# Negative Control Identification of Monetary Policy

Danila Maroz

November 15, 2023

(Click here for the latest version)

**Abstract**

The present work explores a novel identification method in the setting of monetary policy. This method, called the negative control approach, borrows ideas from the conditional exogeneity and the IV methods but dispenses with the exogeneity assumptions while shifting the focus to the confounding mechanism. An appealing conceptual feature of this model is the fact that endogeneity can be exploited as much as exogeneity. Pioneered in Miao, Geng and Tchetgen Tchetgen (2018) in the context of biostatistics, the approach relies on the use of additional variables and a set of orthogonality conditions to construct a proxy of the endogenous component and block the omitted variable bias. This work is the first application of the negative control approach in the macroeconomic setting. In identifying the causal effects of the market interest rates on the macroeconomic outcomes of interest, we employ macroeconomic news and measures of Fed policy in a novel fashion as the additional variables satisfying orthogonality restrictions derived from the analysis of the information flow. Two key orthogonality conditions are used: (1) macroeconomic news only affects the Fed policy and the market rates by influencing the expectations of the unobserved state of the economy and (2) Fed policy only affects the macroeconomic outcomes through the influence on the market interest rates. A simple state space model supporting the reasoning is presented and the empirical application is benchmarked against the IV approach used in Bauer and Swanson (2023*b*). Resulting estimates support the idea that the improper use of additional observables might underestimate the impact of monetary policy. This work also provides ample ground for further research in expanding the set of auxiliary variables and exploring nonlinear extensions.

# 1 Introduction

The exploration of causal effects of monetary policy is of great importance, particularly so within the current context of higher inflation. As the central banks around the world hike interest rates to tame price rises, such actions have pervasive effects on the whole economy as the structure of interest rates governing the intertemporal financing constraints changes significantly.

The question of identifying the causal effects of monetary policy on macroeconomic outcomes has a long tradition. The problem is not trivial. There are many factors that influence both the policy decisions and the macroeconomic developments. The tightness of the labor market, the state of manufacturing and service sectors, business and consumer confidence are just a few example. If not properly addressed, such factors might hinder identification and distort the estimates. A vast set of tools have been developed as a response. One end of the spectrum consists of structural models, where the causal effect is a function of the underlying parameters governing the decision functions of the economic agents. On the other end are the empirical models that attempt to estimate the causal impact by relying on the reduced form relations in the data and imposing additional assumptions driven by economic reasoning.

The present work explores a novel identification approach in the canonical macroeconomic setting. It expands on the front of empirical models by using a novel identification approach, the negative control approach. This approach will borrow ideas from the conditional exogeneity and the instrumental variable methods to construct a proxy of the endogenous variation in order to provide causal identification. This method exploits the analysis of confounding mechanisms as a source of identification. Restrictions on conditional dependence of the variables in the system will warrant a creation of a control variable that will capture the source of endogenous variation and lead to unbiased causal inference.

We answer the question of identifying the causal effects of changes in the market interest rates on the macroeconomic outcomes. This relation is marred by the presence of unobserved confounding by the hidden state of the economy. We address the identification question using the negative control method. The key idea of the negative control identification approach lies in two orthogonality restrictions using two additional variables, a Fed policy measure and macroeconomic news. These variables are called negative control treatment and negative control outcome respectively. One required restriction involves the Fed policy measure and is similar to the instrumental variable exclusion restriction under the presence of an unobserved confounder. In our case it states that the Fed policy (fed funds rate) does not affect the outcome of interest (unemployment) conditional on market interest rates once the hidden state is controlled for. Intuitively, Fed policy is only thought to affect macroeconomic outcomes through its impact on the market interest rates. The second orthogonality restriction involves both additional variables and helps to control for the state of the economy. In our setting it states that macroeconomic news only affect the behavior of the Fed (fed funds rate in particular) and the markets by changing these actors' views of the state of the economy. This restriction implies that macroeconomic news can serve the role of the proxy for the unobserved state. If the proposed restrictions hold and the additional variables have sufficient variation to inform of the movement in the state of the economy, we can employ the negative control method to identify causal effects.

Borrowing ideas from both conditional exogeneity and the instrumental variable approaches, the negative control framework extracts the endogenous part of the variation in macroeconomic news using the fed funds rate as an instrument. This extracted endogenous component is then used as a control variable in the regression of the outcome of interest on the market interest rates which yields unbiased estimates of the causal effect.

The negative control approach as an identification method has been introduced in Miao, Geng and Tchetgen Tchetgen (2018). The roots of the approach come from the biostatistical literature where the idea was used to detect unobserved confounding. Negative control variables draw their name from the fact that they should be irrelevant when confounding factors are controlled for. The usage of negative controls for identification is a recent development. The essence of the approach lies in the retrieval of the confounding mechanism and its subsequent usage for identification.

The current paper is the first to introduce the negative control method to macroeconomics. This work focuses on a linear case as most common in empirical macroeconomics settings although the approach allows for nonlinear extensions. Following the introduction of the identification result and the needed assumptions, we provide the estimation algorithm and discuss its properties. To better understand how negative control approach is placed within the spectrum of existing methods, we compare it against two other commonly used identification techniques – SVARS and IV methods. In the vein of comparison, we apply the proposed method to the setting of Bauer and Swanson (2023*b*) which serves to benchmark the negative control estimates to the recently proposed baseline high-frequency IV estimates.

In terms of the contributions, this work presents a novel angle to the question of identification of monetary policy. In particular, we leverage the information structure underlying the monetary policy identification question to justify of the orthogonality restrictions. The necessary conditions follow from the reasoning of information usage: macroeconomic news inform the markets and the Fed of the state of the economy, the Fed acts upon its information and the markets adjust the interest rates which are the instrument of direct influence on the economy. We expand on these assumptions arguing both from the empirical viewpoint as well as presenting a more structural argument.

We also present a simple structural model to help guide the reasoning and support the orthogonality restrictions. This model takes into account the endogeneity of the macroeconomic news and the high-frequency nature of the news arrival and policy decisions. Additionally, this model it allows to verify that the orthogonality restrictions hold once we consider the aggregation of a high-frequency information flow model to a higher level of aggregation, as needed within the scope of estimation.

Additionally, we document the performance of the negative control approach empirically in a setting that matches as close as possible the baseline layout proposed in Bauer and Swanson (2023*b*). We highlight the similarity of the negative control estimates as compared to the IV estimates of the reference paper.

Last, we highlight potential improvements and directions for future work. The primary direction for extensions lies in using a more extensive set of measure of Fed policy which would allow for a more precise estimation of the causal coefficients of interest. Another avenue of potential improvement lies in the usage of non-linear methods that dispense with the functional form assumptions for the conditional means. On the modeling side, further work can be done in distinguishing the views of the state of the economy by the Fed and the markets and the explicit modeling of the Fed ”information effect”.

The utility of the present work is manifold. It takes a step in the direction of improving the understanding of the interplay between the monetary policy, financial markets and macroeconomic outcomes. The new perspective on the information flow in the monetary policy context allows to exploit the orthogonality relations that ensue the modeling step. This work is also pushing the frontier of empirical macroeconomics by using a novel identification approach that allows to incorporate the modeling knowledge in the estimation of causal effects and argue for a different use of variables in a fully endogenous system.

**Relation to Literature** Our paper relates to several strands of literature. First, we relate the broad literature on structural identification in the VAR setting in the macroeconomic environment. Starting from is Sims (1980), the literature has expanded with with many different approaches to identification of macroeconomic causal effects and monetary policy in particular. This work is links to classic approaches that attempt to control for the state of the economy directly via several avenues: augmenting the set of controls with low-dimensional summaries of large-dimensional data sets (Bernanke and Boivin, 2003), (Bernanke, Boivin and Eliasz, 2005), (Forni et al., 2009);using large datasets directly via the application of Bayesian techniques (Bańbura, Giannone and Reichlin, 2010), (Bok et al., 2018). Other literature explores the IV approach to identification: Kuttner (2001), Nakamura and Steinsson (2018), Bauer and Swanson (2023*b*). This work explores a novel approach that allows to block the confounding biased induces by the unobserved state of the economy with method of negative controls.

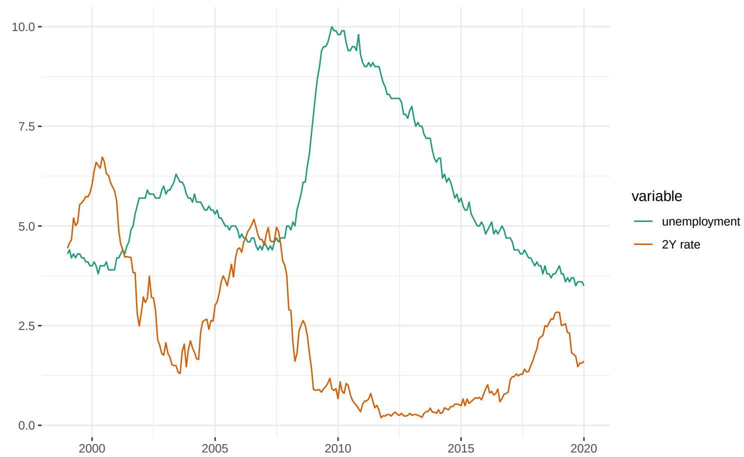
We also relate to the literature analysing the interplay of financial markets and the central bank in the context of monetary policy: Bernanke and Kuttner (2005), Gürkaynak, Sack and Swansonc (2005), Hanson and Stein (2015), Nakamura and Steinsson (2018), Bauer and Swanson (2023*a*). Our approach provides a consistent view of the system of macroeconomic outcomes, news, actions of the Fed and the market where all these variables are endogenous.

Additionally, we relate to the econometric literature on proxy controls and negative controls by elaborating on the linear case and exploring the consequences of applying different estimation approaches in the given conditional independence framework. The negative control approach is exposed in a series of papers coming from the field of biostatistics: Miao, Shi and Tchetgen (2018), Miao, Geng and Tchetgen Tchetgen (2018), Tchetgen et al. (2020). Recent works have extended the approach in applying it to more general identification settings (Singh, 2023), handling a larger number of proxy variables (Deaner, 2021*a*) as well as extending it to the panel data domain (Deaner, 2021*b*).

## 2 Monetary policy identification

The identification issue is best understood by trying to relate the dynamics of the interest rates to the ones of the macroeconomic outcomes of interest. To fix ideas, consider the relation between the 2-year treasury rate and unemployment in Figure 1.

Figure 1: 2Y yield and unemployment




A simple regression of unemployment on the interest rate will provide an estimate that is heavily biased with respect to the causal effect. This regression will recover a strong negative predictive link: a 25 basis point change in the 2Y yield predicts a decrease in the unemployment of 5 basis points. Regressing unemployment on the value of the 2Y rate one year prior recovers a similar negative predictive relation between the two. On the other hand, the expected causal impact of monetary policy, as thought of in theory and estimated by various studies, is never negative. The causal effect should start off near zero and reach a peak value around one year after. As a point of reference, a twenty five basis point increase in the rate should cause a hike of unemployment of the order of 5 basis point or around 100 thousand lost jobs.

The discrepancy between these different estimates is driven by the endogeneity problem. In the simple regression, the error term will be highly correlated with the regressor, the 2Y rate. This is due to a vast array of other determinants of unemployment that also influence the interest rate. Another way to put it is that both unemployment and the interest rates are, to a large degree, driven by a common causes. Such common causes could broadly be labeled as the ”state” of the economy. Fore example, when economy is booming, monetary policy is tighter reflecting the reaction of the Fed to the economic situation and at the same time the unemployment rate is lower reflecting economic growth. The simple regression will reflect this common covariation rather than isolating the causal effect of monetary policy.

In terms of causal effect identification, the endogeneity problem can be framed as omitted variable bias. The key set of omitted variables is the one describing the state of the economy. While several proxies and measures exist, there exists neither a proper definition nor a perfect measure. The absence of a direct measure of the state of the economy exerts a large downward bias in the estimates within a simple regression. The question then is: how does one isolate the influence of the economy state on the estimates of the causal effect of interest rate changes on macroeconomic outcomes?

This question has been offered several types of solutions in the literature to date. Two methods are of particular interest as they tackle the endogeneity problem head on. One approach relies on the assumption of conditional exogeneity and the other uses instrumental variables. As I recap the ideas behind these two methods, I will highlight potential deficiencies and lay ground towards the introduction of the novel identification approach explored in this work.

**Structural VAR approach.** One possible solution to the endogeneity problem is to control for the state of the economy as best as possible in hopes of eliminating the omitted variable bias altogether by employing the conditional exogeneity assumption. This can be done by including in the regression additional variables that represent the proxies of the state of the economy. Given the dynamic nature of the problem and the fact that the unobserved state vector is persistent, the problem can be split into two parts. One can first control for the past dynamics of the state and then add additional controls that capture the state evolution from the past to the contemporaneous value. A common assumption is that controlling for the lagged values of the variables in the system is as good as controlling for the lagged values of the state of the economy. This reasoning justifies the estimation of the vector autoregressive models in such context. Next, there is the question of controlling for the state evolution, that is the discrepancy between the current state and the lagged state value. The classical point of reference here is Sims (1980). The triangularity assumption proposed in the paper implies that adding contemporaneous values of other variables besides the interest rate is sufficient to control for the state evolution component. Summing the arguments together, running a regression of future outcome on the contemporaneous and lagged values of all the variables at hand should identify causal effects of the interest rate changes. In other words, there is an assumption of conditional exogeneity of the interest rate given an appropriate set of controls.

While being the most common approach to causal inference, identification though conditional exogeneity heavily relies on finding the right set of controls. In our context, can one find a set of variables to exactly control for the state of the economy? Many variables can be thought of as proxies of the state. But the assumption that one manages to control for the entire set of mediators between the state of the economy and the interest rate is not easily justifiable. The literature has expanded with several iterations on the basic conditional exogeneity idea: augmenting the set of controls with low-dimensional summaries of large-dimensional data sets (Bernanke and Boivin, 2003), (Bernanke, Boivin and Eliasz, 2005), (Forni et al., 2009);using large datasets directly via the application of Bayesian techniques (Bańbura, Giannone and Reichlin, 2010), (Bok et al., 2018). Still, conditional exogeneity represents the weak link in the causal effect identification reasoning.

**Instrumental variable approach.** Another type of approach employed in the literature relies on instrumental variables. As with the previous approach, there is an exogeneity assumption at play. Instead of conditional exogeneity of the interest rate, one seeks for a variable that is predictive of the interest rate and exogenous in its own right. The first uses of the idea go back to Kuttner (2001), where high-frequency changes in the interest rates around the FOMC announcements were used as instruments for the interest rate changes at a lower frequency. Others have expanded on the idea and refined the approach by perfecting the estimation of the instruments and exploring the interplay between the Fed and the markets in the monetary policy context: (Gürkaynak, Sack and Swanson, 2005), (Nakamura and Steinsson, 2018) ,(Bauer and Swanson, 2023*a*).

Despite the advancements in the direction of IV identification, some problems still persist. Exogeneity assumption of high-frequency instruments has been recently challenged in Bauer and Swanson (2023*b*). The instrument strength is under strain as well – the vast majority of the interest rate variation seems to come from other sources beyond the FOMC announcement driven changes[1]. Thus, the IV approach has its own set of issues that don't have a direct solution as they represent the inherently untestable part of the causal inference assumptions and limitations related to the data.

## 2.1 Motivation of negative control

This paper attempts to tackle the endogeneity problem from a different angle. While the previous two approaches rely on exogeneity assumptions to guide inference, the negative control approach employed in this paper changes perspective by moving the focus to the endogenous variation. The idea underlying this novel approach is that by fully exploring the endogenous variation in the interest rates, one can isolate the exogenous part that drives causal inference. To fully understand the workings of the approach, it is best to proceed by introducing the necessary components one at a time.

To fix ideas, we are interested in the causal effect of a change in a market interest rate (denote it $i_t$) of a medium maturity, e.g. the 2Y treasury rate, on a set of future macroeconomic variables including unemployment and prices (we denote a variable of this type $Y_{t+h}$). The unobserved confounders in this case will correspond to all the unobserved factors influencing the market rates as well as macroeconomic developments and will be called the state of the economy, or simply state and denoted as $S_t$. As one looks at the relation between the market rates and the macroeconomic outcomes without any additional information no progress can be made. In making the next step we will add the Fed policy as an additional entry in the system and argue of its novel use through the lens of the negative control approach.

While the existing literature rarely distinguishes between the Fed policy and market rates of medium maturity, this distinction will be one of the pillars of the negative control approach developed in this work. Recent studies like Nakamura and Steinsson (2018) pool the 2Y yields

[1] Bauer and Swanson (2023*b*) tries to enrich the instrument variability by considering the Fed chair statements in addition to FOMC announcements and the source of exogenous high-frequency variation.

with other shorter yields together in defining the policy instrument. While there is a direct link between the policy and the market rates, there is a net distinction between them. Market interest rates are directly controlled by the financial market participants. While the large part of the variation in the market rates comes from gauging the broad economic conditions and state of the economy as well as anticipating the central bank reaction to such developments, some of the variation is due to the idiosyncratic factors that come from the market interactions. As far as typical monetary policy instruments go, the Fed can only directly control the overnight interest rates and the communication about its future course of actions. As such, it has a degree of influence on the market rates but not the ability to set them to desired values outright. We just argued for the following assumption:

$$i_{0,t} \perp\!\!\!\perp Y_{t+h} | S_t, i_t \tag{1}$$

Fed policy, and fed funds rate in it's simplest form, is definitely a driver of market interest rates. There are many potential channels at play: the pure expectation channel (influencing the expected future rates), the information channel (influencing the market views of the state of the economy) or the fact of the resolution of uncertainty regarding Fed's reaction function. Yet, the Fed policy has arguably no direct causal effect on macroeconomic outcomes conditional on the market interest rates. Besides this, Fed policy is reflective of the state of the economy as it is a direct consequence of the central bank's views of the current state of the economy. These features will be accounted for when building the identification argument in the negative control setting. In particular, these links will allow to produce a proxy of the confounding component of the system by instrumenting for the endogenous variation. This objective cannot be achieved with the commonly known identification methods.

In order to make progress, the negative control approach will require the addition of a variable that will allow to capture the endogenous variation in both the Fed policy and the market rates. In particular, one needs to find a variable that acts as a proxy of the state but is not directly related to neither policy measures not the market interest rates. If there is a variable that only directly relates to the state of the economy and, conditional on the state, does not correlate with the market rates and Fed policy, we can devise a mechanism to purge the endogenous variation. The role of such a variable can be played by macroeconomic news as measured by the discrepancy between the data release and the market expectation just prior to the release, denote such variable $W_t$. Below we argue that for economic news variable the following exclusion restriction holds

$$(i_{0,t}, i_t) \perp\!\!\!\perp W_t | S_t \tag{2}$$

Macroeconomic news do fit the desired role of being pure proxies of the state. While the news are informative of the economic state for both markets and the central bank, this information channel is arguably the only channel of influence between these sets of variables. For example, consider a situation when the inflation data turn out higher than previously anticipated. This fact prompts the Fed and the markets to reconsider their expectations about the state of the economy that is likely more lievely than previously though. These updates on the views of the economy prompt the reaction both from the policy side (hiking rates, communicating a future path for the rates) and the market reaction (taking the state changes and the policy changes into account) setting medium and long maturity interest rates. As long as this channel is the only one that links the news to the actions of the policymaker and the market reaction, there will be a way to exploit this additional proxy of the state. With these properties in mind, how can one then exploit macroeconomic news data for causal inference given this specific dependence structure?

Negative control identification approach elaborates on the idea of constructing a better proxy of the state by exploiting the endogenous variation in the additional variables described above. In particular, it will aim at predicting the endogenous part of the macroeconomic news using the variation in the market interest rates and the central bank policy measures. This predicted endogenous component will be used as a control and, as is shown in the next section, will block the omitted variables bias stemming from the unobserved confounder, the state of the economy.

Summarising the approach so far, we have highlighted two components, Fed policy and the macroeconomic news, that respect certain orthogonality conditions but are not directly suitable for causal inference in the commonly known approaches. The negative control identification framework will make use of these variables in a novel way that provides identification of the causal effect of the market rates on macroeconomic outcomes.

Consider the restrictions mentioned before. The restriction that Fed policy does not affect the future outcomes conditional on the state of the economy and the market interest rates is given in equation (1). The fact that the macroeconomic news are only correlated to the Fed policy and the market rates through the state of the economy is inscribed in equation (2).Armed with these two orthogonality conditions, negative control identification approach provides identification and inference guidelines. This approach has been recently developed in the biostatistical literature (Tchetgen et al. (2020), Miao, Geng and Tchetgen Tchetgen (2018)) and recently expanded upon in several directions (Deaner (2021*b*), Deaner (2021*a*), Singh (2023)). To our knowledge this is a first application of the negative control approach in the monetary policy identification setting.

The theoretical underpinning of the approach will be discussed in the next section, but it is important to provide an intuition behind the functioning of the approach. Conditioning on macroeconomic news in this context would be insufficient to eliminate the omitted variable bias due to the fact that it is an imperfect proxy. However, if we condition on the value of news predicted by the Fed policy measures and market interest rates, we will be extracting the variation in news $W$ that is only due to the variation in the unobserved confounder $S$, expectation of the state of the economy. Therefore, conditioning on the predicted value of $W$ will be the same as conditioning on the unobserved $S$, leading to correct causal inference.

Overall, the negative control approach is both similar and different from the conditional exogeneity and the instrumental variable approaches. First key distinction is that there are no (conditional) exogeneity restrictions involved. Second, instead of focusing on the exogenous part of the variation in the variable of interest, the conceptual focus of the approach is flipped: one attempts to extract the endogenous variation from a proxy variable to be used as a control. In the setting at hand, the goal is to extract the endogenous component of the macroeconomic news and use it as a control in a regression of the macroeconomic outcomes on the market interest rates. This change in paradigm allows to explore the question of monetary policy identification from a completely different angle. Instead of focusing on the exogenous variation, one can focus on trying to capture the all the endogenous variation sources. While this approach is not widely known, we will show a simple way to apply the idea and compare this approach to the existing ones through simple examples.

Next section will discuss all the relevant assumptions and the econometric details of the negative control identification approach, while Section 4 will present a simple state state model that implies the conditional independence structure underpinning the identification.

# 3 Negative control approach

The question of interest will be the identification of the causal effect of the treatment $D_t$ on the outcome $Y_{t+h}$ in the presence of the unobserved confounder $S_t$. In order to make the approach operational one needs two types of variables with particular characteristics. In essence, the first type of the variable, negative control treatment $Z_t$, should induce the variation in the outcome only through the treatment and be related to the unobserved confounder. The second type of variable, negative control outcome $W_t$, should essentially be a proxy of the confounder $S_t$. In particular, it has to be correlated with the treatment $D_t$ and NC treatment $Z_t$ only through $S_t$. While the notation is changed for this section for the sake of better exposition, the variables will directly mirror the ones we use in the monetary policy analysis: $S_t$ will correspond to the expected state of the economy; $Y_{t+h}$ be the macroeconomic outcome of interest; $D_t$ be a market interest rate of medium maturity; $Z_t$ will be a measure of Fed policy, such as the fed funds rate; and $W_t$ will be the macroeconomic news.

Altogether, we consider a system of five variables $H_t = (S_t, D_t, Z_t, W_t, Y_{t+h})$. In the following we construct the identification argument while stating the assumptions necessary for the results to apply.

While the negative control framework has been developed for inference in nonlinear cases, the present work focuses on the linear case. The following equations states the linearity assumption for the relevant conditional means. The linearity restriction will only involve two conditional means because only these functions will be directly involved in constructing the identification argument.

**Assumption 1** (Linearity)**.** *The following conditional means are linear:*

$$\begin{aligned}\mathbb{E}[Y_{t+h}|D_t, Z_t, S_t] =& \beta_0 + \beta_d D_t + \beta_z Z_t + \beta_s S_t &(3)\\ \mathbb{E}[W_t|D_t, Z_t, S_t] =& \eta_0 + \eta_d D_t + \eta_z Z_t + \eta_s S_t &(4)\\ \mathbb{E}[S_t|D_t, Z_t] =& \delta_0 + \delta_d D_t + \delta_z Z_t &(5)\end{aligned}$$

The decision to restrict the conditional means to be linear is driven by the desire to ease the exposition and facilitate intuitive understanding. It allows for a clear objective that is easily comparable to the existing literature: the average treatment effect $\beta_d$. Moreover, within the context of monetary policy analysis, the short data span of the relevant variables does not warrant the use of non-linear methods. Going beyond the linear case is left to future work and represents a straightforward conceptual extension.

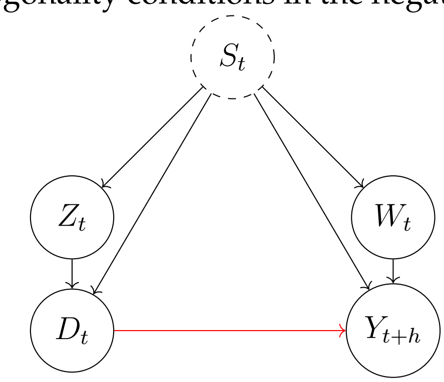
Next, we turn to the orthogonality conditions that drive the identification argument.The orthogonality conditions are one of the necessary parts to the identification result and their use is conceptually close to that of the instrumental variables orthogonality conditions. In the IV case the exclusion restriction serves to extract the exogenous variation from the variable of interest. The NC approach will instead use the exclusion restrictions to instrument for the endogenous component of the proxy $W_t$. Two orthogonality conditions are required:

**Assumption 2** (Orthogonality)**.** *The following orthogonality conditions hold:*

$$Y_{t+h} \perp\!\!\!\perp Z_t \mid D_t, S_t \tag{6}$$
$$W_t \perp\!\!\!\perp (D_t, Z_t) \mid S_t \tag{7}$$

In relation to Assumption 1, Assumption 2 implies that several coefficients are null, namely $\beta_z = \eta_d = \eta_z = 0$. While the equations above succinctly describe the needed restrictions, a graphical representation is also useful to transmit the underlying logic. Below is a prototypical graph that inscribes the exclusion restrictions of the NC approach. Each node represents a random variable. Observed variables are circled with solid lines, while the unobserved confounder $S$ is circled with a dotted line. The directed edges on the graph correspond to causal links between the variables and the red line corresponds to the causal effect of interest.

Figure 2: Orthogonality conditions in the negative control setup




Importantly, orthogonality restrictions are represented by the absence of certain links from the graph. In particular, condition 6 from above implies that the there is no direct link $Z_t \to Y_{t+h}$. Condition 7 implies that there are no direct links $Z_t \to W_t$ and $D_t \to W_t$.

The other key component is semantically close to the notion of instrument relevance. This second piece involves a necessary set of assumptions regarding the properties of joint distributions of some of the variables in the system. In the general theory of negative controls handling a nonparametric setting such assumption involves the usage of completeness conditions. As we focus of the linear case, this greatly simplifies exposure by restricting the focus to projection matrix rank conditions.

Using the linearity introduced by Assumption 1, we note that the conditional mean $\mathbb{E}[W_t|D_t, Z_t]$ is also linear. Building off this, together with the restrictions of Assumption 2, we introduce the rank conditions necessary for identification.

**Assumption 3** (Completeness)**.** *Let Assumptions 1 and 2 hold. Consider the following regressions:*

$$\mathbb{E}[S_t|D_t, Z_t] = \delta_0 + \delta_d D_t + \delta_z Z_t$$
$$\mathbb{E}[W_t|D_t, Z_t] = \gamma_0 + \gamma_d D_t + \gamma_z Z_t$$

*The following rank conditions are assumed to hold:*

$$rank(\delta_z) \geq dim(S_t) \tag{8}$$
$$rank(\gamma_z) = dim(Z_t) \tag{9}$$

Maintaining the comparison with IV, intuitively, Assumption 3 requires $Z_t$ to be a relevant instrument for $W_t$ while it also ensures that the instrumentation does not lose the informational content of $S_t$, all conditional on $D_t$. In particular, inequality (8) implies that $Z_t$ does not lose information from $S_t$. Inequality (9) ensures sufficient dimensionality and variability of $W_t$ as compared to $Z_t$ conditional on $D_t$. It is important to mention that $Z_t$ does not have to perfectly predict $S_t$ but rather be sufficiently informative of it. By the same token, $W_t$ has to be informative of $Z_t$, conditional on $D_t$. The exact usage of these conditions is explained in the discussion part of the section.

The three assumptions above are sufficient to state the negative control identification result, which is based on the replication of Theorem 1 (Miao, Geng and Tchetgen Tchetgen, 2018) for the linear case:

**Proposition 1** (Negative control identification)**.**
*Consider a set of random variables $(Y_{t+h}, D_t, Z_t, W_t, S_t)$ that satisfy assumptions (1)-(3). Then $\beta_d$ is identified by the following regression:*

$$\mathbb{E}[Y_{t+h} \mid D_t, \mathbb{E}[W_t \mid D_t, Z_t]] = \beta_0^* + \beta_d D_t + \beta_s^* \mathbb{E}[W_t \mid D_t, Z_t] \tag{10}$$

In words, we are able to identify the causal effect of $D_t$ on $Y_{t+h}$ by controlling for the value of $W_t$ that is instrumented by $Z_t$ and uses $D_t$ in place of an exogenous variable.

While the identification argument was originally made for the generic non-linear case, we will expose the linear case that is of most interest in the macroeconomic context given the availability of the data and the need for a clear interpretation. The case of the linear conditional means allows to greatly simplify exposure by restricting the necessary focus to conditional means and projection coefficients.

Below we present the part of the proof that transmits the logic of the approach. The first step is to employ the restrictions of Assumption 2 in the specification described by Assumption 1. Assumption 2 implies that several coefficients of the unrestricted specification are equal to zero, simplifying to the following:

$$\begin{aligned} \mathbb{E}[Y_{t+h} \mid D_t, Z_t, S_t] &= \beta_0 + \beta_d D_t + \beta_s S_t \\ \mathbb{E}[W_t \mid D_t, Z_t, S_t] &= \eta_0 + \eta_s S_t \end{aligned} \tag{11}$$

Next, we drop the unobserved confounder $S_t$ from the conditioning set:

$$
\begin{aligned}
\mathbb{E}[Y_{t+h} \mid D_t, Z_t] &= \beta_0 + \beta_d D_t + \beta_s \mathbb{E}[S_t \mid D_t, Z_t] \\
\mathbb{E}[W_t \mid D_t, Z_t] &= \eta_0 + \eta_s \mathbb{E}[S_t \mid D_t, Z_t]
\end{aligned} \tag{12}
$$

By noticing the fact that the right hand side of both equations includes a common unidentified term, the last step involves substituting away the unobserved confounder mean $\mathbb{E}[S_t \mid D_t, Z_t]$. The inversion performed in this step relies on Assumption 3 and will be discussed in further detail later on.

By following these steps we arrive to an expression involving identifiable means and the causal coefficient of interest $\beta_d$:

$$
\mathbb{E}[S_t \mid D_t, Z_t] = (\eta_s' \eta_s)^{-1} \eta_s' (\mathbb{E}[W_t \mid D_t, Z_t] - \eta_0) \tag{13}
$$

$$
\mathbb{E}[Y_{t+h} \mid D_t, Z_t] = \beta_0^* + \beta_d D_t + \beta_s^* \mathbb{E}[W_t \mid D_t, Z_t] \tag{14}
$$

where

$$
\begin{aligned}
\beta_0^* &= \beta_0 - (\eta_s' \eta_s)^{-1} \eta_s' \beta_s \eta_0 \\
\beta_s^* &= (\eta_s' \eta_s)^{-1} \eta_s' \beta_s
\end{aligned}
$$

This concludes the exposition of the negative control approach. We can observe that the practical side of the approach uses the same methodology as the instrumental variable approach. In IV implementation terms, we instrument for $W_t$ using $Z_t$ as instrument and using $D_t$ as an exogenous variable. Although the application is similar to IV, the logic behind it is quite distinct. None of the variables in the system are in fact exogenous. Instead, by instrumenting $W_t$, we recover the endogenous component that blocks the omitted variable bias coming from the unobserved confounder.

## 3.1 Estimation

The estimation can be carried out via a standard application of the Generalized Method of Moments (GMM). Let us first consider the simpler case when we have the same number of NC outcomes and NC treatments and all the variables are mean zero. In this simplest case, the identification arguments suggest the following set of moments:

$$
\mathbb{E}[(Y_{t+h} - \beta_d D_t - \beta_s^* W_t)\, [D_t', Z_t']] = 0 \tag{15}
$$

This leads to the following estimator:

$$\hat{\beta} = \arg\min_{\beta} \sum_t g(H_t, \beta) g(H_t, \beta)'$$

for $g(H_t, \beta) = (Y_{t+h} - \beta_d D_t - \beta_s W_t)\,[Z_t', X_t']$. Under the typical GMM assumptions this estimator provides consistent inference for $\beta_d$ as well as allows to characterise the asymptotic distribution:

$$\sqrt{n}(\hat{\beta} - \beta) \rightarrow \mathcal{N}(0, \Omega)$$

where $\Omega = G^{-1}\Sigma G^{-1'}$, $G = \mathbb{E}\left[\frac{\partial g(H_t,\beta)'}{\partial \beta}\right] = \begin{bmatrix} \Sigma_{dd} & \Sigma_{dz} \\ \Sigma_{wd} & \Sigma_{wz} \end{bmatrix}$ and $\Sigma = \sum_{-\infty}^{\infty} \mathbb{E}\left[g(H_t, \beta) g(H_{t-j}, \beta)'\right]$ is the long-run variance. Given the fact that we will be dealing with time series, the potential autocorrelation of the errors prompts for the use of the appropriate covariance estimator. In the practical application we will use parametric bootstrap standard errors.

Note that this simpler setup assumes the system is exactly identified. In particular, it implies that NC outcomes $W_t$ have the same dimensionality of NC treatment $Z_t$. While this will represent the baseline estimation case, we will also consider the application where $dim(W_t) > dim(Z_t)$, which is elaborated in the following discussion part of this section.

## 3.2 Discussion

The negative control (NC) approach is an identification strategy that hinges on a set of conditional independence assumptions involving the treatment, outcome and additional observed variables. This orthogonality structure is used to purge the OVB stemming from the unobserved confounder. In contrast with the conditional exogeneity approach, no conditional exogeneity assumptions are made or needed. Rather, the negative control approach focuses on the endogenous variation directly. The orthogonality conditions warrants a construction of a control variable that allows for unbiased inference.

The origins of the negative control approach lie in the biostatistical and epidemiological literature[2]. The initial use of the idea was to detect confounding bias in both experimental and observational settings. The terminology derives from the fact that negative controls are variables that must be causally unrelated, therefore "negative". Regressions testing this dependence should produce estimates of no association. In particular, two groups are distinguished: NC treatments and NC outcomes. Negative control treatment are the variables that should not cause the outcome directly. That is, if the confounding has been successfully handled, the regression of the outcome on the variable of interest including the NC treatment should have a zero coefficient for the latter. Following a similar idea, the negative control outcomes are the variables that should not be causally related to both the treatment and the negative control treatment. In a

[2]See Lipsitch, Tchetgen and Cohen (2010), Lousdal et al. (2020)

regression with no confounding, relating NC outcome to both treatments should produce zero coefficients for the two regressors.

The earlier uses of the negative control ideas were confined to the detection of potential confounding. The use of negative control ideas for eliminating confounding is a recent development. Miao, Geng and Tchetgen Tchetgen (2018) and Miao, Shi and Tchetgen (2018) introduce the idea, prove the identification in the non-linear case and provide examples. Recent work generalizes the approach to handle many negative control outcomes (Deaner, 2021*a*), panel data (Deaner, 2021*b*) as well as generalizing the identification result (Singh, 2023). The current work explores the linear case in the novel setting and sheds some light on the comparison of the negative control approach to the classical approaches from the perspective of unbiased inference.

The negative control has several advantages and disadvantages with respect to the conditional exogeneity approach. As for the advantages, the NC approach does not put any restrictions on the distribution of the unobserved confounder, nor does it require its identification. Besides, one does not need the conditional exogeneity of the variable of interest given the controls. While conditional exogeneity approach implies the inclusion all mediators of the confounder, the negative control approach only requires that the additional variables are informative of the confounder. Concessions have to be made to counterbalance the upsides of the approach and the relaxed assumptions. On the cost side, NC heavily exploits the conditional orthogonality conditions by restricting certain conditional dependence structures. In particular, additional two sets of variables respecting the orthogonality conditions need to be introduced in the basic framework. These additional variables will serve the purpose of creating a proxy for the unobserved confounder by extracting the endogenous variation component. The resulting proxy will block the confounding effect and warrant unbiased causal inference. Overall, the NC approach can be viewed as a refinement over the conditional exogeneity approach when additional structural information about the problem at hand is available and one can exploit the expert judgement of the problem to drive inference.

We next discuss some of the features of the NC approach. By focusing on the comparison with existing methods, we will provide additional intuition and put the approach into perspective. We also discuss in more details the role of completeness assumptions in the identification argument. Additionally, we confront the case when nuisance parameters are not uniquely identified in case more NC outcomes are used in estimation.

**NC versus conditional exogeneity**

One of the first questions one might have in facing the negative control approach is the comparison with the most common practice of simply including all the additional variables as controls. Under the dependence structure restrictions described in Assumption 2 the conditional exogeneity of $D_t$ does not hold and therefore such regression will produce biased results. To see the exact bias expression we can further analyse the linear case. Consider the following simplified linear system with no NC treatment $Z_t$:

$$
\begin{aligned}
W_t =& \varepsilon_t^W \ , \ \varepsilon_t^W \sim (0, \Sigma_w) \\
S_t =& K W_t + \varepsilon_t^S \ , \ \varepsilon_t^S \sim (0, \Sigma_s) \\
D_t =& G S_t + \varepsilon_t^D \ , \ \varepsilon_t^D \sim (0, \Sigma_d) \\
Y_{t+h} =& \beta_d D_t + \Gamma S_t + \varepsilon_t^Y \ , \ \varepsilon_t^Y \sim (0, \Sigma_y)
\end{aligned}
$$

A regression of $Y_{t+h}$ on $D_t$ and $W_t$ yields

$$
\mathbb{E}[Y_{t+h}|D_t, W_t] = \beta_d D_t + \Gamma \mathbb{E}[S_t|D_t, W_t] = \tilde{\beta}_d D_t + \tilde{\beta}_w W
$$

where

$$
\tilde{\beta}_d = \beta_d + \Gamma \Sigma_s G' \left( G \Sigma_s G' + \Sigma_d \right)^{-1}
$$

which implies that the bias still remains if we add $W_t$ as a control unless $W_t$ is a perfect proxy of $S_t$. It can be shown that the bias is lower than in the case of no added controls and highlights the sources of bias difference. Including $Z_t$ in this regression would complicate the expression but the main logic will go through – conditioning on $W_t$ is beneficial in terms of bias reduction but it does not eliminate the bias.

**NC versus IV**

As noted in Miao, Shi and Tchetgen (2018), the NC approach can also be seen as a generalization for the instrumental variable procedure. In the simplest case of the relevant variables being unidimensional, the expression for the coefficient of interest becomes

$$
\hat{\beta}_x = \frac{\sigma_{xw}\sigma_{zy} - \sigma_{xy}\sigma_{zw}}{\sigma_{xw}\sigma_{zx} - \sigma_{xx}\sigma_{zw}}
$$

The authors proceed to show that the approach possesses a robustness property: the NC estimate of the causal effect is unbiased if either the NC restrictions specified in the beginning of the section are correct or the variable $Z$ is a valid instrument for $X$. This is important in view of the relation between my work and that of Bauer and Swanson (2023*b*).

**On the relation to the measurement error literature**

The negative control idea is closely linked to the concept of unbiased inference in the case of measurement error. A review of the methods can be found in Schennach (2012) while the linear case is exposed in Stefanski and Buzas (1995). One of the ways to address the mismeasurement of the outcome of interest is the use of the instrumental variables that are related to the outcome of interest but orthogonal to the (mis)measurement mechanism. Similar to the negative control approach, one attempts to construct a measure of the unobserved variable of interest in order to use it in a regression against the outcome. The difference between the two methods lies in the focus of the identification. In measurement error problems, the unobserved component is the variable of interest itself while the negative control approach instruments for the unobserved confounder.

**On completeness**

The explanation of the workings of the negative control approach provided at the beginning of the section provided the intuition but was muted on the exact use of the completeness conditions. At first glance it might seem like the invertibility of the $\eta_s'\eta_s$ matrix in equation (13), which only requires that the rank of $\eta_s$ is greater or equal to the dimension of the space $\mathcal{S}$, is generically satisfied as far as the variable dimensions are met, that is $dim(\mathcal{W}) \geq dim(\mathcal{S})$. However, the proper condition involves multiple projections (thinking of $\eta_s$ as first orthogonalizing $W$ and $S$ with respect to $Z$ and then projecting $W$ on $S$) and thus requires a more refined set of restrictions. Appendix A elucidates the instances of the usage of completeness conditions.

# 4 Model

This section will provide additional arguments in favor of the orthogonality restrictions that underlie the negative control identification approach. With that in mind, the first part of the section will introduce a state-space model that will help guide the economic intuition and be justify the needed orthogonality restrictions. While the simple model will help convey the essential logic, additional empirical arguments will provide further support to the orthogonality restrictions.

## 4.1 State space model

A state space model can provide a clear, although simplified, account of the economic interactions that the present work aims to exploit. In the following we will present a state space model that will pursue this goal as well as justify the set of orthogonality restrictions employed by the NC approach. While the proposed state space model is fairly basic at heart, several features will render it distinct among the literature counterparts. First, the market interest rates will be introduced as a control variable in the state equation. This will reflect the idea that the market rates have a direct causal effect on the future evolution of the economy. Second, we will make a distinction between the information that is observed by the market participants and the part of it that is available to the econometrician. Third, we will introduce the response of the Fed and the response of the market in a cohesive way while maintaining the distinction between the two. The Fed will have a direct effect on the market interest rates but no direct effect over the macroeconomic dynamics, as conceived in the introduction.

The goal of this subsection is to merge the economic intuition behind the causal connections with a simple structural model. This is indented to clarify the concepts in play, shed light on the causal links and share the perspective through which the orthogonality conditions can be understood.

The model falls in line with many classical applications of the empirical state-space approaches in macroeconomics. Examples include Ang and Piazzesi (2003), Diebold, Rudebusch and Borag˘an Aruoba (2006), Gürkaynak and Wright (2012), and Bauer and Rudebusch (2020). The present model will also borrow ideas on the interplay between the Fed and the financial markets investigated in Romer and Romer (2000), Piazzesi (2005), Lakdawala and Schaffer (2019), Jarociński and Karadi (2020), Karnaukh and Vokata (2022), Schmeling, Schrimpf and Steffensen (2022), Caballero and Simsek (2022) and Hillenbrand (2022).

The proposed model should be viewed from the perspective of daily time intervals. While the eventual empirical application will use monthly data, the present model will be used for higher frequency structural reasoning and, via aggregation to the monthly level, will justify the relevant orthogonality conditions.

The heart of the model is a Markov state space model whereby the state of the economy $S_t$ is dynamically controlled by the market yield curve $(i_{m,t})_m$, where $m > 0$ stands for maturity. The observation equation is most typical with the data $Y_t$ being generated using fixed loadings and the state $S_t$.

$$S_t = GS_{t-1} + \alpha(L)(i_{m,t})_m + u_t \tag{16}$$
$$Y_t = FS_t + v_t \tag{17}$$

For simplicity, we assume that both the error terms are normally distributed, which allows the application of the usual Kalman filtering reasoning. Using this reasoning, we can state that the state expectations, the unobserved confounder in the NC reasoning, evolves over time using the logic of Kalman updating. Thus, every period the state expectations are updated by the Fed and the market:

$$S_t^e = S_{t-1}^e + K(Y_t - \mathbb{E}[Y_t|\mathcal{I}_{t-1}]) = S_{t-1}^e + KW_t \tag{18}$$

where $K$ stands for the Kalman gain term and $\mathbb{E}[Y_t|\mathcal{I}_{t-1}]$ is the discrepancy between the previously expected and actually realised macroeconomic values. We will denote this discrepancy $W_t$ and refer to is as the macroeconomic news.

The crux of the problem is that only some of the macroeconomic news are observed by the econometrician. Therefore, even under the assumption of correct model specification, the econometrician would not be able to perfectly predict the Fed and the market's view on the state of the economy. This assumption is easily defendable from the practical standpoint given the limited data availability. However, one should also keep in mind that the present model is only intended to capture the essence of the problem. In the real world other issues, such as opinion aggregation, information delays etc could affect the consensus estimate. Thus, the assumption of the inability to fully control for the state expectation seems fairly plausible in the context at hand. Next we turn to how the state expectations are acted on by the markets and the Fed.

The Fed acts upon it's views of the state of the economy via several channels. First, it sets the fed funds rate as a function of the current state expectations:

$$i_{0,t} = \left(g_t(S_t^e) + \varepsilon_t^0\right) \mathbb{1}\{Fed_t = 1\} + i_{0,t-1}\mathbb{1}\{Fed_t = 0\} \tag{19}$$

where $g(\cdot)$ is the reaction function of the central bank describing how it reacts to the perceived state of the economy and the policy is only updates in time periods when the Fed takes an action, $Fed_t = 1$.

Second, the Fed can affect the longer term yields by communicating its plans about the future path of the fed funds rate, the so-called forward guidance. This can be viewed as a change in the prospected reaction function for some period $h$ steps ahead: $g_{t+h}(\cdot)$.

Next, we introduce the reaction of the markets. The markets set the yield curve $(i_{m,t})_m$ by considering the central bank policy as well as the idiosyncratic factors.

$$i_{m,t} = \alpha_m + \frac{1}{m}\sum_{j=0}^{m-1} i^e_{0,t+j} + \varepsilon^m_t \tag{20}$$

$$i^e_{0,t} = i_{0,t} \tag{21}$$

$$i^e_{0,t+j} = g^e_{t+j}(S^e_{t+j|t}) = g^e_t(G^j S^e_t) \tag{22}$$

where $\alpha_m$ is a potential term premium term, $i^e_{0,t+j}$ is the expected future spot rate and the reasoning follows the expectation hypothesis of the interest rate yields. Given the dynamic structure of the state, the expected future short term rates $i^e_{0,t+j}$ are thought of as results of the application of the Fed's expected future response function $g^e_{t+j}(\cdot)$to the expected future state of the economy. The last equation is a simplification assuming that only the expected mean of the of the future state would be relevant for decision making.

The last modeling step regards the time aggregation. As news and policy actions happen in specific time instances, it is important to make sure that the aggregated dynamics of the system respect the orthogonality conditions of the negative control approach. As the empirical application will focus on monthly data, we will reflect this idea in the aggregation exposition although the concept can be applied to any aggregation level structure.

Consider the following time aggregation structure. Let $t$ be the higher frequency (daily) time label. Let $\tau$ and $\tilde{\tau}$ be the low frequency (monthly) time label. The $\tau$ index will correspond to simply taking the last value of the underlying time period while $\tilde{\tau}$ will aggregate the data by taking the value of the variable corresponding to the time period of the action of the Fed within the higher frequency window. The following equations clarify the aggregation schedule:

$$Y_\tau = Y_t \tag{23}$$

$$Y_{\tau+1} = Y_{t+31} \tag{24}$$

$$S^e_{\tau+1} = S^e_\tau + KW_{\tau+i} \tag{25}$$

$$i_{\tau+1} = i_{t+31} \tag{26}$$

$$W_{\tau+1} = \left(\sum_{i=1}^{31} \mathbb{1}\{O^j_{t+i} = 1\} W^j_{t+i}\right)^J_{j=1} \tag{27}$$

$$S^e_{\tilde{\tau}+1} = \sum_{j=1}^{31} \mathbb{1}\{\mathit{Fed}_{t+j} = 1\} S^e_{t+j} + S^e_{\tilde{\tau}} \mathbb{1}\left\{\left(\sum_{j=1}^{31} \mathit{Fed}_{t+j}\right) = 0\right\} \tag{28}$$

$$i_{\tilde{\tau}+1} = \sum_{j=1}^{31} \mathbb{1}\{\mathit{Fed}_{t+j} = 1\} i_{t+j} + i_{\tilde{\tau}} \mathbb{1}\left\{\left(\sum_{j=1}^{31} \mathit{Fed}_{t+j}\right) = 0\right\} \tag{29}$$

where $O^j_{t+i}$ is the indicator variable that denotes whether the news variable $j$ in time $t+i$ is observed [3].

The claim is that within the model framework delineated above, the relevant orthogonality restrictions hold true.

**Proposition 2.**
*Under the state-space model specified in equations 16-22, the following orthogonality conditions hold*

$$i_{0,\tilde{\tau}} \perp\!\!\!\perp Y_{\tau+h} | S^e_{\tilde{\tau}}, i_{\tilde{\tau}}$$
$$(i_{0,\tilde{\tau}}, i_{\tilde{\tau}}) \perp\!\!\!\perp W_\tau | S^e_{\tilde{\tau}}$$

*once the variables are aggregated to the higher frequency.*

Let's first focus on the orthogonality condition involving the orthogonality of the fed funds rate and the future output conditional on the expected state of the economy and the market interest rate. The key causal chain that links the two is the following: $i_{0,\tilde{\tau}} \to (i_{m,\tilde{\tau}})_m \to S_{\tau+h} \to Y_{\tau+h}$. That is, the fed funds rate, which is an overnight rate, influences the market rates which in turn influence the state evolution up to period $h$ which in turn causes an impact in the future outcome $Y_{\tau+h}$. The influence on the market rates is thought the effect on the shortest end as well through the updates in the state expectations that happen in between the observation period. As the fed funds is decided based on the concurrent state expectations $S^e_\tau$ with only an orthogonal discretionary component at hand, there is no other significant source of correlation between the fed funds rate and the future macroeconomic outcomes[4].

The second orthogonality condition states that the market interest rates together with the fed funds rate are independent on the news conditional on the state expectations $S^e_{\tilde{\tau}}$. On the surface the relation is simple once you consider the contemporaneous high-frequency structure. In fact, if one considers the relation between the variables within the period $t$, as the news only affect the reaction of the market and the Fed through the state expectations. Plus, the news are only related to other past variables and the current value of the true state while the yield curve then only goes to affect the variables in the next period. This simple reasoning is undermined by the fact of temporal aggregation. As we aggregate to the lower frequency, the news variable will collect the all the news that were observed throughout the interval between the lower frequency time stamps. For example, suppose that only news series $j$ was observed and it happened in $t=1$, while we then use other variables as they appear at the end of the higher frequency cycle, at $t=30$. Then we would need to make sure that the following orthogonality condition holds: $(i_{0,\tilde{\tau}}, i_{\tilde{\tau}}) \perp\!\!\!\perp W_\tau | S^e_{\tilde{\tau}}$. This conditional independence relation will also hold in this model. It is easier to first look at the more general case. The state expectations only evolve with the arrival of

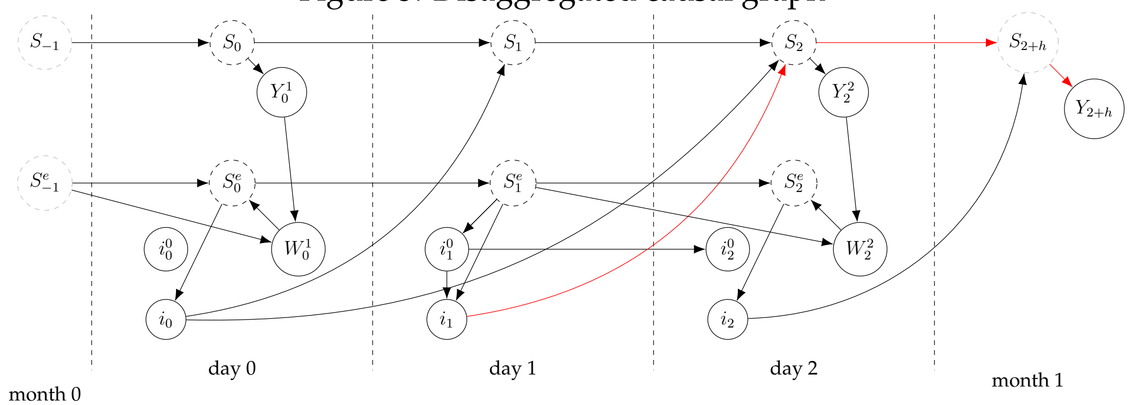
[3]For simplicity we explicitly assume that the Fed can only act once in the lower aggregation level, but we can also accommodate the case when multiple actions are taken within the aggregation period by for example considering taking only the last value.

[4]There might be a minor amount of correlation guided by the influence of the interest rates on the state of the economy and therefore macroeconomic data within the aggregation time frame that is represented in the data. However, this effect should be negligible on the real and nominal macroeconomic within short time windows.

macroeconomic news. Thus, conditioning on the policy period state expectations is the same as conditioning on all the intermediate state expectations as the only difference between the two is the vector of news. If we additionally consider the fact that the policy is either perpetuated from the past or is a function of the state values for the periods in which it is enacted, conditioning on the state value of the policy time period is sufficient to block the correlation between the policy measure and all the macroeconomic news.

The graphical representation of an example of the information flow in the model is provided in Figure 3.

Figure 3: Disaggregated causal graph

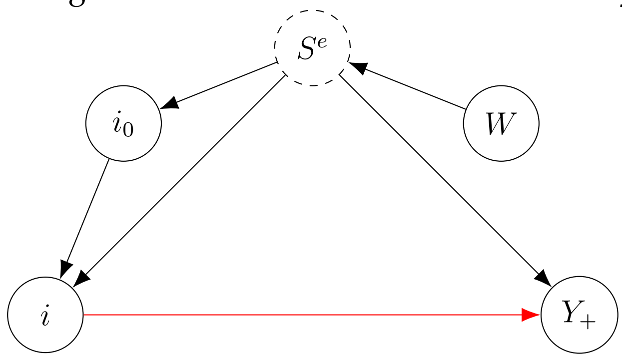



On the graph the policy is enacted in period 1, while news arrive in periods 0 and 2, all of which happens within a lower frequency aggregation period, e.g. days within a month. The red line represents the causal effect of interest, that is the effect of the market rates on the future macroeconomic outcomes. The focus is on period 1 because that is the period in which a policy change happens and the correct selection of the variables allows to exploit the orthogonality structure of the problem. In this case the relevant variables are the market rates in period 1 $i_1$, the fed funds rate $i_{0,1}$, the state expectations $S_1^e$, macro news $(W_0^0, W_1^1)$ and the outcomes $Y_{2+h}$.

## 4.2 Orthogonality restrictions

A simplified graphical representation of the model and the implies restrictions is presented in Figure 4.

Figure 4: Negative control identification of monetary policy




This figure displays a simplified version of the information flow in the monetary policy context. All the variables besides the outcome belong to the same observation period, while the outcome variable $Y_+$ is in the future, hence the subscript. While the representation looks static, it does transmit the information flow structure in a concise way. News affect the state expectations and only through them do they affect the fed funds rate and the market interest rates. The fed funds rate only affects the outcome through the market rates. And the red line represents the causal effect of interest. As such, the graph in Figure 4 only aggregates the underlying structure of the full causal graph described in the model (equations 16 - 29) and represented in Figure 3.

Armed with the simplifying graph, we can review the orthogonality conditions from a top down perspective. Let us now reconsider the set of restrictions that the model imposes. In addition to the state space model presented earlier, several additional justifying arguments are to be made.

**Macroeconomic news shocks $W$ and interest rates**

In the proposed framework there is a clear fit for the chosen set of what will constitute the negative control outcome variables: macroeconomic news shocks. First, let us first clarify what the empirical counterparts to these proposed variables are. In short, $W$ are the variables that represent the surprise component of the market expectations on a set of macroeconomic variable news releases. The data we use come from Bloomberg and consist of the median of the market views on a particular macroeconomic news release value as well as the release value itself. Taking the difference between the two we obtain a measure of the surprise in the financial market. As with all the analysis involving the market reactions, it is key to look at the surprises rather than the news directly since, arguably, only the news component should constitute an input for new repricing actions. The efficient market hypothesis would guarantee this type of behavior but it is not strictly needed to make the broad reasoning go through.

The key restriction that has to hold for the negative control outcome $W$ is that it should be independent of $i_0$ and $i$ conditional on the unobservables.

$$(i_0, i) \perp\!\!\!\perp W | S^e$$

In the proposed setting this corresponds to the independence of the macroeconomic news shocks from the short and medium-to-long points of the yield curve conditional on the expectations of the state of the economy. Before proceeding with the analysis of the plausibility of this assumption it is paramount to better describe unobserved confounder. Since both the negative control treatment (fed funds rate) and the treatment (market yields), the correct unobserved confounder in this case seems to be the consensus expectations of the Fed and the markets about the state of the economy. In this work we do not make the distinction between the two. While the existing literature puts some focus on the distinction between the Fed and the market views as well as the derivative Fed "information effect", this is not necessary for the approach to work. The interplay of the markets and the Fed however presents an interesting avenue of extension and is deferred to future work.

Let us recall the causal flow structure of the problem. First, there is an arrival of macroeconomic news. Then, the unexpected component of the news is the relevant part of information used to update the expectations of the state of the economy (think of the Kalman filter type updating in a latent state model that had the macro news providing an input for the updating equations). Next, the updated expectations are acted acted upon: market participants rebalance their portfolios and reach a new set of equilibrium prices and the related interest rates. The past steps are quite condensed in time, but the further ones are much less so. Next, the real economy reacts to the new interest rates by adjusting investment and consumption decisions and there is an eventual reaction of the usual battery of the macroeconomic variables of interest: inflation, unemployment and output.

Given the temporal structure above, the reasoning in terms of the independence assumption proper seems to be quite straightforward. Since the should not be any other channel that mediates the effect of the market news arrival on the interest rates (the only effect is mediated by the updated market expectations of the state of the economy) the conditional independence assumption should be easily satisfied.

So far, no restrictions on the dimensionality of $S^e$ were made. However, the implementation of the negative control approach will need to place restrictions on the size of $S^e$. This is important when thinking about nonlinearities. Suppose that the Fed and the markets react differently to macroeconomics news shocks differently depending on the sign of the shock. Say, negative unemployment news produce larger movements in the interest rates given the expectations about the state of the economy. This would invalidate the exogeneity assumption because the second moments of the yield curve would be correlated with the macro news shocks. One might then argue that it would be possible to suitably expand the dimension of $S^e$ to incorporate the potential non-linear relations. While this is possible in theory, this would require an unfeasibly rich set of negative controls, as discussed in the future sections, and thus is not attainable in practice. Therefore, in general, this work will focus on a logic of linear relations in hope of elucidating the relevant identification mechanisms and using the feasible set of data.

**Fed funds, yield curve and the economy**

The second independence assumption that the graphical model implies is that the fed funds rate $i_0$ has no effect on the economy conditional on the market-controlled longer end of the yield curve $i$.

$$i_0 \perp\!\!\!\perp Y_+ | S^e$$

The reasoning about the negative control treatment is a bit more complicated. First, one has to make a choice of what constitutes the treatment and what constitutes the NC treatment. In the baseline example, the treatment variable is a longer maturity point on yield curve and the NC treatment is the very short end of the yield curve. This choice is pinned by two considerations. One of the reasons is that two separate variables were needed to satisfy the conditional independence wrt $W$. The more important reason is that the proposed negative control treatment would need to be independent of the outcome variable given the treatment and the unobserved confounder. The last piece of reasoning is an important link in this story that merits further comment.

Why can we argue that the NC treatment exclusion restriction is satisfied? As a preamble, this assumption does not fit the usual narrative in the literature. The typical analysis and reasoning would link the short end of the yield curve directly to the outcomes. The prototypical VAR links the effects of the changes to the fed funds rate to the key macro outcomes (inflation and output) and there are no arguments as to whether the longer run yields act as mediators (or potentially sole mediators) of this effect. Moreover, recent research by Nakamura and Steinsson (2018) and Bauer and Swanson (2023*a*) employs 1 and 2-year treasuries as the monetary policy instrument directly, without focusing much on the distinction between what is actually directly controlled by the Fed and what is not.

The separation between the fed funds rate $i_0$ and a market rate $i$, e.g. 2Y treasury yield, can be justified by the fact that most of the decisions that involve financing require debt that directly involved longer maturities. Typical consumption decisions that involve financing include car loans and mortgages and both of these involve financing horizons that span from one to thirty years. On the investment side, most business decisions have a horizon of years. And even the liquidity management aspect of the businesses that can involve much shorter rates as the reference points rarely relies on the overnight rates, which are essentially under the direct control of the Fed. At this point one can ask whether a very short maturity instrument, like a 3-month bill, will fall together with the set of market rates $i$. According to this reasoning, it will fall in there, as the identifying assumptions will still hold. However, it will be harder to produce precise estimates given the fact that such a short rate covaries significantly with the overnight fed funds rate.

The question of the covariation of the interest rates is also an important one. Typically, the literature models the yield curve as a cointegrated vector, where the base dynamics is guided by a process that in the simpler versions is close to a random walk. This leaves a question of whether it is overall possible to regard the interest rates as sources of different information. The key to answering this question is looking into the rate correlations once the unit root dynamics are taken out of the picture. One feasible measure of the codependence between the variables is the partial $R^2$ coefficient in a regression $i_t = \alpha_0 + \alpha_1 i_{0,t} + \alpha_2 i_{0,t-1} + \alpha_3 i_{t-1} + \epsilon_i$. The lagged interest rate values are supposed to imperfectly control for the persistent part of the state expectations, while the coefficient $\alpha_1$ should provide a measure of the direct relation between the two. The value of the partial $R^2$ for $\alpha_1$ in this regression is on the order of 20%. While the model-driven reason for this value of the statistic would require further assumptions about the relation of both variables to the unobservable state expectations, this value is indicative of the fact that, while being related, there is a substantial amount of the variation that influences the market rates $i$ through other channels.

The last comment of the section regards a possible interpretation of the distinction between the central bank and the market rates. Using the microeconometrics terminology, one can interpret the fed funds rate as the treatment assignment and the market rates as the actual treatment value. This perspective is useful for a number of reasons. First, it elucidates the fact that the Fed cannot directly affect the rates relevant for the real economy, or, in other words, there might be imperfect compliance. Second, it clarifies the difference between the estimates using the two rates. In the case the focus is on the fed funds rate, one aims to estimate the ITT, the intention to treat effect. Using the market rates can then be interpreted as gearing towards estimating the average treatment effect (ATE) proper.

Summing up, the assumption of the irrelevance of the fed funds rate conditional on the market rates seems reasonable, while it is also reasonable to expect a significant amount of the covariation needed to drive the estimates. Next, we will proceed to addressing the other relevant points of the proposed model and highlight the key conceptual features aside the orthogonality conditions.

## 4.3 Discussion

The following subsection presents the discussion of the proposed information model as well as the resulting orthogonality restrictions subsequently used by the negative control framework.

**Time aggregation** As highlighted before, the question time aggregation is instrumental for the justification of the orthogonality restrictions. If one uses the typical aggregation schedule of taking the last values of every sub-period, the desired restrictions will not hold. Taking the example from Figure 3, $i_2^0 \not\perp\!\!\!\perp Y_{2+h} | S_2^e, i_2$. To see this, consider the the following subgraph: $i_2^0 \leftarrow i_1^0 \rightarrow i_1 \rightarrow S_2 \rightarrow S_{2+h} \rightarrow Y_{2+h}$. This subgraph shows one route for generating the correlation between the fed funds rate and the outcome. Conditioning on $S^e - 2$ and $i_2$ does not block the proposed subgraph (no d-separation occurs by the conditioning set [5]) which implies that the dependence between the fed funds rate and the future outcome is not eliminated. The same logic holds for the active path $i_2^0 \leftarrow i_1^0 \leftarrow S_1^e \leftarrow S_0^e \rightarrow i_0 \rightarrow S_0 \rightarrow S_{2+h} \rightarrow Y_{2+h}$, which is not

[5] See Pearl (2009) for the exposition on the graphical model terminology and concepts.

blocked by the conditioning variables. The intuition for this fact is that conditioning on the last values of the state expectations and the market interest rates does not perfectly condition on the state expectations and the market rates on the time of the policy action. End of the month values reflect the updates that occurred throughout the month after the policy action and thus have additional noise as compared to the state expectations on the policy day.

The aggregation of macroeconomic news does not present similar complexities. The statement of Proposition 2 invokes a simple end-of-period aggregation of the macroeconomic news $W_\tau$. In that case it is irrelevant whether the macroeconomic news occur before or after the policy action period. On Figure 3, $W_0^1$ arrives before the policy and is only correlated to the fed funds rate and the market interest rate via the state expectations channel: $W_0^1 \to S_0^e \to S_1^e \to (i_1^0, i_1)$. Same goes for the future news $W_2^2$. The path that drives the correlation is $(i_1^0, i_1) \leftarrow S_1^e \to W_2^2$.

**Information span of the NC variables** As a reminder, the identification argument involves several key components. First, there are the two orthogonality conditions that pin down the ability to block the confounding variation. Additionally, there is the assumption of completeness that ensures that the relevant variation is picked up to a sufficient degree. While this assumption is conceptually similar to IV strength, it is different as it refers to the unobserved confounding factors that can reside in a multidimensional space.

The reasoning behind the completeness assumption relies on the fact that the variation in the state expectations vector has to be related to the variation in both the Fed policy measure and the macroeconomic news. In case the state expectation vector is unidimensional, the sufficient condition to justify completeness is to state that the fed funds rate and the macroeconomic news do correlate with the state expectations along its support. In case of multidimensional confounding, one would also require the negative control outcome and treatment to have at least the same dimension as the dimension of the confounding space. This point represents a weakness of the approach as the present application only uses one negative control treatment, the fed funds rate. While the direct response of the fed funds rate represents an important instrument and a direct measure of Fed policy, augmenting the set of variables describing the direct action of the Fed is the primary direction of extension of the present work. Using other measures of Fed policy would allow to capture the variation in the possibly multidimensional vector of the state of economy expectations and improve inference.

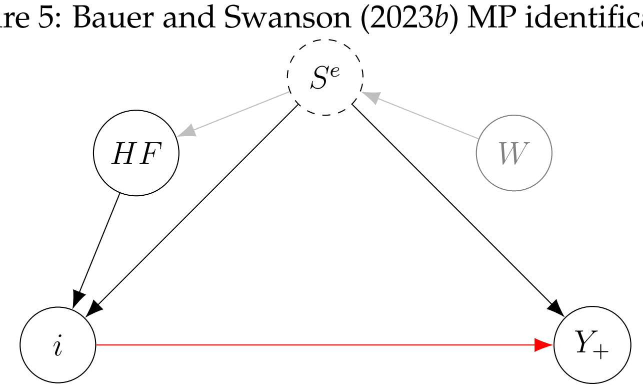
**Distinction between state and state expectations** Another crucial point of distinction from the typical state space models is the idea that the econometrician does not observe the same data as the Fed and the market which does not allow to perfectly recover the state expectations of these agents. This difference in the set of observables drives the difference between the latent state and the agents' state expectations. The latent state is not observed by anyone but the consensus state expectations are only observed by the market participants and the Fed. The distinction between state and state expectations is also important in guiding the reasoning about the structure of the confounding mechanism. If the latent state was the relevant confounder, macroeconomic news would act as a mediator for the Fed policy measures and the market interest rate response.

Instead, once the state expectations are considered as the relevant confounder, the role of the macroeconomic news changes. News then can be viewed as a proxy of the state expectations but including the news as controls would not eliminate the bias as would be the case in the mediator perspective.

**Comparison to high-frequency IV identification** The negative control approach is closely related to the high-frequency IV approach used in Bauer and Swanson (2023*b*). The approach proposed by the authors relies on different identification assumptions and focuses on the exogenous variation of in the interest rates but similarly employs macroecnomic news to fight the possible endogeneity. The authors highlight that the macroeconomic news are predictive of the Fed response and thus the high-frequency variation of the market interest rates around the FOMC announcement release. The proposed solution is then to orthogonalize the high-frequency shocks with respect to the news that predict the shocks and cause endogeneity concerns.

In the reasoning of the authors, orthogonalizing with respect to the information that predicts the instrument renders the instrument exogenous. However, if we cast this reasoning within the framework proposed in this section, the effect of orthogonalization is not exactly the one desired by the authors. The difference lies in the consideration of what constitutes the unobserved confounder. If the latent state was considered as the confounder, the approach of the authors would work as the news would be the only mediator of the confounder that would cause the endogeneity of the instrument. However, if the state expectations are considered as the relevant confounding factors, the macroeconomic news occupy a different role in the causal structure. News inform the state expectations and can be viewed as a proxy of these expectations but are not a mediator of the confounding effect. Thus, as shown in Section 3, controlling for the news will reduce but not eliminate the bias. This reasoning can be demonstrated in a causal graph presented in Figure 5.

Figure 5: Bauer and Swanson (2023*b*) MP identification




The negative control approach can be seen as an improvement over the orthogonalization approach used in Bauer and Swanson (2023*b*). Although it does rely on a different set of assumptions, the approach correctly exploits the orthogonality structure derived from the information flow analysis of the monetary policy problem and allows for correct identification.

As was shown in Section 3, the approach can in general be seen as an improvement on the IV approach as it is also valid in case the instruments are truly exogenous.

The comparison of empirical performance of the negative control approach to the IV approach of Bauer and Swanson (2023*b*) is explored in Section 5. As will be shown, the NC application does improve on the baseline unadjusted IV case although the estimates will be lower in magnitude as compared to the orthogonalized IV approach.

# 5 Empirical application

The following section presents the empirical results from the application of the negative control framework to the estimation of monetary policy effects. The primary goal of the empirical application will be the comparison of the resulting estimates to the best practice estimates provided in Bauer and Swanson (2023*b*), B&S in short. The authors address the identification question by elaborating on the IV identification idea that has been prominent in the recent literature.

Throughout this work we have exposed the negative control approach as an alternative to other two commonly known approaches: conditional exogeneity and IV. The former has not received extensive developments in the recent literature. The latter, however, has been thoroughly examined in the recent works of Gertler and Karadi (2015), Ramey (2016), Plagborg-Møller and Wolf (2021), and Miranda-Agrippino and Ricco (2021). B&S summarize the literature employing the IV approach and provide best practice estimates, which makes for an illustrative comparison with the negative control approach. As we present specification and estimation details, we will stick as close as possible to the B&S baseline specification to make sure that the difference in the presented estimates is only driven by the difference in the identification approach. In practical terms, the proposed NC implementation will only differ from the B&S baseline version in the short-run identification step, while the further propagation of the shocks will be common, which will allow for most direct comparison.

Of particular interest is the explicit modeling of the confounding mechanism that distinguishes the present work from Bauer and Swanson (2023*b*). within the proposed framework, macroeconomic news occupy a specific role in the causal chain. It is a proxy but not a mediator of the confounder. B&S employ the news measures to address deficiencies of the high-frequency instrument by orthogonalizing the instrument with respect to the news. The negative control reasoning presented in Section 4 argues that using macroeconomic news as controls (or controls in IV setting) will reduce not eliminate the omitted variable bias once state expectations rather than the state itself is taken as the unobserved confounder. Using news as negative control outcome variable has the potential of eliminating the confounding bias and the empirical application serves to analyse this theoretic implication.

## 5.1 Data

This work employs the data sets from several sources: data employed in Bauer and Swanson (2023*b*) originally coming from FRED, treasury yield curve data from the US Department of the Treasury, and data on macroeconomic news from Bloomberg.

**Macroeconomic data** The data on macroeconomic outcomes of interest are borrowed from Bauer and Swanson (2023*b*). We will focus on the baseline specification proposed by the authors and will employ the same set of six macroeconomic series.

- IP – Industrial Production

- u – Unemployment
- Pcomm – Commodity Price Index in levels
- CPI – Consumer Price Index in levels
- EBP – Excess Bond Premium
- 2y Treasury – 2Y treasury bond yield
- ff – Effective federal funds rate

**Yield curve** The yield curve data come from the US Department of the Treasury. The US Treasury One-Year Forward Rate Curve comes from NASDAQ Data.

**Macroeconomic news** The data on macroeconomic news come from Bloomberg. In particular, I construct the news measure as the actual data release value minus the mean market expectation value prior to the release. The following news series are used: Consumer Confidence, Core CPI, Employment Cost, Federal Funds Rate, GDP Third Release, Initial Claims, NAPM Manufacturing Prices, New Home Sales, Non-Farm Payroll, Retail Sales, Unemployment Rate, and Producer Price Index.

## 5.2 Empirical specification

The empirical implementation will borrow several ideas from Bauer and Swanson (2023*b*). In its heart, negative control approach is most conducive to applying the method of local projections (LP) given the nature of the restrictions that relate the contemporaneous values of the market rates, Fed policy and macroeconomic news to the outcomes in the next periods. However, LP estimates exhibit large amounts of volatility with potentially small gains in bias and, following suggestions from Li, Plagborg-Møller and Wolf (2022), we will follow the best practice approach proposed in the reference work. We will first obtain one step ahead causal estimates and then use the VAR-implied dynamics to propagate these estimates along the response horizon. Prior to showing the results, we will introduce the replication specification and then talk about the implementation in the negative control case, highlighting the differences between the two.

### 5.2.1 Bauer and Swanson (2023b) replication

The replication specification directly follows the paper of reference. It uses an IV identification strategy employing high-frequency changes in the yields following Fed policy actions as an instrument for the 2Y yield. This exogenous variation in the 2Y yield is takes as the exogenous monetary policy variation and is used to estimate the causal effects of monetary policy.

Two particular estimates from B&S will be the focus of comparison. First is the estimate using the existing high-frequency instrument directly with no additional adjustments. The instrument series $z_t$ is constructed by using the first principal component of four series of Eurodollar quarterly futures contracts using current quarter and the three following quarters. The estimates using $z_t$ as the instrument will be denoted as *IV*. The other series or reference will be an improved version of the former. B&S highlight the failure of exogeneity of $z_t$ given its predictability using macroeconomic news. To address this deficiency, they construct a new series $z_t^{\perp} = z_t - \mathbb{E}[z_t|news_t]$ that captures the variation orthogonal to the macroeconomic news over the same period. The estimates using $z_t^{\perp}$ as the instrument will be denoted as *IV+*.

The data used for estimation are comprised of macroeconomic series or interest and the series of high-frequency monetary policy shocks. Six macroeconomic data series are gathered in the vector $Y_t$: Industrial Production, unemployment, Commodity Price Index in levels, Consumer Price Index in levels, Excess Bond Premium, and 2Y treasury bond yield. In addition to macroeconomic data, we use the two high-frequency instrument series provided by B&S: standard high-frequency shock series and the orthogonalized version of the same series. The only difference between our implementation and the baseline B&S version is the fact that we don't use the Fed chair announcements as the additional source of shocks identifying monetary policy as these data are not yet readily available.

The estimation of the causal effects follows Stock and Watson (2012) and Stock and Watson (2018). The following regression

$$
\begin{aligned}
Y_t \quad &= \beta_0 + A(L)Y_{t-1} + \beta_i i_{2Y,t} + u_t && (30)\\
i_{2Y,t} &= \delta_0 + \delta_z z_t^* + v_t && (31)
\end{aligned}
$$

is estimated using two-stage least squares using either $z_t$ or $z_t^{\perp}$ as instruments $z_t^*$ for the 2Y treasury yield $i_{2Y,t}$. In accordance with B&S, we estimate $A(L)$ on a larger sample and then use the residuals for the estimation of $\beta_i$.

The regression above identifies the contemporaneous correlations between the 2Y yield structural shock and the other variables in the system up to scale with $\beta_i$. The scale is then normalized to have a unit monetary policy shock correspond to a 25 basis point change in the 2Y yield on impact. Once the contemporaneous effects are identified, the causal effects are propagated forward using the VAR dynamics, $A(L)$:

$$
\Psi_h^i = \frac{\partial Y_{t+h}}{u_t}\beta_i \tag{32}
$$

where $\Psi_h$ is the coefficient matrix of the structural VMA process corresponding to horizon $h$ and $\Psi_h^i$ is the vector of the dynamic causal effects of monetary policy (exogenous 2Y yield changes) at horizon $h$. The figures presented further in the section will present both the standard high-frequency instrument version (*IV*) and the version using orthogonalized instruments (*IV+*).

### 5.2.2 Negative control specification

The baseline negative control specification is the following:

$$Y_{t+1} = \beta_0 + \beta_i i_{2Y,t} + \beta_w W_t + A(L)Y_{t-1} + u_t \tag{33}$$

$$W_t = \alpha_0 + \alpha_i^0 i_{0,t} + \alpha_i^{2Y} i_{2Y,t} + v_t \tag{34}$$

In this specification, $\beta_i$ is the coefficient of interest and is the causal effect of a change in the market interest rate of a medium maturity (baseline version focuses on 2Y yield). The estimation approach is equivalent to using the IV approach where macroeconomic news series $W_t$ is instrumented using the fed funds rate $i_{0,t}$ and the 2Y yield $i_{2Y,t}$ is an exogenous regressor.Implementation is done via GMM, following the procedure described in Section 3.

The estimation proceeds in several steps. First, we estimate a VAR in the outcome variables $Y_{t+1} = \alpha + A(L)Y_t + u_t$. Following B&S, we use 12 lags, though results are not much sensitive to lag choice overall. Exploiting data availability, we employ a larger dataset, spanning years 1973 to 2020, to estimate this VAR. Second, we use the VAR residuals to estimate equations 33-34 on a shorter sample spanning years 1999 to 2020. Thus we obtain the one step ahead causal effects $\beta_i$.Next, using the estimated coefficient values, we calculate the structural impulse response values using the formula 32. We normalize these impulse responses so as to have the horizon zero causal effect on the 2Y yield be equal to 25 basis points. The standard errors are estimated using bootstrap.

## 5.3 Results

The main specification focuses on the causal effects of changes in the 2Y treasury yields. The estimation results are presented in Figure 6. The figure plots the impulse responses of three implementations. *IV* is the classical high-frequency IV implementation from B&S. *IV+* is the B&S IV implementation using orthogonalized instruments. *NC* is the negative control implementation. The shaded region represents a 90% bootstrapped confidence interval around the NC estimates.

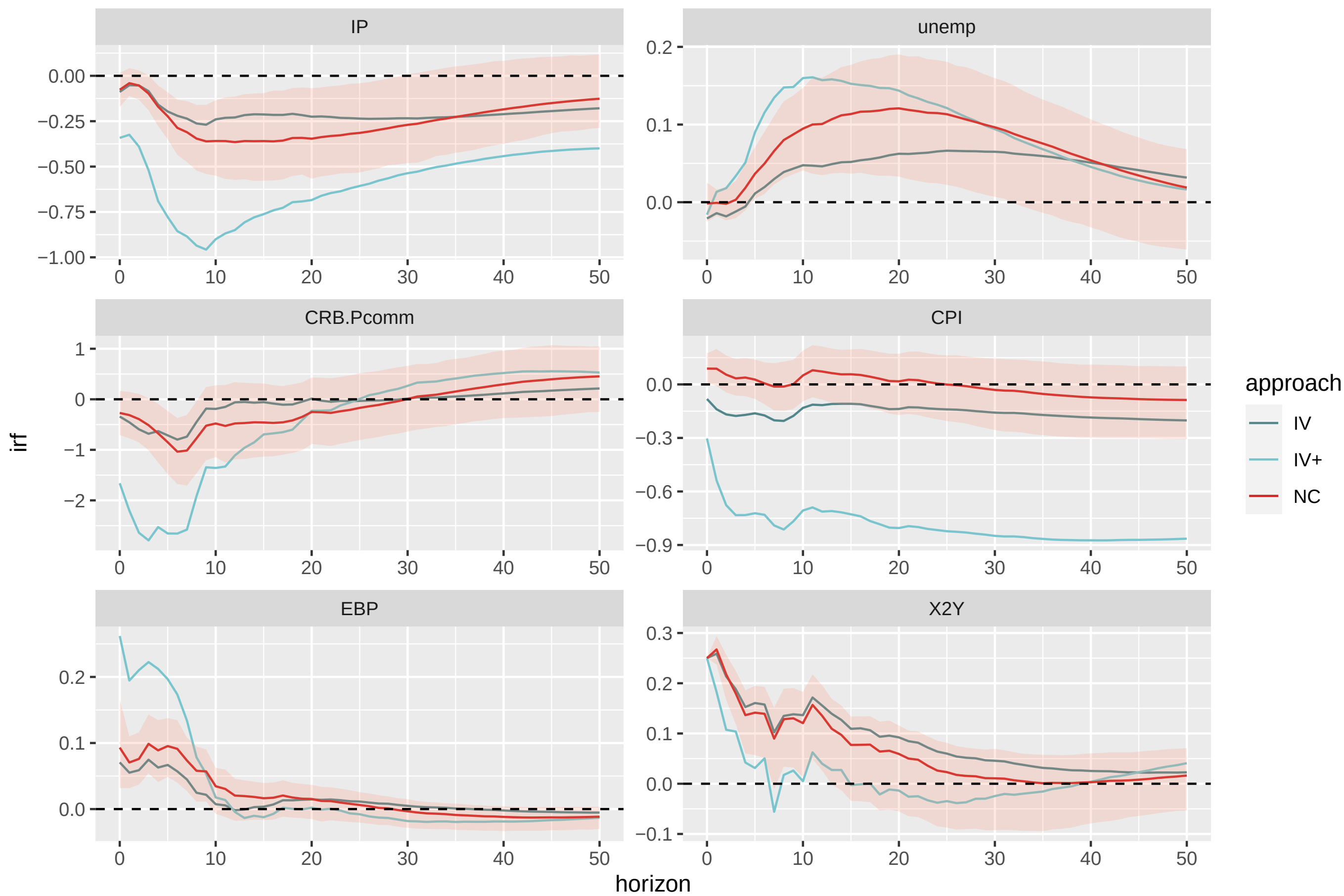



Figure 6: Effects of 2Y yield changes

Comparison of the negative control approach to IV approaches of Bauer and Swanson (2023*b*). Impulse responses to a 25 basis point 2Y yield shock. The shaded region represents a 90% bootstrapped confidence interval. IP = Industrial Production, unemp = Unemployment, CRB.Pcomm = Commodity Price Index in levels, CPI = Consumer Price Index in levels, EBP = Excess Bond Premium, X2Y = 2Y treasury bond yield.

We start by highlighting a large difference in B&S *IV* and *IV+* estimates. Orthogonalizing the high-frequency instrument with respect to predictive news results in estimates that are far more pronounced in several directions. *IV* estimates an negative response of IP that reaches its lowest point around 10 months out at -0.25%. The orthogonalized version *IV+* produces estimates that are four times larger, reaching the lowest point at -1%. The unemployment response is quite different as well. The standard approach estimates a persistent increase in unemployment of around 0.6% while the orthogonalized approach exhibits a far more humped-shaped response peaking at 1.6% around 10 months after the shock. Commodity price index exhibits almost no reaction to a 2Y treasury shock as estimated by *IV*, while *IV+* estimates a drop of around 2% over the first year converging to zero by the end of a two-year horizon. Consumer price index estimates under *IV* show a persistent drop of nearly 0.2%. *IV+* results are far more pronounced for CPI, with most of the reaction happening over the first year after the shock, to -0.75%. Excess bond premium response is far larger with *IV*+ as compared to *IV*, averaging 0.2% against 0.6 % along the first six months. The dynamics of the 2Y yield itself is also quite different, with the key distinction being the difference in persistence. The non-orthogonalized version exhibits very slow dropoff while the orthogonalized approach predicts a fast dropoff towards zero by the end of first six months.

Now let's focus on the negative control approach. The overarching point is that the *NC* estimates tend to lay in between the *IV* and *IV*+ estimates, though some responses are different in other respects. Industrial production estimates lay close to *IV*, reaching the lowest point of -0.38 towards the end of the first year. Unemployment response is exactly midway between the two high-frequency instrumental variable approaches, peaking at -0.12% around 20 months in. The response of price indices is closer to *IV*, with a slight price puzzle for CPI response. EBP response is also close to *IV*, averaging 0.1% in the first six months. The dynamics of the 2Y yield exhibit the same level of persistence as *IV*, diminishing towards zero only near the end three years after the shock.

Comparing the *NC* estimates to *IV* and *IV*+ reveals a great deal of similarity in the resulting estimates. While the approaches are conceptually different and employ different sets of variables to identify the immediate causal impacts, the dynamics are quite close. To a degree, this is driven by the fact that the same VAR dynamics are used to propagate the impulse response functions beyond the first horizon step. But for the most part these dynamics are driven by the causal effects identified at either the contemporaneous horizon in the case of *IV* and *IV*+ or the one step ahead horizon in the case of *NC*. Under the assumptions of Proposition 1, the NC approach should provide for unbiased causal inference. Approaches that involve controlling for macroeconomic news $W_t$ should improve in terms of absolute bias levels but not eliminate the bias.

Adding macroeconomic news reduces the bias but does not eliminate it and the amount of the reduction depends on two factors, namely how predictive the news are of the state expectation changes and how much idiosyncratic variance there is in state expectations. Overall we see that the responses provided by the NC approach go in the direction of bias reduction as well although they don't surpass the *IV*+ estimates as the theoretical analysis predicts. There might be several reasons for this: other news series, sampling uncertainty, using other variables (fed funds rate) as the instruments and overall a different perspective. Considering the limitations it is particularly interesting that the NC approach seems to have produces less bias than the standard *IV* approach, although it followed a completely different procedure. Further elaboration on the method and expanding on the set of measures of Fed policy could shed more light on the problem with the expectation that the current estimates are still downward biased. On the other hand, the present work only focuses on the simpler case of negative control framework, where only the fed funds rate is used as the negative control treatment. This fact might hinder inference when the unobserved confounder is multidimensional. With the potential improvements in mind, the fact that the approach delivers estimates that possess such degree of similarity to the baseline versions of B&S is reassuring for the negative control approach.

**Effects of longer maturity yields** Next we present the estimates of the impulse responses where the market interest rate of interest is of longer maturity. We focus here on the estimates of the causal effects of a change in a 10Y treasury yield, as it represents a rate of longer maturity that might be even more prominent in terms of the macroeconomic implications. Causal effect estimates for rates of other maturities are presented in the Appendix C.

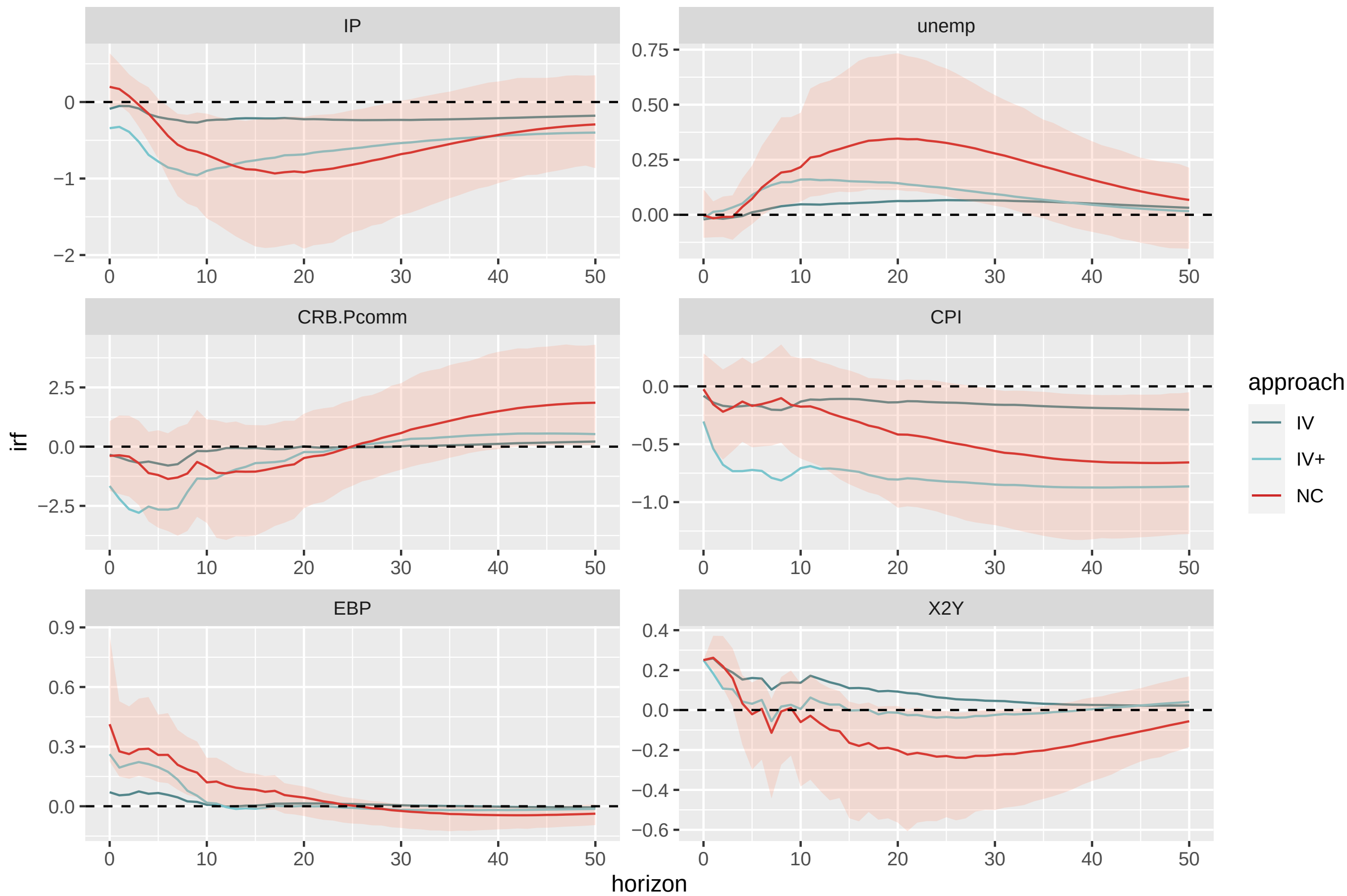



Figure 7: Effects of 10Y yield changes

Comparison of the negative control approach to IV approaches of Bauer and Swanson (2023*b*). Impulse responses to a 10Y yield shock scaled to have 25 basis point impact on 2Y yield. The shaded region represents a 90% bootstrapped confidence interval. IP = Industrial Production, unemp = Unemployment, CRB.Pcomm = Commodity Price Index in levels, CPI = Consumer Price Index in levels, EBP = Excess Bond Premium, X2Y = 2Y treasury bond yield.

The causal effect estimates of the effects of 10Y yield changes are far more pronounces than the effects of 2Y yield changes. IP and unemployment estimates become comparable to the *IV+* estimates, with unemployment response surpassing that of *IV+* and peaking at 0.3% towards a year and a half after the shock. The decrease in the commodity price index is similar to *IV+* although less pronounced. CPI effects on the other hand are low in magnitude on impact but continuously grow over time surpassing -0.5% at a three year mark. Excess bond premium response surpasses that of *IV+* and is slightly more persistent while the response of the 2Y yield is as short lived as the *IV+* response with the rate response becoming negative after the first year as a consequence of the large real economy response. The fact that the estimated causal effects are more pronounced for the 10Y yield changes tells us that the real economy might be more responsive to the changes in the longer rates as the financing structure of the real economy might impact the transmission of monetary policy. A deeper look in this direction is warranted and the differential analysis of the impact of different maturity rate changes is of separate interest.

# 6 Conclusion

The present work develops a negative control framework for identifying the causal effects of monetary policy. The proposed framework relies on exploiting orthogonality restrictions derived from the analysis of the information flow between the latent state, the central bank and the markets. Using the endogenous variation of the fed funds rate, the approach creates a proxy of the confounding factor by extracting the endogenous component of macroeconomic news by using the fed funds rate as an instrument for this variation. The approach presents and alternative to conditional exogeneity and IV approaches by exploiting the endogeneity of the variables in the system.

The view adopted in this paper is useful for several reasons. First, we show how the bottom down analysis of the causal flow can inform the question of causal inference in the context of monetary policy. Second, the perspective explored in this paper illustrates how the endogenous variation can be exploited in identifying causal effects.

The empirical application finds that the negative control approach produces estimates that are close in magnitude to the literature benchmark high-frequency instrumental variable implementation. The NC estimates improve over the typical IV estimates in terms of bias but are lower in magnitude than the estimates using orthogonalized high-frequency instruments.

Further work is needed in the direction of improvement of the approach. One is related to the expansion of the negative control variable set. Namely, additional direct measures of Fed policy would amplify the ability of the approach to handle multidimensional confounding. Additionally, expanding the set of macroeconomic news and other types of signals that provide the markets and the Fed with the information about the state of the economy would also improve precision and potentially allow to explore non-linear application. Exploring the interplay between the markets and the Fed by distinguishing the information sets of the two is another avenue of future work.

## References

**Ang, Andrew, and Monika Piazzesi.** 2003. "A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables." *Journal of Monetary economics*, 50(4): 745–787.

**Bańbura, Marta, Domenico Giannone, and Lucrezia Reichlin.** 2010. "Large Bayesian vector auto regressions." *Journal of applied Econometrics*, 25(1): 71–92.

**Bauer, Michael D, and Eric T Swanson.** 2023*a*. "An alternative explanation for the "fed information effect"." *American Economic Review*, 113(3): 664–700.

**Bauer, Michael D, and Eric T Swanson.** 2023*b*. "A reassessment of monetary policy surprises and high-frequency identification." *NBER Macroeconomics Annual*, 37(1): 87–155.

**Bauer, Michael D, and Glenn D Rudebusch.** 2020. "Interest rates under falling stars." *American Economic Review*, 110(5): 1316–1354.

**Bernanke, Ben S, and Jean Boivin.** 2003. "Monetary policy in a data-rich environment." *Journal of Monetary Economics*, 50(3): 525–546.

**Bernanke, Ben S, and Kenneth N Kuttner.** 2005. "What explains the stock market's reaction to Federal Reserve policy?" *The Journal of finance*, 60(3): 1221–1257.

**Bernanke, Ben S, Jean Boivin, and Piotr Eliasz.** 2005. "Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach." *The Quarterly journal of economics*, 120(1): 387–422.

**Bok, Brandyn, Daniele Caratelli, Domenico Giannone, Argia M Sbordone, and Andrea Tambalotti.** 2018. "Macroeconomic nowcasting and forecasting with big data." *Annual Review of Economics*, 10: 615–643.

**Caballero, Ricardo J, and Alp Simsek.** 2022. "Monetary policy with opinionated markets." *American Economic Review*, 112(7): 2353–2392.

**Deaner, Ben.** 2021*a*. "Many proxy controls." *arXiv preprint arXiv:2110.03973*.

**Deaner, Ben.** 2021*b*. "Proxy Controls and Panel Data." *arXiv preprint arXiv:1810.00283*.

**Diebold, Francis X., Glenn D. Rudebusch, and S. Borag˘an Aruoba.** 2006. "The macroeconomy and the yield curve: a dynamic latent factor approach." *Journal of econometrics*, 131(1): 309–338.

**Forni, Mario, Domenico Giannone, Marco Lippi, and Lucrezia Reichlin.** 2009. "Opening the black box: Structural factor models with large cross sections." *Econometric Theory*, 25(5): 1319–1347.

**Gertler, Mark, and Peter Karadi.** 2015. "Monetary policy surprises, credit costs, and economic activity." *American Economic Journal: Macroeconomics*, 7(1): 44–76.

**Gürkaynak, Refet S, and Jonathan H Wright.** 2012. "Macroeconomics and the term structure." *Journal of Economic Literature*, 50(2): 331–367.

**Gürkaynak, Refet S, Brian Sack, and Eric Swanson.** 2005. "The sensitivity of long-term interest rates to economic news: Evidence and implications for macroeconomic models." *American economic review*, 95(1): 425–436.

**Gürkaynak, Refet S, Brian Sack, and Eric T Swansonc.** 2005. "Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements." *International Journal of Central Banking*.

**Hanson, Samuel G, and Jeremy C Stein.** 2015. "Monetary policy and long-term real rates." *Journal of Financial Economics*, 115(3): 429–448.

**Hillenbrand, Sebastian.** 2022. "The secular decline in long-term yields around fomc meetings." hbs working paper.

**Jarociński, Marek, and Peter Karadi.** 2020. "Deconstructing monetary policy surprises—the role of information shocks." *American Economic Journal: Macroeconomics*, 12(2): 1–43.

**Karnaukh, Nina, and Petra Vokata.** 2022. "Growth forecasts and news about monetary policy." *Journal of Financial Economics*, 146(1): 55–70.

**Kuttner, Kenneth.** 2001. "Monetary Policy Surprises and Interest Rates: Evidence from the Fed Funds Futures Market." *Journal of Monetary Economics*, 47(3): 523–544.

**Lakdawala, Aeimit, and Matthew Schaffer.** 2019. "Federal reserve private information and the stock market." *Journal of Banking & Finance*, 106: 34–49.

**Li, Dake, Mikkel Plagborg-Møller, and Christian K Wolf.** 2022. "Local projections vs. vars: Lessons from thousands of dgps." National Bureau of Economic Research.

**Lipsitch, Marc, Eric Tchetgen Tchetgen, and Ted Cohen.** 2010. "Negative controls: a tool for detecting confounding and bias in observational studies." *Epidemiology (Cambridge, Mass.)*, 21(3): 383.

**Lousdal, Mette Lise, Timothy L Lash, W Dana Flanders, M Alan Brookhart, Ivar Sønbø Kristiansen, Mette Kalager, and Henrik Støvring.** 2020. "Negative controls to detect uncontrolled confounding in observational studies of mammographic screening comparing participants and non-participants." *International Journal of Epidemiology*, 49(3): 1032–1042.

**Miao, Wang, Xu Shi, and Eric Tchetgen Tchetgen.** 2018. "A confounding bridge approach for double negative control inference on causal effects." *arXiv preprint arXiv:1808.04945*.

**Miao, Wang, Zhi Geng, and Eric J Tchetgen Tchetgen.** 2018. "Identifying causal effects with proxy variables of an unmeasured confounder." *Biometrika*, 105(4): 987–993.

**Miranda-Agrippino, Silvia, and Giovanni Ricco.** 2021. "The transmission of monetary policy shocks." *American Economic Journal: Macroeconomics*, 13(3): 74–107.

**Nakamura, Emi, and Jon Steinsson.** 2018. "High-Frequency Identificaiton of Monetary Non-Neutrality: The Information Effect." *Quarterly Journal of Economics*, 133(3): 1283–1330.

**Pearl, Judea.** 2009. *Causality.* Cambridge university press.

**Piazzesi, Monika.** 2005. "Bond yields and the Federal Reserve." *Journal of Political Economy*, 113(2): 311–344.

**Plagborg-Møller, Mikkel, and Christian K Wolf.** 2021. "Local projections and VARs estimate the same impulse responses." *Econometrica*, 89(2): 955–980.

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

**Romer, Christina D, and David H Romer.** 2000. "Federal Reserve information and the behavior of interest rates." *American economic review*, 90(3): 429–457.

**Schennach, Susanne M.** 2012. "Measurement error in nonlinear models: A review." CEMMAP working paper.

**Schmeling, Maik, Andreas Schrimpf, and Sigurd AM Steffensen.** 2022. "Monetary policy expectation errors." *Journal of Financial Economics*, 146(3): 841–858.

**Sims, Christopher A.** 1980. "Macroeconomics and reality." *Econometrica: journal of the Econometric Society*, 1–48.

**Singh, Rahul.** 2023. "Kernel methods for unobserved confounding: Negative controls, proxies, and instruments." *arXiv preprint arXiv:2012.10315*.

**Stefanski, LA, and JS Buzas.** 1995. "Instrumental variable estimation in binary regression measurement error models." *Journal of the American Statistical Association*, 90(430): 541–550.

**Stock, James H, and Mark W Watson.** 2012. "Disentangling the channels of the 2007-09 recession." *Brookings Papers on Economic Activity*, 81–135.

**Stock, James H, and Mark W Watson.** 2018. "Identification and estimation of dynamic causal effects in macroeconomics using external instruments." *The Economic Journal*, 128(610): 917–948.

**Tchetgen, Eric J Tchetgen, Andrew Ying, Yifan Cui, Xu Shi, and Wang Miao.** 2020. "An introduction to proximal causal learning." *arXiv preprint arXiv:2009.10982*.

# A Completeness usage

Recall the completeness conditions:

$$rank(\delta_z) \geq dim(S_t) \tag{8}$$
$$rank(\gamma_z) = dim(Z_t) \tag{9}$$

where the coefficients come from the following regressions

$$\mathbb{E}[S_t|D_t, Z_t] = \delta_0 + \delta_d D_t + \delta_z Z_t$$
$$\mathbb{E}[W_t|D_t, Z_t] = \gamma_0 + \gamma_d D_t + \gamma_z Z_t$$

To look deeper intro the workings of the approach and understand where the completeness conditions come into play it is useful to consider employing a conveniently chose set of projections and elaborate on those. Below is the deeper look into the identification question which highlights where the completeness conditions are used. Let $P^D_{Y|Z,D}$ be the projection coefficient on $D$ in a projection of $Y$ on $(Z, D)$. The, using the orthogonality restrictions, we can express the projections in the following way:

$$P_{Y|Z,D} = P_{Y|D,S} \begin{bmatrix} 0 & I_D \\ P^Z_{S|Z,D} & P^D_{S|Z,D} \end{bmatrix} \tag{35}$$
$$P_{W|Z,D} = P_{W|S} P_{S|Z,D} \tag{36}$$

which (since that holds for any values of $Z, D$) implies the equality of all the coefficients

$$\begin{aligned}
P^Z_{Y|Z,D} &= P^S_{Y|D,S} P^Z_{S|Z,D} \\
P^D_{Y|Z,D} &= P^D_{Y|D,S} + P^S_{Y|D,S} P^D_{S|Z,D} \\
P^Z_{W|Z,D} &= P_{W|S} P^Z_{S|Z,D} \\
P^D_{W|Z,D} &= P_{W|S} P^D_{S|Z,D}
\end{aligned}$$

First, let's use the first two equations by substituting in $P^S_{Y|D,S}$. This uses the the completeness condition (8) which is equivalent to the assumption that $\det\left(P^Z_{S|Z,D} {P^Z_{S|Z,D}}^\top\right) > 0$ .

$$\begin{aligned}
P^S_{Y|D,S} &= P^Z_{Y|Z,D} {P^Z_{S|Z,D}}^\top \left(P^Z_{S|Z,D} {P^Z_{S|Z,D}}^\top\right)^{-1} \\
P^D_{Y|Z,D} &= P^D_{Y|D,S} + P^S_{Y|D,S} P^D_{S|Z,D} \\
&= P^D_{Y|D,S} + P^Z_{Y|Z,D} {P^Z_{S|Z,D}}^\top \left(P^Z_{S|Z,D} {P^Z_{S|Z,D}}^\top\right)^{-1} P^D_{S|Z,D}
\end{aligned}$$

Next, we can use the second set of equations to obtain the final expression linking the causal effect of interest to the observables:

$$
\begin{aligned}
P_{W|S} =& P^{Z}_{W|Z,D} {P^{Z}_{S|Z,D}}^{\top} \left( P^{Z}_{S|Z,D} {P^{Z}_{S|Z,D}}^{\top} \right)^{-1} \\
P^{D}_{W|Z,D} =& P_{W|S} P^{D}_{S|Z,D} \\
=& P^{Z}_{W|Z,D} {P^{Z}_{S|Z,D}}^{\top} \left( P^{Z}_{S|Z,D} {P^{Z}_{S|Z,D}}^{\top} \right)^{-1} P^{D}_{S|Z,D} \\
\left( {P^{Z}_{W|Z,D}}^{\top} P^{Z}_{W|Z,D} \right)^{-1} {P^{Z}_{W|Z,D}}^{\top} P^{D}_{W|Z,D} =& {P^{Z}_{S|Z,D}}^{\top} \left( P^{Z}_{S|Z,D} {P^{Z}_{S|Z,D}}^{\top} \right)^{-1} P^{D}_{S|Z,D}
\end{aligned}
$$

where we used completeness assumption (9) which is equivalent to $\det\left( {P^{Z}_{W|Z,D}}^{\top} P^{Z}_{W|Z,D} \right) > 0$. Finally we get the expression for bias of the regression of $Y$ on $(Z, D)$.

$$
P^{D}_{Y|Z,D} = P^{D}_{Y|D,S} + P^{Z}_{Y|Z,D} \left( {P^{Z}_{W|Z,D}}^{\top} P^{Z}_{W|Z,D} \right)^{-1} {P^{Z}_{W|Z,D}}^{\top} P^{D}_{W|Z,D}
$$

The inversions performed in the steps above require restrictions on the matrix ranks and the completeness condition in the linear case does exactly that.

# B Using additional NC outcomes

To recall, the key moment condition that identifies $\beta_d$ is the following:

$$
\mathbb{E}[(Y_{t+h} - \beta_d D_t - \beta^{*}_{s} W_t) \, [D'_t, Z'_t]] = 0 \tag{15}
$$

When the number of negative control outcomes $W_t$ surpasses the number of NC treatments $Z_t$, the direct GMM estimation using (15) is not possible as the coefficients $\beta^{*}_{s}$ are not uniquely identified. Let $W'_t = [W^1_t ... W^j_t ... W^J_t]$ and let $dim(Z_t) = 1$. One way to go around it is to realise that the orthogonality conditions hold for each $W^j_t$ individually. Then we can stack the moments for each $W^j_t$ to obtain overidentification and employ the information in the entire vector $W_t$ to improve the precision of $\beta_d$. The stacked moments have the following form:

$$
\mathbb{E} \begin{bmatrix} (Y_{t+h} - \beta_d D_t - \gamma_1 W^1_t) \, [D_t, Z_t] \\ ... \\ (Y_{t+h} - \beta_d D_t - \gamma_J W^J_t) \, [D_t, Z_t] \end{bmatrix} = 0
$$

As the system is overidentified ($J + 1$ parameters and $2 * J$ equations), GMM can proceed by choosing the optimal weighting matrix the usual way.

## C Causal effects of yields of different maturities

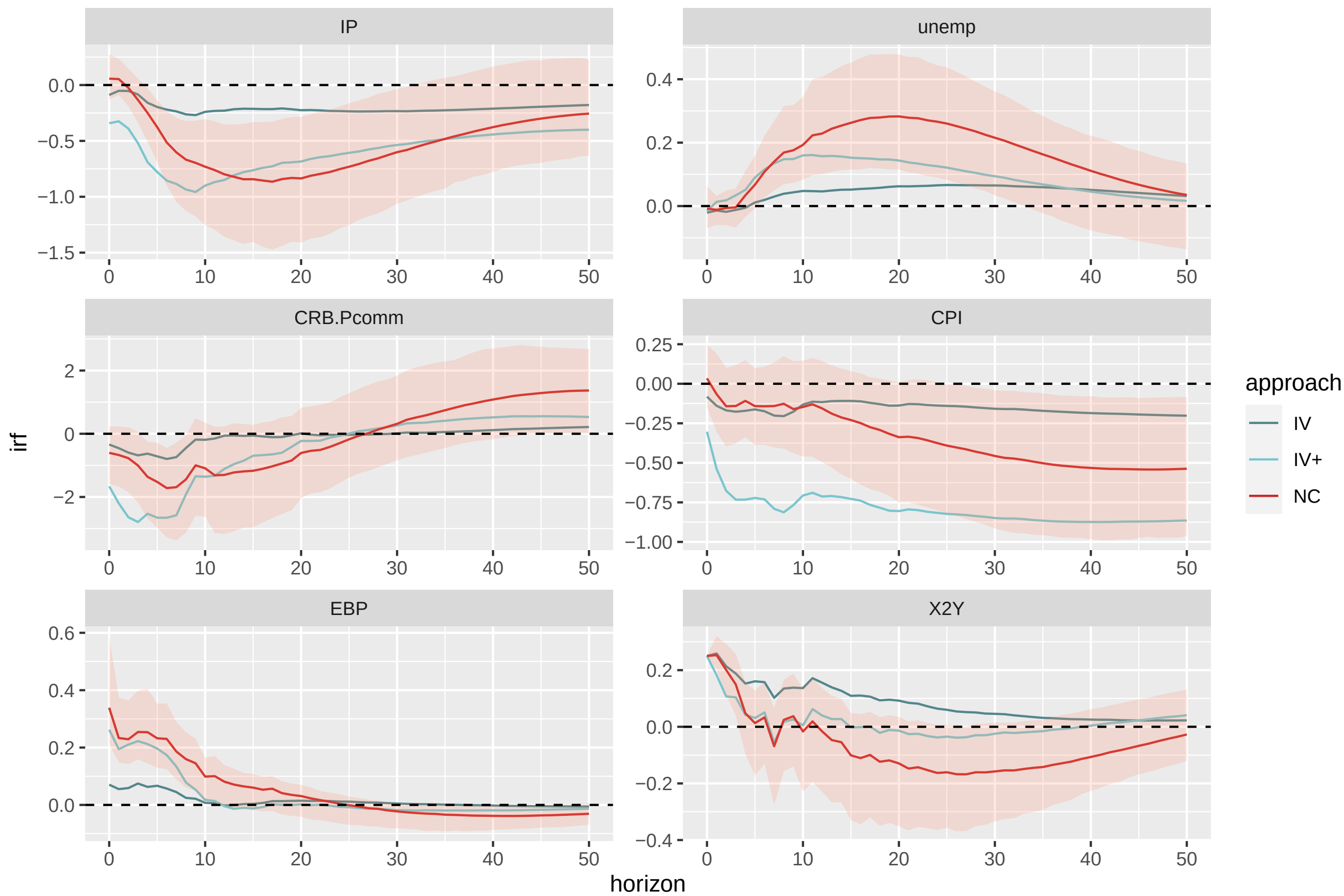



Figure 8: Effects of 5Y yield changes

Comparison of the negative control approach to IV approaches of Bauer and Swanson (2023*b*). Impulse response functions to a 5Y yield shock scaled to have 25 basis point impact on 2Y yield. The shaded region represents a 90% bootstrapped confidence interval. IP = Industrial Production, unemp = Unemployment, CRB.Pcomm = Commodity Price Index in levels, CPI = Consumer Price Index in levels, EBP = Excess Bond Premium, X2Y = 2Y treasury bond yield.

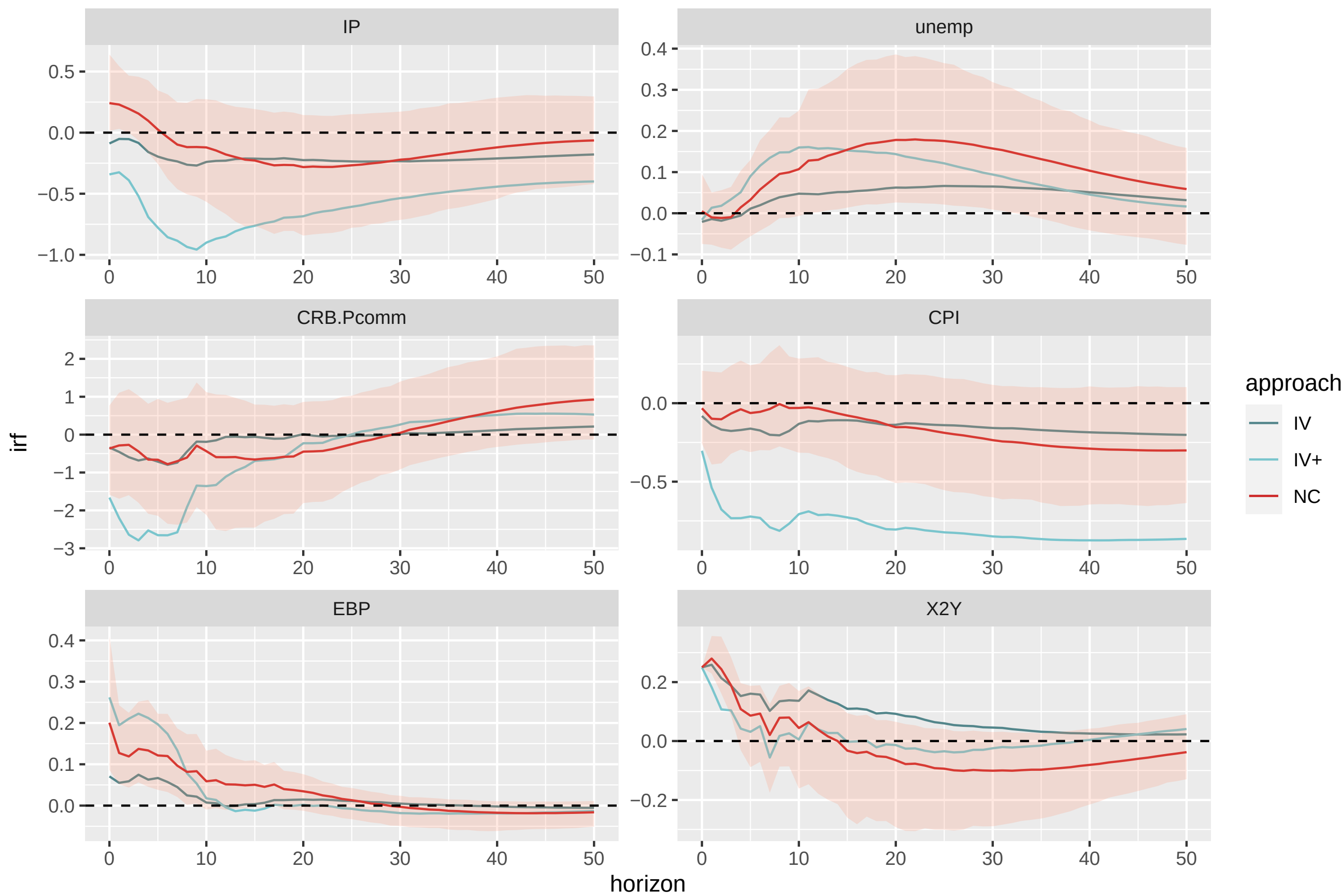



Figure 9: Effects of 20Y yield changes

Comparison of the negative control approach to IV approaches of Bauer and Swanson (2023*b*). Impulse response functions to a 20Y yield shock scaled to have 25 basis point impact on 2Y yield. The shaded region represents a 90% bootstrapped confidence interval. IP = Industrial Production, unemp = Unemployment, CRB.Pcomm = Commodity Price Index in levels, CPI = Consumer Price Index in levels, EBP = Excess Bond Premium, X2Y = 2Y treasury bond yield.