tilde{u}_t$ . In the latter regression, we successively increase the variance of  $D_t$  and record how much the performance of the event study estimator approaches that of the oracle estimator in the idealized regression. In particular, we determine the threshold value for the variance of  $D_t$  that ensures that the event study estimator performs as well as the oracle. Lastly, we verify whether the sample variance of  $D_t$  in Nakamura and Steinsson (2018a) is larger than this threshold. If it is, relative exogeneity likely holds and we can interpret  $\sigma_{e,t}^2$  as being large enough relative to the variance of  $\tilde{u}_t$  for the OLS event study estimator to have low enough bias to perform well in practice.

Let  $T_C$  denote the sample size of the control sample  $t \in \mathbf{C}$ . The policy sample consists of all regularly scheduled FOMC meeting days from 1/1/2000 to 3/19/2014. The control sample includes all Tuesdays and Wednesdays that are not FOMC meeting days over the same period. This yields  $T_P = 74$  and  $T_C = 762$  when the dependent variable is the 2-Year real yield, and  $T_P = 106$  and  $T_C = 1130$  when it is the 5-Year real yield.

Consider the following steps:

1. Estimate the mean, variance and autoregressive coefficient of order 1 from all  $D_t$  in the control sample ( $t \in \mathbf{C}$ ). Denote these by  $\overline{D}_C$ ,  $\hat{\sigma}_{D,C}^2$  and  $\hat{\rho}_{D,C,1}$ . Similarly, obtain  $\hat{\sigma}_{Y,C}^2$  and  $\hat{\rho}_{Y,C,1}$  from all  $Y_t$  in the control sample ( $t \in \mathbf{C}$ ). Test whether  $\hat{\rho}_{D,C,1}$  and  $\hat{\rho}_{Y,C,1}$  are statistically significant different from zero.
2. Obtain the sample variances,  $\hat{\sigma}_{D,P}^2$  and  $\hat{\sigma}_{Y,P}^2$ , from all  $D_t$  and  $Y_t$  in the policy sample ( $t \in \mathbf{P}$ ).
3. If  $Y_t$  is constructed as a change over a substantially longer frequency than  $D_t$ , compute

$\hat{\sigma}_u^2 = \hat{\sigma}_{Y,P}^2 \hat{\sigma}_{D,C}^2 / \hat{\sigma}_{D,P}^2$ . Note that  $\hat{\sigma}_u^2$  is a proxy for the variance of  $Y_t$  in the control sample. Since  $Y_t$  is the 1-day change in the forward rate while  $D_t$  is the change in a 30-minute window, the variance of  $D_t$  is substantially smaller than that of  $Y_t$  in the control sample and the variance of  $Y_t$  and  $\tilde{u}_t$  are very similar. So we choose  $\hat{\sigma}_u^2$  such that the ratio of the variances of  $Y_t$  in the policy and control sample is the same as that of  $D_t$ , i.e.,  $\hat{\sigma}_{Y,P}^2 / \hat{\sigma}_u^2 = \hat{\sigma}_{D,P}^2 / \hat{\sigma}_{D,C}^2$ .<sup>19</sup> Otherwise, compute  $\hat{\sigma}_u^2 = \hat{\sigma}_{Y,C}^2 - \beta^2 \hat{\sigma}_{D,C}^2$  with  $\beta$  set equal to 0.99, which is the OLS estimate in Nakamura and Steinsson (2018a).<sup>20</sup>

4. For  $t = 1, \dots, T_P$  generate  $\tilde{D}_t \sim i.i.d. \mathcal{N}(\bar{D}_C, \hat{\sigma}_{D,C}^2)$  if  $\hat{\rho}_{D,C,1}$  is not statistically different from zero. Otherwise, generate  $\tilde{D}_t = (1 - \hat{\rho}_{D,C,1})\bar{D}_C + \hat{\rho}_{D,C,1}\tilde{D}_{t-1} + u_{D,t}$  where  $u_{D,t} \sim i.i.d. \mathcal{N}(0, (1 - \hat{\rho}_{D,C,1}^2)\hat{\sigma}_{D,C}^2)$  for  $t = 1, \dots, T_P$  and  $\tilde{D}_0 = 0$ .
5. For  $t = 1, \dots, T_P$  generate  $\tilde{u}_t \sim i.i.d. \mathcal{N}(0, \hat{\sigma}_u^2)$  if  $\hat{\rho}_{Y,C,1}$  is not statistically different from zero. Otherwise generate  $\tilde{u}_t$  according to  $\tilde{u}_t = \hat{\rho}_{Y,C,1}\tilde{u}_{t-1} + \tilde{v}_t$  where  $\tilde{v}_t \sim i.i.d. \mathcal{N}(0, (1 - \hat{\rho}_{Y,C,1}^2)\hat{\sigma}_u^2)$  and  $\tilde{u}_0 = 0$ .
6. Generate  $\tilde{Y}_t = \beta\tilde{D}_t + \tilde{u}_t$  for  $t = 1, \dots, T_P$  where  $\tilde{D}_t$  is generated in 4.,  $\tilde{u}_t$  in 5. and  $\beta$  is set equal to 0.99 as explained in 3. Run OLS for the regression of  $\tilde{Y}_t$  on  $\tilde{D}_t$  to obtain the OLS estimator  $\hat{\beta}_{\text{orcale}}$ .
7. Repeat 4-6. 5,000 times and compute the bias, MAE and MSE of  $\hat{\beta}_{\text{orcale}}$ .
8. Repeat the regression in 6. but now allow for correlation  $\rho$  between  $\tilde{u}_t$  and  $\tilde{D}_t$ . For a given  $\rho \in [-1, 1]$ , generate  $\tilde{u}_t = \rho(\hat{\sigma}_u / \hat{\sigma}_{D,C})(\tilde{D}_t - \bar{D}_C) + \sqrt{1 - \rho^2}\eta_t$ , where  $\eta_t \sim i.i.d. \mathcal{N}(0, \hat{\sigma}_u^2)$  for  $t = 1, \dots, T_P$ .<sup>21</sup> The case  $\rho = 0$  corresponds to the regression in 6. Generate  $\tilde{Y}_t = \beta\tilde{D}_t + \tilde{u}_t$  for  $t = 1, \dots, T_P$  with  $\beta$  as in 6. and  $\tilde{D}_t$  as in 4. Run OLS for the regression of  $\tilde{Y}_t$  on  $\tilde{D}_t$  and label this estimator  $\hat{\beta}_{\text{ES}}(\rho)$ . Repeat this 5,000 times and compute its bias, MAE and MSE.
9. Consider the regression in 8. but start to raise the variance of  $\tilde{D}_t$ . That is, define  $\tilde{D}_{\delta,t} = \bar{D}_C + (\tilde{D}_t - \bar{D}_C)\sqrt{1 + \delta}$  for some  $\delta \geq 0$  where  $\tilde{D}_t$  is generated in 4.<sup>22</sup> Generate  $\tilde{Y}_t = \beta\tilde{D}_{\delta,t} + \tilde{u}_t$  for  $t = 1, \dots, T_P$  with  $\beta$  and  $\tilde{u}_t$  as in 8. Run OLS for the regression of  $\tilde{Y}_t$  on  $\tilde{D}_{\delta,t}$  and label this estimator  $\hat{\beta}_{\text{ES}}(\rho, \delta)$ .<sup>23</sup> Repeat this 5,000 times and compute its bias, MAE and MSE.

<sup>19</sup>Note that  $\hat{\sigma}_u^2$  approximates the variance of  $\tilde{u}_t$  for  $t \in \mathbf{P}$  when  $D_t$  and  $\tilde{u}_t$  are independent and the variance of  $\tilde{u}_t$  is similar across the policy and control samples.

<sup>20</sup>We could choose any other value for  $\beta$ , the intuition behind the procedure and the results would not change.

<sup>21</sup>Note that the correlation between  $\tilde{u}_t$  and  $\tilde{D}_t$  is equal to  $\rho$  and  $\tilde{u}_t$  has variance  $\hat{\sigma}_u^2$ .

<sup>22</sup>Note that  $\tilde{D}_{\delta,t}$  is a rescaled version of  $\tilde{D}_t$  with variance  $\hat{\sigma}_{D,C}^2(1 + \delta)$ .

<sup>23</sup> $\hat{\beta}_{\text{ES}}(\rho)$  in 8. corresponds to  $\hat{\beta}_{\text{ES}}(\rho, 0)$ .

10. Find the value of  $\delta$  such that  $\text{MSE}(\widehat{\beta}_{\text{ES}}(\rho, \delta)) = \text{MSE}(\widehat{\beta}_{\text{oracle}})$  [or  $\text{MAE}(\widehat{\beta}_{\text{ES}}(\rho, \delta)) = \text{MAE}(\widehat{\beta}_{\text{oracle}})$ ] and label it  $\delta^*(\rho)$ . Define  $\widehat{\sigma}_{D,*}^2(\rho) = \widehat{\sigma}_{D,C}^2(1 + \delta^*(\rho))$ .
11. Repeat 8.-10. for a grid of  $\rho$  values in  $[-1, 1]$ .

Steps 1.-11. provide a detailed sensitivity analysis that computes the values of the variance of  $D_t$  that allows the event study estimator to perform as well as the oracle estimator over a range of values of endogeneity  $\rho$ , namely  $\widehat{\sigma}_{D,*}^2(\rho)$ . If  $\widehat{\sigma}_{D,P}^2 \geq \widehat{\sigma}_{D,*}^2(\rho)$  one can be confident that the variance of the policy shock in Nakamura and Steinsson (2018a) is relatively large enough for a degree of endogeneity that satisfies  $|\text{Corr}(\tilde{u}_t, D_t)| \leq |\rho|$  for the event study estimator to have low enough bias for its MSE (MAE) to be as low as in the oracle regression with no endogeneity. If one is not interested in the full sensitivity analysis, but rather just in evaluating the values of  $\rho$  for which the policy shock in Nakamura and Steinsson (2018a) is large enough to produce a credible event study estimator, one could replace  $1 + \delta$  in 9. by  $\widehat{\sigma}_{D,P}^2/\widehat{\sigma}_{D,C}^2$  and instead of 10., find the value of  $\rho$  such that  $\text{MSE}(\widehat{\beta}_{\text{ES}}(\rho, \widehat{\sigma}_{D,P}^2/\widehat{\sigma}_{D,C}^2 - 1)) = \text{MSE}(\widehat{\beta}_{\text{oracle}})$  [or  $\text{MAE}(\widehat{\beta}_{\text{ES}}(\rho, \widehat{\sigma}_{D,P}^2/\widehat{\sigma}_{D,C}^2 - 1)) = \text{MAE}(\widehat{\beta}_{\text{oracle}})$ ].

The validity of the sensitivity analysis relies on the following two properties: (i) the model in (4.1) is correct also in the control sample, i.e., for  $t \in \mathbf{C}$  and (ii) the order of magnitude of  $\widehat{\sigma}_u^2$  is a good proxy for the order of magnitude of  $T_P^{-1} \sum_{t \in \mathbf{P}} \sigma_{u,t}^2$ . Property (i) is implicitly assumed in typical VAR estimation of simultaneous equations models since typically the relationship between monetary policy and real economic variables is estimated with VARs at monthly, quarterly, yearly frequencies and each observation aggregates both FOMC announcement and non-announcement days. Property (ii) is reasonable since  $\tilde{u}_t$  for  $t \in \mathbf{P}$  represents news and latent factors that are not specific to the announcement and would be present even if there were no FOMC announcements. Thus, its variation is expected to be similar to that in non-announcement days which in turn has the same order of magnitude as that of  $Y_t$  in non-announcement days.

Table 1 shows the bias, MAE and MSE of  $\widehat{\beta}_{\text{oracle}}$  and  $\widehat{\beta}_{\text{ES}}(\rho, \delta)$  for different values of  $\rho$  and  $\delta$ . We first compare  $\widehat{\beta}_{\text{oracle}}$  and  $\widehat{\beta}_{\text{ES}}(\rho, \delta)$  for  $\rho = 0, 0.25, \dots, 1$  and  $\delta = 0$ . As we raise the endogeneity parameter  $\rho$  from 0 to 1 the bias, MAE and MSE of  $\widehat{\beta}_{\text{ES}}$  increases substantially.

To provide numerical support for the theoretical results of Theorem 1-4, we look at the change in the bias, MAE and MSE of  $\widehat{\beta}_{\text{ES}}(\rho, \delta)$  as  $\delta$  increases for a given  $\rho \neq 0$  when looking at the 2-Year real yields. The results show that these summary statistics decrease quickly as  $\delta$  grows. For example, for a small degree of endogeneity  $\rho = 0.25$ , it is sufficient to increase the variance of the policy variable in the announcement windows nine-fold ( $\delta = 8$ ) for the event study estimator to exhibit an MSE that is as small as that of the oracle estimator. A similar feature applies to MAE, though a slightly larger  $\delta$  is needed. As we raise  $\delta$  further, the MAE and MSE of  $\widehat{\beta}_{\text{ES}}$  become

smaller and smaller relative to those of  $\widehat{\beta}_{\text{oracle}}$ . In particular, both the MAE and MSE of  $\widehat{\beta}_{\text{ES}}$  converge to zero quickly as  $\delta$  increases for  $\rho = 0.25$ .

Table 1: Bias, MAE and MSE of  $\widehat{\beta}_{\text{oracle}}$  and  $\widehat{\beta}_{\text{ES}}$ 

| 2-Year U.S. Treasury instantaneous real forward rates |  |        |        |  |       |       |  |       |       |  |       |       |
|---|--|--------|--------|--|-------|-------|--|-------|-------|--|-------|-------|
| $\rho$  | $\widehat{\beta}_{\text{oracle}}$                |        |        | $\widehat{\beta}_{\text{ES}}(\rho, \delta = 0)$  |       |       | $\widehat{\beta}_{\text{ES}}(\rho, \delta = 1)$  |       |       | $\widehat{\beta}_{\text{ES}}(\rho, \delta = 8)$  |       |       |
|   | Bias   | MAE    | MSE    | Bias   | MAE   | MSE   | Bias   | MAE   | MSE   | Bias   | MAE   | MSE   |
| 0.  | -0.004   | 0.158  | 0.039  | 0.000  | 0.155 | 0.037 | -0.004   | 0.126 | 0.025 | 0.000  | 0.069 | 0.008 |
| 0.25  |  |        |        | 0.294  | 0.302 | 0.128 | 0.278  | 0.282 | 0.101 | 0.177  | 0.179 | 0.039 |
| 0.50  |  |        |        | 0.580  | 0.580 | 0.367 | 0.550  | 0.550 | 0.322 | 0.349  | 0.349 | 0.128 |
| 0.75  |  |        |        | 0.869  | 0.869 | 0.777 | 0.826  | 0.826 | 0.696 | 0.524  | 0.524 | 0.278 |
| 1   |  |        |        | 1.167  | 1.163 | 1.362 | 1.098  | 1.098 | 1.211 | 0.699  | 0.699 | 0.490 |
| $\rho$  | $\widehat{\beta}_{\text{ES}}(\rho, \delta = 16)$ |        |        | $\widehat{\beta}_{\text{ES}}(\rho, \delta = 32)$ |       |       | $\widehat{\beta}_{\text{ES}}(\rho, \delta = 40)$ |       |       | $\widehat{\beta}_{\text{ES}}(\rho, \delta = 48)$ |       |       |
|   | Bias   | MAE    | MSE    | Bias   | MAE   | MSE   | Bias   | MAE   | MSE   | Bias   | MAE   | MSE   |
| 0   | 0.000  | 0.052  | 0.004  | 0.000  | 0.038 | 0.002 | -0.001   | 0.034 | 0.002 | 0.000  | 0.031 | 0.002 |
| 0.25  | 0.133  | 0.134  | 0.022  | 0.098  | 0.099 | 0.012 | 0.091  | 0.091 | 0.000 | 0.081  | 0.082 | 0.008 |
| 0.50  | 0.267  | 0.267  | 0.074  | 0.197  | 0.179 | 0.041 | 0.177  | 0.177 | 0.032 | 0.163  | 0.164 | 0.028 |
| 0.75  | 0.401  | 0.401  | 0.163  | 0.296  | 0.296 | 0.088 | 0.266  | 0.266 | 0.072 | 0.245  | 0.245 | 0.061 |
| 1   | 0.534  | 0.534  | 0.285  | 0.398  | 0.398 | 0.155 | 0.355  | 0.355 | 0.126 | 0.326  | 0.326 | 0.107 |
| 5-Year U.S. Treasury instantaneous real forward rates |  |        |        |  |       |       |  |       |       |  |       |       |
| $\rho$  | $\widehat{\beta}_{\text{oracle}}$                |        |        | $\widehat{\beta}_{\text{ES}}(\rho, \delta = 0)$  |       |       | $\widehat{\beta}_{\text{ES}}(\rho, \delta = 1)$  |       |       | $\widehat{\beta}_{\text{ES}}(\rho, \delta = 8)$  |       |       |
|   | Bias   | MAE    | MSE    | Bias   | MAE   | MSE   | Bias   | MAE   | MSE   | Bias   | MAE   | MSE   |
| 0   | 0.0017   | 0.1171 | 0.0217 | -0.002   | 0.119 | 0.022 | 0.000  | 0.095 | 0.014 | -0.002   | 0.052 | 0.004 |
| 0.25  |  |        |        | 0.268  | 0.272 | 0.093 | 0.249  | 0.250 | 0.076 | 0.159  | 0.159 | 0.029 |
| 0.50  |  |        |        | 0.529  | 0.530 | 0.299 | 0.502  | 0.502 | 0.263 | 0.319  | 0.319 | 0.108 |
| 0.75  |  |        |        | 0.796  | 0.796 | 0.646 | 0.750  | 0.751 | 0.572 | 0.478  | 0.478 | 0.231 |
| 1   |  |        |        | 1.061  | 1.061 | 1.131 | 1.002  | 1.002 | 1.006 | 0.638  | 0.637 | 0.407 |
| $\rho$  | $\widehat{\beta}_{\text{ES}}(\rho, \delta = 16)$ |        |        | $\widehat{\beta}_{\text{ES}}(\rho, \delta = 32)$ |       |       | $\widehat{\beta}_{\text{ES}}(\rho, \delta = 40)$ |       |       | $\widehat{\beta}_{\text{ES}}(\rho, \delta = 48)$ |       |       |
|   | Bias   | MAE    | MSE    | Bias   | MAE   | MSE   | Bias   | MAE   | MSE   | Bias   | MAE   | MSE   |
| 0   | 0.000  | 0.040  | 0.003  | 0.000  | 0.029 | 0.002 | 0.000  | 0.025 | 0.001 | 0.000  | 0.023 | 0.000 |
| 0.25  | 0.123  | 0.123  | 0.017  | 0.090  | 0.090 | 0.009 | 0.081  | 0.081 | 0.008 | 0.074  | 0.074 | 0.006 |
| 0.50  | 0.243  | 0.243  | 0.061  | 0.179  | 0.179 | 0.033 | 0.162  | 0.162 | 0.027 | 0.149  | 0.149 | 0.023 |
| 0.75  | 0.367  | 0.364  | 0.134  | 0.269  | 0.269 | 0.073 | 0.243  | 0.243 | 0.059 | 0.223  | 0.223 | 0.050 |
| 1   | 0.487  | 0.487  | 0.237  | 0.359  | 0.359 | 0.129 | 0.324  | 0.324 | 0.105 | 0.297  | 0.297 | 0.088 |

The bias, MAE and MSE of  $\widehat{\beta}_{\text{oracle}}$  and  $\widehat{\beta}_{\text{ES}}(\rho, \delta)$ . In the top panel, the dependent variable in each regression is calibrated to the 2-Year real forward rate and  $T_P = 74$ . In bottom panel, the dependent variable in each regression is calibrated to the 5-Year real forward rate and  $T_P = 106$ . The number of replications is 5,000.

For a larger value of  $\rho$  it takes larger values of  $\delta$  to decrease the MAE and MSE of  $\hat{\beta}_{\text{ES}}$  to those of  $\hat{\beta}_{\text{oracle}}$ . For example, for  $\rho = 0.5$  increasing the variance of the policy variable in the policy sample nine-fold is not sufficient, while increasing it roughly thirty three times is. Similar features are found when the dependent variable is calibrated to the 5-Year real yield.

The bias of  $\hat{\beta}_{\text{oracle}}$  is always smaller than that of  $\hat{\beta}_{\text{ES}}$ . This follows because for  $\hat{\beta}_{\text{oracle}}$  positive and negative deviations from the true value  $\beta$  cancel each other across simulation replications, while for  $\hat{\beta}_{\text{ES}}$  these deviations tend to be positive since  $\rho > 0$ . On the other hand, these deviations for  $\hat{\beta}_{\text{ES}}$  tend to be smaller in magnitude, making its variance smaller than that of  $\hat{\beta}_{\text{oracle}}$  for large values of  $\delta$ . Focusing only on bias does not account for the higher precision of  $\hat{\beta}_{\text{ES}}$ . A user concerned about both accuracy and precision should rather focus on performance measures such as MSE or MAE, which incorporate both bias and variance.

We now turn to the most important step of our sensitivity analysis. Given that the regression is calibrated to that in [Nakamura and Steinsson \(2018a\)](#), we can verify up to which degree of endogeneity  $\rho$  the variance of the policy shock is large enough for the event study estimator to be expected to perform well in this setting. This entails finding for each  $\rho > 0$  the value  $\delta^*(\rho)$  such that  $\text{MSE}(\hat{\beta}_{\text{oracle}}) = \text{MSE}(\hat{\beta}_{\text{ES}}(\rho, \delta^*(\rho)))$  and then determining whether the estimate  $\hat{\sigma}_{D,P}^2$  from [Nakamura and Steinsson \(2018a\)](#) is larger than  $\hat{\sigma}_{D,*}^2(\rho) = \hat{\sigma}_{D,C}^2(1 + \delta^*(\rho))$ .

We report this information in [Table 2](#). As  $\rho$  rises the value of  $\delta^*$  increases since stronger endogeneity requires higher variance in the policy variable to make the event study estimator less biased. From [Nakamura and Steinsson \(2018a\)](#)  $\hat{\sigma}_{D,P}^2 = 32.43\hat{\sigma}_{D,C}^2$ . Thus, when  $Y_t$  is the 2-Year real yield we have  $\hat{\sigma}_{D,P}^2 \approx \hat{\sigma}_{D,*}^2(0.50)$  while when  $Y_t$  is the 5-Year real yield we have  $\hat{\sigma}_{D,P}^2 \approx \hat{\sigma}_{D,*}^2(0.40)$ . These imply that the event study estimator can be expected to perform well for all degrees of endogeneity no larger than roughly  $\rho = 0.50$  ( $\rho = 0.40$ ) for the 2-Year (5-Year) real yield.

Values of  $\rho = 0.50$  and  $0.40$  represent quite strong empirical contemporaneous correlation for any pair of 30-minute or 1-day changes in common macroeconomic and financial variables. These high-frequency contemporaneous correlations should not be that large in practice for two reasons. First, it is well-known that taking first-differences of trending variables reduces their variability and their contemporaneous correlation is smaller than that corresponding to the series in levels. Second, many relationships between macroeconomic and financial variables are in the form of lead-lag which implies that significant portions of dependence between any two variables is not contemporaneous. We verify this empirically by computing the pairwise contemporaneous correlations between all the time series used by [Nakamura and Steinsson \(2018a\)](#) and the main macroeconomic time series available from FRED. Even though these pairwise correlations are

Table 2: Values of  $\delta^*$  for each  $\rho$ 

| $\rho$ | 2-Year real forward | 5-Year real forward |
|--------|---------------------|---------------------|
|        | $\delta^*$          | $\delta^*$          |
| 0.10   | 0.78                | 1.50                |
| 0.15   | 2.50                | 4.00                |
| 0.20   | 4.60                | 7.30                |
| 0.25   | 7.60                | 11.9                |
| 0.30   | 11.40               | 17.80               |
| 0.35   | 15.80               | 24.80               |
| 0.40   | 21.00               | 32.40               |
| 0.45   | 26.80               | 41.20               |
| 0.50   | 33.50               | 50.80               |
| 0.55   | 40.80               | 61.50               |
| 0.60   | 49.30               | 73.50               |
| 0.65   | 56.80               | 86.50               |
| 0.70   | 66.00               | 100.50              |
| 0.75   | 75.90               | 115.30              |
| 0.80   | 86.50               | 132.10              |
| 0.85   | 97.80               | 148.20              |
| 0.90   | 109.7               | 166.50              |
| 0.95   | 122.3               | 186.70              |
| 1      | 135.5               | 207.50              |

The Values of  $\delta^*$  for each  $\rho = 0.10, \dots, 1.$ . The dependent variable is the 2-Year real forward rate (first column) and the 5-Year real forward rate (second column). The number of replications is 5,000.

typically larger than zero in absolute value, they never exceed 0.5 and they average about 0.2.

To analyze whether the variance of the policy shock is relatively large only for some announcement days and not for others, we compute the sample variance of the policy variable  $\widehat{\text{Var}}(D_t)$  over disjoint sub-samples. We consider the full sample from 1/1/2000 to 3/19/2014 and the sample post-2004.<sup>24</sup> The sample sizes are  $T = T_P + T_C = 1,236$  ( $T_P = 106$ ,  $T_C = 1130$ ) and  $T = T_P + T_C = 836$  ( $T_P = 74$ ,  $T_C = 762$ ), respectively. For each sample we construct several sub-samples with different window lengths according to the rule  $n_P = \lfloor T_P^{4/5} \rfloor$  and  $m_P = \lfloor T_P/n_P \rfloor$  where  $n_P$  is the number of observations in each window and  $m_P$  is the number of windows in the policy sample.<sup>25,26</sup> The

<sup>24</sup>We consider the sample after 2004 because the data for 2-year forward rates, are available from 2004 onward.

<sup>25</sup>The available data for the control sample is up to and including 2012. Thus, in constructing the sub-samples we consider all announcement and non-announcement days until the end of 2012.

<sup>26</sup>The choice of the window length  $n_P = \lfloor T_P^{4/5} \rfloor$  is optimal for nonparametric smoothing under an MSE criterion.

number of windows in the control sample is set equal to that in the policy sample, i.e.,  $m_C = m_P$ , and we set  $n_C = \lfloor T_C/m_C \rfloor$  so that each corresponding window in the policy and control sample brackets the same period.

The evidence in Table 3 shows that there is some time-variation in  $\widehat{\text{Var}}(D_t)$  in both the policy and control sample, though it does not deviate much from the average value computed over the policy and control sample, respectively. In particular, the ratio of  $\widehat{\text{Var}}(D_t)$  in the corresponding policy and control sub-samples displays some time-variation even though it does not fall substantially below the ratio computed over the full policy and control sample. For example, for the period 2004-2012 the rule selects two windows. In the first window, the ratio of  $\widehat{\text{Var}}(D_t)$  in the policy and control sample is 26.27. This is not much smaller than that corresponding to the sample 2004-2012, 36.72.

We also compute the variance of the policy variable in the sub-sample that includes the last two non-announcement days prior to an FOMC meeting that are available from the control sample. That is, we estimate the variance of the 30-minute changes in the policy variable across all Tuesdays and Wednesdays of the week before the announcement, two weeks before the announcement and three weeks before the announcement, labeling these as  $\widehat{\text{Var}}(\overleftarrow{D}_{t,1})$ ,  $\widehat{\text{Var}}(\overleftarrow{D}_{t,2})$  and  $\widehat{\text{Var}}(\overleftarrow{D}_{t,3})$ . The results do not show any significant evidence of leakage occurring in the 30-minute window of the three preceding weeks of an FOMC announcement:  $\widehat{\text{Var}}(\overleftarrow{D}_{t,1})$  is even smaller than  $\widehat{\text{Var}}(\overleftarrow{D}_{t,2})$  and  $\widehat{\text{Var}}(\overleftarrow{D}_{t,3})$ . Of course, this does not exclude the possibility that there is some leakage outside the 30-minute window [2:05pm-2:35pm] or in the other days that precede the announcement. This can be analyzed by applying the same approach to additional high-frequency data (e.g., for Thursdays, Fridays and Mondays). Here we only consider the data from Nakamura and Steinsson (2018a).

As discussed in Section 2.5, an implication of leakage is serial dependence in  $D_t$  in the days prior to the FOMC announcement. We evaluate this and consider policy variables constructed both as 30-minute and 1-day changes. For any given announcement day, we estimate the autoregressive coefficients in the regressions,

$$D_{t-j} = c + \rho_{j+1}D_{t-j-1} + v_{t-j}, \quad t \in \mathbf{P} \quad \text{and} \quad j = 0, \dots, 5.$$

The results in Table 3 show that only  $\hat{\rho}_2$  is significantly different from zero when  $D_t$  is the 1-day change. For the 30-minute change, none of the  $\hat{\rho}_j$ 's are statistically significant. Overall, there is little evidence of leakage in the available data. Again, this does not exclude the possibility that there is some leakage on the Thursdays, Fridays and Mondays that precede an FOMC meeting.

We conclude with a final remark. The sensitivity analysis discussed above allows for endo-



Table 3: Estimates of  $\text{Var}(D_t)$  and of  $\rho_j$ 

| <i>Panel A.</i>                            | 2004-2014                                  | 2000-2014                                  |                     | 2004-2012                   | 2000-2012                   |
|--|--|--|---------------------|-----------------------------|-----------------------------|
|  | $\widehat{\text{Var}}(D_t)$                | $\widehat{\text{Var}}(D_t)$                |                     | $\widehat{\text{Var}}(D_t)$ | $\widehat{\text{Var}}(D_t)$ |
| Policy                                     | 0.000808                                   | 0.001200                                   | Policy, 1st window  | 0.000725                    | 0.002011                    |
| Control                                    | 0.000022                                   | 0.000037                                   | Control, 1st window | 0.000028                    | 0.000054                    |
| Ratio                                      | 36.72                                      | 32.43                                      | Policy, 2nd window  | 0.001374                    | 0.001259                    |
|  |  |  | Control, 2nd window | 0.000016                    | 0.000020                    |
|  |  |  | Ratio 1st window    | 26.27                       | 37.24                       |
|  |  |  | Ratio 2nd window    | 87.85                       | 62.95                       |
| <i>Panel B.</i> 2000-2014                  |  |  | <i>Panel C.</i>     | 30-Minute                   | 1-Day                       |
| $\widehat{\text{Var}}(\overline{D}_{t,1})$ | $\widehat{\text{Var}}(\overline{D}_{t,2})$ | $\widehat{\text{Var}}(\overline{D}_{t,3})$ | $\widehat{\rho}_1$  | 0.14                        | 0.03                        |
| 0.000025                                   | 0.000052                                   | 0.000040                                   |                     | (0.84)                      | (0.18)                      |
|  |  |  | $\widehat{\rho}_2$  | 0.05                        | -0.19                       |
|  |  |  |                     | (0.12)                      | (0.15)                      |
|  |  |  | $\widehat{\rho}_3$  | 0.18**                      | 0.03                        |
|  |  |  |                     | 0.08                        | (0.11)                      |
|  |  |  | $\widehat{\rho}_4$  | -0.22                       | 0.07                        |
|  |  |  |                     | (0.19)                      | (0.12)                      |
|  |  |  | $\widehat{\rho}_5$  | 0.12                        | -0.11                       |
|  |  |  |                     | (0.18)                      | 0.16                        |
|  |  |  | $\widehat{\rho}_6$  | 0.07                        | -0.01                       |
|  |  |  |                     | (0.05)                      | (0.12)                      |

The estimates of the average variance of the policy variable  $D_t$  in different samples or sub-samples (Panel A and B), and the estimates of  $\rho_r$  for  $r = 1, \dots, 6$  (Panel C). The samples considered are from 1/1/2004 to 3/19/2014 and from 1/1/2000 to 3/19/2014. The sub-samples are constructed within each of the latter two samples using window lengths according to the MSE criterion. For the sample 2004-2014,  $m_P = 2$ ,  $n_P = 27$  and  $n_C = 381$ . For the sample 2000-2014,  $m_P = 2$ ,  $n_P = 38$  and  $n_C = 565$ .

geneity in the form of correlation between  $\widetilde{D}_t$  and  $\widetilde{u}_t$ . Although this is a natural specification, it is possible that  $\text{corr}(\widetilde{D}_t, \widetilde{u}_t) = 0$ , yet  $\mathbb{E}(\widetilde{u}_t | \widetilde{D}_t) \neq 0$ .<sup>27</sup> This is possible with a nonlinear relationship between  $\widetilde{D}_t$  and  $\widetilde{u}_t$ . To accommodate this, one could change step 6. above. For example, one could specify  $\widetilde{u}_t = \rho(\widehat{\sigma}_{\widetilde{u}}/\widehat{\sigma}_{D,C})(\widetilde{D}_t - \overline{D}_C)^2 + \sqrt{1 - \rho^2}\eta_t$  with  $\rho \in (-1, 1)$  so that  $\text{Cov}(\widetilde{D}_t, \widetilde{u}_t) = 0$  and  $\mathbb{E}(\widetilde{u}_t | \widetilde{D}_t) \neq 0$ . Then, one could proceed with the other steps as above where now higher values of  $\rho$  correspond to stronger nonlinear relationship between  $\widetilde{u}_t$  and  $\widetilde{D}_t$ , and would require a larger  $\delta$  for relative exogeneity to hold. The case of zero correlation and nonlinear dependence is likely extreme in practice, so the original sensitivity analysis above should suffice for most empirical applications.

<sup>27</sup>Of course this would require non-normality of either  $\widetilde{u}_t$  or  $\widetilde{D}_t$ .

## 4.2 Response of Blue Chip Forecasts on Output to Monetary Policy News

We consider the high-frequency event study regression in Nakamura and Steinsson (2018a):

$$BCrev_t = c + \beta D_t + \tilde{u}_t, \quad (4.2)$$

where  $BCrev_t$  is the monthly change in Blue Chip survey expectations about real GDP and  $D_t$  is the policy news shock that occurs in that month. See Bauer and Swanson (2023a) and Nakamura and Steinsson (2018a) for details on how to construct  $BCrev_t$  and  $D_t$ . It is likely that  $D_t$  is endogenous for several reasons. As shown by Bauer and Swanson (2023a)  $D_t$  is correlated with publicly known macroeconomic and financial market data, say  $X_t$ , that predate the FOMC announcement.<sup>28</sup> Since  $X_t$  is omitted from the regressors,  $\tilde{u}_t$  and  $D_t$  are correlated. Nevertheless, the event study regression is valid provided that relative exogeneity holds. The latter requires the policy shock to dominate any other variable in the event window. This means that the variance of  $D_t$  cannot be an order of magnitude smaller than the variance of  $BCrev_t$ .  $BCrev_t$  is constructed as the change in the average of the 1-, 2-, and 3-quarter ahead consensus forecasts. We denote the latter as  $BCrev_t-1q$ ,  $BCrev_t-2q$  and  $BCrev_t-3q$ , respectively.

Table 4 reports the sample variances of  $D_t$ ,  $BCrev_t$ ,  $BCrev_t-1q$ ,  $BCrev_t-2q$  and  $BCrev_t-3q$ , and of the Treasury yields over the full-sample 1995-2014 as well as over the sub-samples 2000-2014, 2000-2007 and 1995-2000. Strikingly, the variance of  $D_t$  is much smaller than that of  $BCrev_t$ . For example, in the sample 2000-2014 the variances of  $BCrev_t$  and  $BCrev_t-1q$  are thirteen and thirty four times larger than the variance of  $D_t$ . This is likely due to the fact that the event window for the dependent variable is one month but it is 30 minutes for the policy variable. Intuitively, while the policy surprise  $D_t$  is constructed as a 30-minute change, when  $Y_t$  is the one-month change in the Blue Chip forecasts, it may have too large a variance relative to the policy shock as it aggregates all news and factors that are relevant over the month.<sup>29</sup> This implies that relative exogeneity does not provide a good approximation when  $Y_t$  is a much lower-frequency change than  $D_t$ , and any correlation between  $D_t$  and  $\tilde{u}_t$  is not overwhelmed by the high variance of the policy shock.

This explanation is consistent with the evidence in Bauer and Swanson (2023a) who showed that once macroeconomic and financial data that predate the FOMC announcement are controlled for, the coefficient estimates revert back to having signs consistent with standard macroeconomic

---

<sup>28</sup>The index  $t$  of  $X_t$  should not create confusion. Since these data releases occur before the FOMC announcement,  $X_t$  collects information that is known before date  $t$  but is still observable at date  $t$  or is correlated with variables or news that are realized or occur at time  $t$ .

<sup>29</sup>An alternative explanation could be that relative exogeneity holds and  $\beta$  is very large in absolute value. We rule out this possibility as  $\beta$  would need to be implausibly large in order to generate such a large difference.

models. The authors attributed this to the correlation between  $D_t$  and  $X_t$ . To see this, assume that the true model is linear and the treatment effect is homogeneous. Then, standard macroeconomic theory suggests that  $\beta < 0$  when  $Y_t$  is the one-month change in the Blue Chip forecasts for real GDP. The correlation estimates in [Bauer and Swanson \(2023a\)](#) suggest that the omitted economic news that predate the announcement are positively correlated with  $D_t$ . Given that the variance of  $D_t$  is not substantially larger than the variance of  $Y_t$ , the bias  $\Delta_t$  is positive and so the resulting event study estimate  $\hat{\beta}_{\text{ES}}$  is an upward biased estimate of the causal effect of  $D_t$  on real GDP forecast revisions up to even having the wrong sign.

Table 4: The sample variances of  $D_t$ ,  $BCrev_t$  and Treasury yields

|                | 1995-2014 | 2000-2014 | 2000-2007 | 1995-2000 |
|----------------|-----------|-----------|-----------|-----------|
| $D_t$          | 0.0012    | 0.0012    | 0.0017    | 0.0012    |
| $BCrev_t$      | 0.0134    | 0.0158    | 0.0132    | 0.0095    |
| $BCrev_t-1q$   | 0.0334    | 0.0411    | 0.0346    | 0.0207    |
| $BCrev_t-2q$   | 0.0152    | 0.0176    | 0.0147    | 0.0116    |
| $BCrev_t-3q$   | 0.0096    | 0.0089    | 0.0084    | 0.0105    |
| 2-Year Forward |           | 0.0065    | 0.0051    |           |
| 5-Year Forward |           | 0.0054    | 0.0026    |           |

The sample variance of  $D_t$ ,  $BCrev_t$  and Treasury yields over different sub-samples.  $BCrev_t-1q$ ,  $BCrev_t-2q$  and  $BCrev_t-3q$  denote the 1-, 2-, and 3-quarter ahead Blue Chip forecast revisions about real GDP, respectively. For the Treasury yields the sample starts in January 2004.

This argument should also apply to the event study regression with Treasury yields as dependent variable. However, [Bauer and Swanson \(2023a\)](#) found that in regressions where the dependent variable is the change in an asset price or Treasury yield, controlling for macroeconomic and financial data that predate the FOMC announcement does not change the point estimates and their statistical significance. This difference likely arises because the changes in Treasury yields are constructed using a 30-minute or 1-day window around the announcement so that relative exogeneity is likely to provide a good approximation. As can be seen from [Table 4](#), the changes in Treasury yields based on a 1-day window have an order of magnitude similar to those of the policy variable. Since relative exogeneity provides a good approximation in this case, controlling for  $X_t$  does not result in a change in the point estimates even though  $D_t$  and  $X_t$  are correlated. The same omitted economic news that predate the announcement do not generate bias when  $Y_t$  is the change of an asset price or Treasury yield over a similarly-sized narrow window used to construct the policy surprise. The key point here is that the validity of the event study approach does not require the

absence of endogeneity. Rather, it requires that the policy shock dominates any other variable that is present in the event window. When this condition holds, identification of causal effects from an event study does not require the inclusion of controls in the event study regression. It is interesting to note that this implies that, when the identification conditions are met, event study regressions are immune to forms of p-hacking that involve searching through different control specifications, a potential strength of the high-frequency event study method.

This discussion suggests that the event study approach is more credible when a narrow window is used for both the dependent and independent variable. Both 30-minute and 1-day windows are good choices as relative exogeneity is more likely to hold. This was also informally discussed on p. 1289 in [Nakamura and Steinsson \(2018a\)](#). If longer windows (e.g., one-month windows) are used to form the outcome variable, as is the case for Blue Chip forecasts, then relative exogeneity is less likely to hold and the researcher should make more effort to appropriately control for omitted variables and simultaneity. For example, orthogonalizing surprises rather than using the original surprises as suggested by [Bauer and Swanson \(2023a\)](#) could provide a solution in these contexts.

## 5 Conclusions

We establish nonparametric conditions for identification of casual effects in high-frequency event studies. We show that identification can be achieved via a separability condition on the policy shock from the other variables present in the window, and relative exogeneity which refers to the variance of the policy shock being an order of magnitude larger than that of the other variables. Under these conditions we establish the causal meaning of the event study estimand, the consistency and asymptotic distribution of the event study estimator and its robustness to nonlinearities. We propose a simple procedure that can be used to assess relative exogeneity as an approximation and apply it to [Nakamura and Steinsson's \(2018a\)](#) analysis on the real effects of monetary policy.

## References

- ACEMOGLU, D., T. A. HASSAN, AND A. TAHOUN (2018): “The Power of the Street: Evidence from Egypt’s Arab Spring,” *Review of Financial Studies*, 31(1), 1–42.
- AI, H., AND R. BANSAL (2018): “Risk Preferences and the Macroeconomic Announcement Premium,” *Econometrica*, 86(4), 1383–1430.
- ANGRIST, J. D., AND G. M. KUERSTEINER (2011): “Causal Effects of Monetary Shocks: Semi-parametric Conditional Independence Tests with a Multinomial Propensity Score,” *Review of Economics and Statistics*, 93(3), 725–747.
- BAUER, M. D., AND E. T. SWANSON (2023a): “An Alternative Explanation for the “Fed Information Effect,”” *American Economic Review*, 113(3), 664–700.
- (2023b): “A Reassessment of Monetary Policy Surprises and High-Frequency Identification,” in *NBER Macroeconomics Annual 2022*, vol. 37.
- BERNILE, G., J. HU, AND Y. TANG (2016): “Can Information be Locked Up? Informed Trading Ahead of Macro-News Announcements,” *Journal of Financial Economics*, 121(3), 496–520.
- BIANCHI, F., T. KIND, AND H. KUNG (2020): “Threats to Central Bank Independence: High-Frequency Identification with Twitter,” *Journal of Monetary Economics*, 135, 37–54.
- CABALLERO, R. J., AND A. SIMSEK (2022): “Monetary Policy with Opinionated Markets,” *American Economic Review*, 112(7), 2353–2392.
- (2023): “A Monetary Policy Asset Pricing Model,” *Unpublished manuscript*.
- CAMPBELL, J. R., C. L. EVANS, J. D. M. FISHER, AND A. JUSTINIANO (2012): “Macroeconomic Effects of Federal Reserve Forward Guidance,” *Brookings Papers on Economic Activity*, 42(1), 1–54.
- CIESLAK, A. (2018): “Short-Rate Expectations and Unexpected Returns in Treasury Bonds,” *Review of Financial Studies*, 31(9), 3265–306.
- CIESLAK, A., AND M. MCMAHON (2023): “Tough Talk: The Fed and the Risk Premium,” *Unpublished Manuscript*, Available at SSRN: <https://ssrn.com/abstract=4560220>.
- CIESLAK, A., A. MORSE, AND A. VISSING-JORGENSEN (2019): “Stock Returns over the FOMC Cycle,” *Journal of Finance*, 74(5), 2201–2248.
- CIESLAK, A., AND A. SCHRIMPF (2019): “Non-Monetary News in Central Bank Communication,” *Journal of International Economics*, 118, 293–315.
- COCHRANE, J. H., AND M. PIAZZESI (2002): “The Fed and Interest Rates: A High-Frequency Identification,” *American Economic Review*, 92(2), 90–95.
- COOK, T., AND T. HAHN (1989): “The Effect of Changes in the Federal Funds Rate Target on

- Market Interest Rates in the 1970s,” *Journal of Monetary Economics*, 24(3), 331–351.
- DUBE, A., E. KAPLAN, AND S. NAIDU (2011): “Coups, Corporations, and Classified Information,” *Quarterly Journal of Economics*, 126(3), 1375–1409.
- FAUST, J., J. H. ROGERS, S.-Y. B. WANG, AND J. H. WRIGHT (2007): “The High-Frequency Response of Exchange Rates and Interest Rates to Macroeconomic Announcements,” *Journal of Monetary Economics*, 54, 1051–1068.
- GARRED, J., L. STICKLAND, AND N. WARRINNIER (2023): “On Target? Sanctions and the Interests of Elite Policymakers in Iran,” *The Economic Journal*, 133(649), 159–200.
- GUIDOLIN, M., AND E. LA FERRARA (2007): “Diamonds Are Forever, Wars Are Not: Is Conflict Bad for Private Firms?,” *American Economic Review*, 97(5), 1978–1993.
- GÜRKAYNAK, R. S., B. SACK, AND E. T. SWANSON (2005): “Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements,” *International Journal of Central Banking*, 1(1), 55–93.
- GÜRKAYNAK, R. S., AND J. H. WRIGHT (2013): “Identification and Inference Using Event Studies,” *The Manchester School, University of Manchester*, 81, 48–65.
- GÜRKAYNAKY, R. S., B. KISACIKOÇLU, AND J. H. WRIGHT (2020): “Missing Events in Event Studies: Identifying the Effects of Partially-Measured News Surprises,” *American Economic Review*, 110(12), 3871–3912.
- HANSEN, S., M. MCMAHON, AND A. PRAT (2018): “Transparency and Deliberation within the FOMC: a Computational Linguistics Approach,” *Quarterly Journal of Economics*, 133(2), 801–870.
- HANSON, S. G., AND J. C. STEIN (2015): “Monetary Policy and Long-Term Real Rates,” *Journal of Financial Economics*, 115(3), 429–448.
- HECKMAN, J. J., AND E. VYTLACIL (2001): “Policy-Relevant Treatment Effects,” *American Economic Review*, 91(2), 107–111.
- HU, G. H., J. PAN, J. WANG, AND H. ZHU (2022): “Premium for Heightened Uncertainty: Explaining Pre-Announcement Market Returns,” *Journal of Financial Economics*, 145(3), 909–936.
- IMBENS, G. W., AND J. D. ANGRIST (1994): “Identification and Estimation of Local Average Treatment Effects,” *Econometrica*, 62(2), 467–475.
- JAROCIŃSKI, M., AND P. KARADI (2020): “Deconstructing Monetary Policy Surprises—The Role of Information Shocks,” *American Economic Journal: Macroeconomics*, 12(2), 1–43.
- KUTTNER, K. N. (2001): “Monetary Policy Surprises and Interest Rates: Evidence from the Fed

- Funds Futures Market,” *Journal of Monetary Economics*, 47(3), 523–544.
- LI, J., V. TODOROV, AND G. TAUCHEN (2017): “Jump Regressions,” *Econometrica*, 85(1), 173–195.
- LUCCA, D. O., AND E. MOENCH (2015): “The Pre-FOMC Announcement Drift,” *Journal of Finance*, 70(1), 329–371.
- MICHELACCI, C., AND L. PACIELLO (2020): “Ambiguous Policy Announcements,” *Review of Economic Studies*, 87(5), 2356–2398.
- MIRANDA-AGRIPPINO, S., AND R. RICCO (2021): “The Transmission of Monetary Policy Shocks,” *American Economic Journal: Macroeconomics*, 13(3), 74–107.
- NAKAMURA, E., AND J. STEINSSON (2018a): “High Frequency Identification of Monetary Non-Neutrality: The Information Effect,” *Quarterly Journal of Economics*, 133(3), 1283–1330.
- (2018b): “Identification in Macroeconomics,” *Journal of Economic Perspectives*, 32(3), 59–86.
- NEUHIERL, A., AND M. WEBER (2019): “Monetary Policy Communication, Policy Slope, and the Stock Market,” *Journal of Monetary Economics*, 108, 140–155.
- PLAGBORG-MØLLER, M. (2019): “Bayesian Inference on Structural Impulse Response Functions,” *Quantitative Economics*, 10(1), 145–184.
- RAMBACHAN, A., AND N. SHEPHARD (2021): “When Do Common Time Series Estimands Have Nonparametric Causal Meaning?,” *Unpublished Manuscript, Department of Economics, Harvard University*.
- RIGOBON, R., AND B. SACK (2003): “Measuring the Reaction of Monetary Policy to the Stock Market,” *The Quarterly Journal of Economics*, 118(2), 639–669.
- (2004): “The Impact of Monetary Policy on Asset Prices,” *Journal of Monetary Economics*, 51(8), 1553–1575.
- ROMER, C. D., AND D. H. ROMER (2000): “Federal Reserve Information and the Behavior of Interest Rates,” *American Economic Review*, 90(3), 429–457.
- RUBIN, D. (1974): “Estimating Causal Effects of Treatments in Randomized and Non-Randomized Studies,” *Journal of Educational Psychology*, 66(5), 688–701.
- SWANSON, E. T. (2021): “Measuring the Effects of Federal Reserve Forward Guidance and Asset Purchases on Financial Markets,” *Journal of Monetary Economics*, 118, 32–53.
- TRAPANI, L. (2016): “Testing for (In)finite Moments,” *Journal of Econometrics*, 191(1), 57–68.

## A Mathematical Proofs

We begin with the following lemma.

**Lemma A.1.** *Let Assumptions 1-4 hold. Then, for all  $t \in \mathbf{P}$  and all  $d \in \mathbf{D}$ ,*

$$\mathbb{E}(\tilde{Y}_t(d) (\tilde{D}_t - \mathbb{E}(\tilde{D}_t))) \quad \text{and} \quad \mathbb{E}(Y_t^*(d) (D_t^* - \mathbb{E}(D_t^*)))$$

are monotonically decreasing to zero as  $\sigma_{e,t}^2 \rightarrow \infty$ , where  $Y_t^*(d) = \sigma_D^{-1} Y_t(d)$  with  $\sigma_D^2 = \lim_{T_P \rightarrow \infty} T_P^{-1} \sum_{t=1}^{T_P} \text{Var}(D_t)$  and  $T_P$  is the number of observations in the policy sample.

*Proof.* We only provide the proof for  $\mathbb{E}(\tilde{Y}_t(d) (\tilde{D}_t - \mathbb{E}(\tilde{D}_t)))$  since the proof for  $\mathbb{E}(Y_t^*(d) (D_t^* - \mathbb{E}(D_t^*)))$  is nearly identical. Using the structural and reduced forms for  $Y_t$  and  $D_t$  in Assumption 1-2, we have

$$\begin{aligned} & \mathbb{E}(\tilde{Y}_t(d) (\tilde{D}_t - \mathbb{E}(\tilde{D}_t))) \\ &= \mathbb{E}(\sigma_{D,t}^{-1} (\varphi_{Y,D}(d, t) + \varphi_{Y,u}(Z_t, u_t, t)) (\tilde{D}_t - \mathbb{E}(\tilde{D}_t))) \\ &= \sigma_{D,t}^{-2} \mathbb{E}[(\varphi_{Y,D}(d, t) + \varphi_{Y,u}(Z_t, u_t, t)) \\ & \quad \times (g_{D,e}(e_t, t) + g_{D,u}(Z_t, u_t, t) - \mathbb{E}(g_{D,e}(e_t, t) + g_{D,u}(Z_t, u_t, t)))] \\ &= \sigma_{D,t}^{-2} \varphi_{Y,D}(d, t) \mathbb{E}(g_{D,e}(e_t, t) + g_{D,u}(Z_t, u_t, t) - \mathbb{E}(g_{D,e}(e_t, t) + g_{D,u}(Z_t, u_t, t))) \\ & \quad + \sigma_{D,t}^{-2} \mathbb{E}(\varphi_{Y,u}(Z_t, u_t, t) (g_{D,e}(e_t, t) + g_{D,u}(Z_t, u_t, t) - \mathbb{E}(g_{D,e}(e_t, t) + g_{D,u}(Z_t, u_t, t)))) \\ &= 0 + \sigma_{D,t}^{-2} \mathbb{E}(\varphi_{Y,u}(Z_t, u_t, t) (g_{D,e}(e_t, t) - \mathbb{E}(g_{D,e}(e_t, t)))) \tag{A.1} \\ & \quad + \sigma_{D,t}^{-2} \mathbb{E}(\varphi_{Y,u}(Z_t, u_t, t) (g_{D,u}(Z_t, u_t, t) - \mathbb{E}(g_{D,u}(Z_t, u_t, t)))). \end{aligned}$$

By Assumption 3,  $e_t$  is independent of  $(Z_t, u_t)$  and so

$$\sigma_{D,t}^{-2} \mathbb{E}(\varphi_{Y,u}(Z_t, u_t, t) (g_{D,e}(e_t, t) - \mathbb{E}(g_{D,e}(e_t, t)))) = 0. \tag{A.2}$$

By Assumption 4(iii),

$$\begin{aligned} & \sigma_{D,t}^{-2} \mathbb{E}(\varphi_{Y,u}(Z_t, u_t, t) (g_{D,u}(Z_t, u_t, t) - \mathbb{E}(g_{D,u}(Z_t, u_t, t)))) \tag{A.3} \\ &= \sigma_{D,t}^{-2} C, \end{aligned}$$

for some  $C < \infty$ . Thus, (A.1)-(A.3) imply the statement of the lemma since  $\sigma_{D,t}^2 \rightarrow \infty$  as  $\sigma_{e,t}^2 \rightarrow \infty$  by Assumption 4(ii).  $\square$



### A.1 Proof of Theorem 1

Recall that  $\tilde{Y}_t = \tilde{Y}_t(\tilde{D}_t)$ . By the fundamental theorem of calculus and Assumption 5-6, we have

$$\begin{aligned}\tilde{Y}_t &= \tilde{Y}_t(\underline{d}) + \int_{\underline{d}}^{\tilde{D}_t} \frac{\partial \tilde{Y}_t(d)}{\partial d} dd \\ &= \tilde{Y}_t(\underline{d}) + \int_{\underline{d}}^{\bar{d}} \frac{\partial \tilde{Y}_t(d)}{\partial d} \mathbf{1}\{d \leq \tilde{D}_t\} dd.\end{aligned}$$

Using this and the fact that  $\text{Var}(\tilde{D}_t) = 1$ , we have

$$\begin{aligned}\beta_{\text{ES},t} &= \frac{\text{Cov}(\tilde{Y}_t, \tilde{D}_t)}{\text{Var}(\tilde{D}_t)} = \text{Cov}(\tilde{Y}_t, \tilde{D}_t) \\ &= \text{Cov}(\tilde{Y}_t(\underline{d}), \tilde{D}_t) + \mathbb{E} \left( \int_{\underline{d}}^{\bar{d}} \frac{\partial \tilde{Y}_t(d)}{\partial d} \mathbf{1}\{d \leq \tilde{D}_t\} (\tilde{D}_t - \mathbb{E}(\tilde{D}_t)) dd \right) \\ &= \Delta_t + \int_{\underline{d}}^{\bar{d}} \mathbb{E} \left( \frac{\partial \tilde{Y}_t(d)}{\partial d} \mathbf{1}\{d \leq \tilde{D}_t\} dd (\tilde{D}_t - \mathbb{E}(\tilde{D}_t)) \right) \\ &= \Delta_t + \int_{\underline{d}}^{\bar{d}} \frac{\partial \tilde{Y}_t(d)}{\partial d} \mathbb{E}(\mathbf{1}\{d \leq \tilde{D}_t\} (\tilde{D}_t - \mathbb{E}(\tilde{D}_t))) dd,\end{aligned}$$

where the fourth equality holds by Fubini's Theorem and Assumption 5-6 and the final equality follows from Assumption 1 since  $\partial \tilde{Y}_t(d)/\partial d = \partial \varphi_{Y,D}(d, t)/\partial d$ , proving the first statement of the theorem.

To see that the weights are non-negative, note that for  $d \in [\underline{d}, \bar{d}]$  we have

$$\begin{aligned}&\mathbb{E}(\mathbf{1}\{d \leq \tilde{D}_t\} (\tilde{D}_t - \mathbb{E}(\tilde{D}_t))) \\ &= \mathbb{E}(\mathbf{1}\{d \leq \tilde{D}_t\} \tilde{D}_t) - \mathbb{E}(\mathbf{1}\{d \leq \tilde{D}_t\}) \mathbb{E}(\tilde{D}_t) \\ &= \mathbb{E}(\tilde{D}_t | d \leq \tilde{D}_t) \mathbb{E}(\mathbf{1}\{d \leq \tilde{D}_t\}) - \mathbb{E}(\mathbf{1}\{d \leq \tilde{D}_t\}) \mathbb{E}(\tilde{D}_t) \\ &= (\mathbb{E}(\tilde{D}_t | d \leq \tilde{D}_t) - \mathbb{E}(\tilde{D}_t)) \mathbb{P}(d \leq \tilde{D}_t) \geq 0,\end{aligned}$$

since  $\mathbb{E}(\tilde{D}_t | d \leq \tilde{D}_t) - \mathbb{E}(\tilde{D}_t) \geq 0$  for  $d \in [\underline{d}, \bar{d}]$ . To see that the weights integrate to one, note that

$$\tilde{D}_t = \underline{d} + \int_{\underline{d}}^{\tilde{D}_t} d\tilde{d} = \underline{d} + \int_{\underline{d}}^{\bar{d}} \mathbf{1}\{\tilde{d} \leq \tilde{D}_t\} d\tilde{d}.$$

Using this we have,

$$\begin{aligned}
 1 = \text{Var}(\widetilde{D}_t) &= \mathbb{E} \left[ (\widetilde{D}_t - \underline{d}) (\widetilde{D}_t - \mathbb{E}(\widetilde{D}_t)) \right] \\
 &\quad - \mathbb{E} \left[ \mathbb{E} \left( \int_{\underline{d}}^{\bar{d}} \mathbf{1}\{d \leq \widetilde{D}_t\} dd \right) (\widetilde{D}_t - \mathbb{E}(\widetilde{D}_t)) \right] \\
 &= \mathbb{E} \left[ (\widetilde{D}_t - \underline{d}) (\widetilde{D}_t - \mathbb{E}(\widetilde{D}_t)) \right] \\
 &\quad - \mathbb{E} \left( \int_{\underline{d}}^{\bar{d}} \mathbf{1}\{d \leq \widetilde{D}_t\} dd \right) \mathbb{E}(\widetilde{D}_t - \mathbb{E}(\widetilde{D}_t)) \\
 &= \mathbb{E} \left[ (\widetilde{D}_t - \underline{d}) (\widetilde{D}_t - \mathbb{E}(\widetilde{D}_t)) \right] \\
 &= \mathbb{E} \left[ \int_{\underline{d}}^{\bar{d}} \mathbf{1}\{d \leq \widetilde{D}_t\} d\widetilde{d} (\widetilde{D}_t - \mathbb{E}(\widetilde{D}_t)) \right] \\
 &= \int_{\underline{d}}^{\bar{d}} \mathbb{E}(\mathbf{1}\{d \leq \widetilde{D}_t\} (\widetilde{D}_t - \mathbb{E}(\widetilde{D}_t))) dd.
 \end{aligned}$$

Finally, the fact that  $\Delta_t$  is decreasing in  $\sigma_{e,t}^2$  follows directly from (A.3) which continues to hold when Assumption 4 is replaced by Assumption 7.  $\square$

## A.2 Proof of Theorem 2

The result follows directly from Theorem 1 and Lemma A.1.  $\square$

## A.3 Proof of Theorem 3

We have

$$\begin{aligned}
 \widehat{\beta}_{\text{ES}} &= \frac{\sigma_D^{-2} T_P^{-1} \sum_{t=1}^{T_P} (D_t - \overline{D}) (Y_t - \overline{Y})}{\sigma_D^{-2} T_P^{-1} \sum_{t=1}^{T_P} (D_t - \overline{D})^2} \\
 &= \frac{T_P^{-1} \sum_{t=1}^{T_P} (D_t^* - \overline{D}^*) (Y_t^* - \overline{Y}^*)}{T_P^{-1} \sum_{t=1}^{T_P} (D_t^* - \overline{D}^*)^2} \\
 &\xrightarrow{\mathbb{P}} \int_0^1 c(D^*, Y^*, s) ds
 \end{aligned}$$

by Assumption 8. Given the definition of  $c(D^*, Y^*, s)$ , the statements of the theorem then follow from the same arguments as in the proof of Theorem 1.  $\square$

#### A.4 Proof of Theorem 4

Note that Theorem 3 is obtained under Assumption 7 while the current theorem uses Assumption 4. The difference between the two assumptions is that  $\mathbb{E}(g_{D,e}(e_t, t)^2) < \infty$  in Assumption 7 while  $\mathbb{E}(g_{D,e}(e_t, t)^2) \rightarrow \infty$  as  $\sigma_{e,t}^2 \rightarrow \infty$  in Assumption 4. Given that we take the limits sequentially, we can use the limiting result from Theorem 3 before taking the limit as  $\sigma_{e,t}^2 \rightarrow \infty$ . Thus, the theorem follows directly from Theorem 3 and Lemma A.1.  $\square$

#### A.5 Proof of Theorem 5

We have

$$\hat{\beta}_{\text{ES}} - \beta_{\text{ES}} = \frac{T_P^{-1} \sum_{t=1}^{T_P} (D_t^* - \bar{D}^*) (Y_t^* - \bar{Y}^* - \beta_{\text{ES}} (D_t^* - \bar{D}^*))}{T_P^{-1} \sum_{t=1}^{T_P} (D_t^* - \bar{D}^*)^2} = \frac{T_P^{-1} \sum_{t=1}^{T_P} (D_t^* - \bar{D}^*) \varepsilon_t}{T_P^{-1} \sum_{t=1}^{T_P} (D_t^* - \bar{D}^*)^2}.$$

Using Assumption 8(ii) and Assumption 9, we obtain the statement of the theorem.  $\square$

#### A.6 Proof of Corollary 1

The result follows directly from Theorem 5 and Lemma A.1.  $\square$