Stage-Based Identification of Policy Effects*

Christian Alemán
Université Libre de Bruxelles
ECARES

Christopher Busch
LMU Munich
CESifo

Alexander Ludwig
Goethe University Frankfurt
ICIR, UAB and CEPR

Raül Santaeulàlia-Llopis
NYU Abu Dhabi
UAB, BSE and CEPR

March 6, 2023

Abstract

We develop a method that identifies the effects of policy implemented nationwide across all regions at the same time. Our method consists of a normalization that maps the time-paths of regional outcomes onto a reference path—using only pre-policy data. After normalization, the normalized pre-policy outcome paths are identical—up to a minimization error. The stage of a regional outcome is defined as its location on (the support of) the reference path. Since regions can differ by stage at any point in time, our normalization uncovers heterogeneity in the stage at the time of policy implementation—even when implementation occurs at the same time across regions. We use this stage variation to identify the policy effects: a stage-leading region delivers the counterfactual path inside a window—in the stage domain—in which non-leading regions are subject to the policy whereas the leading region is not.

Keywords: Policy Effects, Identification, Stages
JEL Classification: C01, H00, E01, E22, E25

*We thank Alberto Abadie, Jaap Abbring, Francesco Agostinelli, Manuel Arellano, Brant Callaway, Jeffrey Campbell, Davide Cantoni, Xu Cheng, Hanming Fang, Hans-Peter Kohler, José-Víctor Ríos-Rull, Frank Schorfheide and participants at seminars and conferences for helpful comments and suggestions. Raül Santaeulàlia-Llopis acknowledges financial support from the AGAUR 2020PANDE00036 "Pandemies" Grant 2020-2022, from the Spanish Ministry of Economy and Competitiveness, through the Proyectos I+D+i 2019 Retos Investigacion PID2019-110684RB-I00 Grant, Europa Excelencia EUR2021-122011 and the Severo Ochoa Programme for Centres of Excellence in R&D (CEX2019-000915-S).
Alemán: christian.c.aleman@gmail.com; Busch: chris.busch.econ@gmail.com; Ludwig: mail@alexander-ludwig.com; Santaeulàlia-Llopis: loraulet@gmail.com

1 Introduction

Motivation. The empirical assessment of a policy requires a credible counterfactual. Standard empirical strategies critically rely on cross-regional heterogeneity in the time of policy implementation as source of identification—e.g., the existence of one untreated region or a staggered rollout.[1] Further, the credibility of the counterfactual requires the pre-policy paths of the outcome of interest to be similar across regions with differences not exceeding a constant gap over time—the so-called parallel trends assumption.[2] However, many relevant policy contexts violate these conditions. First, many policies are implemented nationwide, carried out in all regions at the same time, which eliminates the source of identification for standard empirical strategies. Second, the pre-policy outcome paths can be non-linear and differ across regions—e.g., in their starting date, speed, or magnitude—in a way that violates the parallel trends assumption. In panel (a) of Figure 1, we illustrate such a scenario. Our goal in this paper is to provide an identification strategy for such policy contexts.

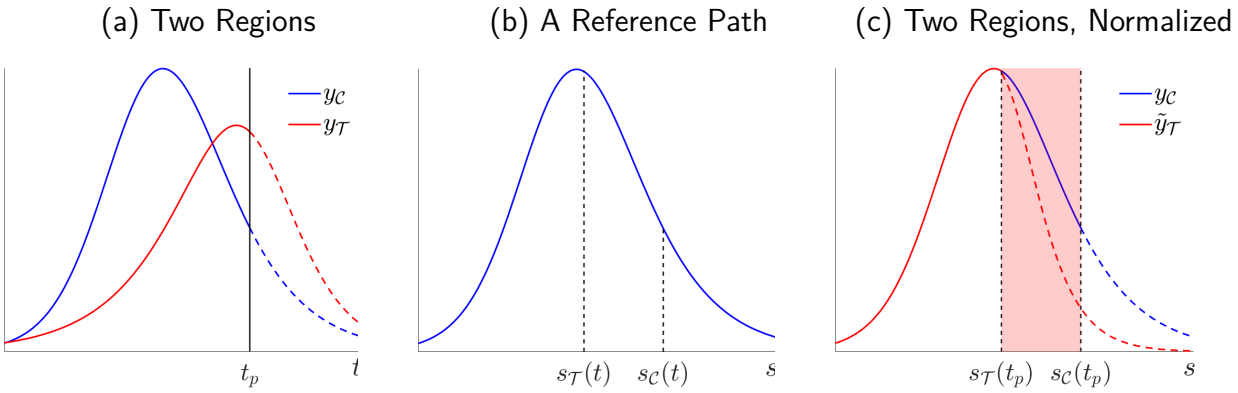
Idea. We develop the idea to track outcome paths in terms of stages (instead of time), where we define a stage of a regional outcome at time t as its location on the support of a reference outcome path. We illustrate this in panel (b) of Figure 1, which shows a reference outcome path that evolves over stages. At some time t, the outcome path of one region can be at a different stage—in the figure, the outcome of region $\mathcal{C}$ is at a more advanced stage than the outcome of region $\mathcal{T}$. Therefore, there is potential cross-regional heterogeneity in stages at the time of policy implementation, t_p. The core of our Stage-Based Identification (SBI) method is to trace out and to exploit this heterogeneity.

The Method. Our method consists of a normalization step and an identification step. First, we normalize the pre-policy outcome time paths of non-reference regions to the outcome time path of a reference region. Precisely, we conduct a transformation of time into stages (i.e. normalized time) and also normalize the level of the outcome of interest. Formally, this mapping is described by low degree polynomials and our normalization determines the coefficients of these polynomials—the normalization coefficients—such that the distance between the pre-policy outcome path of the reference region and each of the normalized pre-policy outcome paths of the non-reference regions is minimized. As a result, the pre-policy regional paths are identical—up to a minimization error—in the stage domain before the stage at which policy is implemented first

[1] See, among others, Angrist and Krueger (1999), Blundell and Macurdy (1999) and Imbens and Rubin (2015). See also the more recent discussions in Athey and Imbens (2017) and Card (2022).

[2] This assumption is relaxed in Abadie (2005) when there exist untreated regions and, more recently, in Callaway and Sant'Anna (2021) for staggered policies. See also recent related work by Rambachan and Roth (2021).

Figure 1: A Stage-Based Identification for Nationwide Policy: An Illustration



Notes: Panel (a) shows the time paths of an outcome variable for two regions, $\mathcal{C}$ and $\mathcal{T}$; policy is implemented in both regions at the same time t_p; dashed sections indicate post-policy paths. Panel (b) shows a reference path, where at some t, region $\mathcal{C}$ is at a more advanced stage thatn region $\mathcal{T}$. Panel (c) shows the result of mapping $y_{\mathcal{T}}$ onto $y_{\mathcal{C}}$, resulting in the normalized path $\widetilde{y}_{\mathcal{T}}$; the pink shaded area indicates the identification window.

across regions; see panel (c) of Figure 1 in which we use the outcome time path of region $\mathcal{C}$ as reference and normalize the outcome time path of region $\mathcal{T}$ to the reference.[3]

Second, our identification is based on the uncovered heterogeneity in the stages at which the policy is implemented across regions—hence the label, SBI. For example, in the illustration, policy is implemented at an earlier stage in region $\mathcal{T}$ than in region $\mathcal{C}$. The identification assumption is that the normalization coefficients that minimize the distance between the pre-policy outcome paths across regions in the stage domain are unaffected by policy. That is, we assume that these normalization coefficients would make the post-policy paths line up as well in the absence of policy. Applying these normalization coefficients on post-policy data opens an identification window in stages where a stage-leading region—in the example, region $\mathcal{C}$—is not subject to policy whereas the other region—in the example, region $\mathcal{T}$—is subject to policy, see pink shaded area in panel (c) of Figure 1.[4] Thus, the stage-leading region serves as control region for the other region, which is considered treatment.[5] The identified policy effects are captured by the area between the control region and treatment region inside the identification window.

Method Performance. We apply SBI to model-generated data in order to assess whether the identified policy effects recover the true policy effects that emerge from the model. We focus

[3]We show examples in which the normalization coefficients can be solved analytically, which provides an interpretation for these coefficients. In particular, our closed-form solutions show that the normalization coefficients reshape the structural parameters that determine the outcome path of the non-reference region into those of the reference region before policy is implemented; see Section 2.3.

[4]That is, the stage-leading region is the one in which policy is implemented at a more advanced (later) stage.

[5]Note that since the stage at which each region receives the policy is a result of our normalization, SBI does not require an ex-ante assignment of control or treatment across regions—this assignment is determined endogenously by the normalization coefficients in our approach, see Section 2.2.

on three nationwide policies that resemble our empirical applications: a stay-home policy against a pandemic using a model where economic activity shapes and is shaped by the pandemic in equilibrium; the approval of oral contraceptives in a model with women fertility and education choices; and the removal of an institutional barrier to economic growth in a model of structural transformation. Within the model framework, we know the true counterfactual path that would occur if policy were absent throughout, and thus we know the true policy effect. We apply SBI to the set of model-generated data that would be available to a policy evaluator and find that SBI can successfully identify the true effects of nationwide policy generated by the model.

We further assess whether and when our normalization procedure comes to its limits and plausible identification is not feasible using SBI. For this assessment, we perform a Monte Carlo study that numerically characterizes the bounds within which SBI is able to recover the true (model-generated) effects of policy. We consider one benchmark region which we pair up with a large set of regions (drawing from a large set of structural parameters) one-by-one. Applying SBI to these pairs, the error by which SBI captures the true policy effects systematically increases when moving away farther from the benchmark region in the space of the normalization coefficients. Under the interpretation that the normalization reshapes the structural parameters that determine the outcome paths, SBI requires—for sucessful identification of policy effects—the structural parameters to not be too dissimilar across regions before policy implementation. Along similar lines, we study how SBI fares in contexts where there are potential confounding factors such as time-varying latent heterogeneity, confounding policy and endogenous policy. Using model-generated data, we find that SBI is able to identify the true policy effects in these contexts, as long as the confounding factors keep the regional outcome paths sufficiently close in the space of the normalization coefficients. We also establish in a set of Placebo diagnoses that SBI successfully estimates a zero policy effect when there is none. We further show how to conduct inference on our identified policy effects with data that incorporates a stochastic component.

Three Applications. We apply SBI to study nationwide policy in three empirical applications. First, we assess the effectiveness of the stay-home policy implemented nationwide in response to the first wave of the Covid-19 pandemic in Spain. SBI assigns Madrid as the stage-leading (control) region at the time of policy implementation. We find that the stay-home policy significantly reduces the amount of deaths by 24.7% in the rest of Spain inside an identification window of seven days. In other words, had the stay-home policy not been implemented, there would have been 1,734 more deaths over the course of one week. Second, we assess the effects of the Food and Drugs Administration (FDA) approval of oral contraceptives (the pill) in the United States (U.S.) in 1960. We find that the pill reduced the crude fertility rate (number of births per 10,000) by 8.36%; with the stage-leading region of West Virginia and the effects are measured for the rest

of the U.S. We also find that the pill increased the proportion of college women by 24.9% during the decade that followed the FDA approval; where the stage-leading region is Washington DC and the effects are measured for the rest of the U.S. Third, we study the effects of the German reunification in 1990 on income per capita in West Germany where SBI assigns Hessen as the stage-leading region. Using the path of GDP per capita of Hessen as no-policy counterfactual, we find that the German Reunification significantly reduces income per capita of the rest of West Germany by 3.29% in a window of approximately 7 years.

Heterogeneous effects across stages. In order to assess how the policy effects potentially differ by stage, we focus on the Covid-19 application and map the paths of Spanish regions one-by-one onto the path of the leading region, Madrid. We find that the policy effect systematically varies with the stage at policy implementation (e.g., the number of prevented deaths is 65% in Murcia where the policy is implemented two weeks earlier in the stage domain than Madrid and 12% in the Basque Country where policy arrives two days earlier in the stage domain than Madrid). We further create a set of artificial regions using the power set of the non-reference regions (where each artificial region combines some autonomous regions) and reach similar insights.

Non-nationwide policy. SBI can also be applied to non-nationwide policy. When there is one untreated region, SBI delivers a right-open identification window bounded from below with the stage at which the policy is implemented in the treated region and unbounded from above. In the case of a staggered rollout where the policy is implemented across all regions but at different points in time, SBI delivers an identification window that—for a pair of regions—is bounded from below with the stage at which the policy is implemented first across regions and bounded from above with the stage in which the policy is implemented last across regions. It is worth noting that SBI can endogenously flip which region is considered control and which region is considered treatmemt with respect to the standard procedure in other empirical strategies studying staggered rollout. This flipping occurs when the region that implements the policy first in time is at a later stage than the stage at which the other region(s)—that implement the policy later in time—are.

Related literature. Our method is directly related to the standard empirical strategies designed for settings that resemble natural experiments. These strategies rely on a difference-in-differences methodology in order to generate the counterfactual path (or potential outcome as in Imbens and Rubin, 2015) that serves as control for a treated region—i.e., the region subject to policy. We emphasize two main differences of SBI. First, a critical common factor in previous strategies is that the source of identification relies on the heterogeneity in the time of policy implementation across regions either with the existence of one untreated region (e.g., Card, 1990; Card and Krueger, 2000) or a staggered policy adoption (e.g., Athey and Imbens, 2021; Borusyak et al.,

2021).[6] This is not the case in our method. Precisely, our main point of departure with respect to previous work is that SBI is able to deliver identification of policy effects for contexts in which the cross-regional heterogeneity in the time of policy implementation is absent. In this paper, we provide a new identification that uncovers cross-regional heterogeneity in the stage of the outcome of interest. Then, we use this cross-regional heterogeneity in stages at the time of policy implementation to identify the effects of policy, including nationwide policy.

Second, a relevant concern is that there might be cross-regional differences in the pre-policy determinants of the outcome of interest that also determine the outcome paths after policy. In standard empirical strategies, the policy effects are only credibly identified after controlling for these determinants, an idea that is typically conveyed through the parallel trends assumption (e.g. Bertrand et al., 2004, among many others).[7] Our approach to this question is rather different and does not rely on a parallel trends assumption for identification. Instead, we normalize pre-policy paths in a stage domain in a way that aims to minimize the cross-regional differences in the pre-policy determinants of outcomes, whether observable or not. Our identification assumption is that, absent policy, the normalization coefficients—obtained using pre-policy data only—would also map the post-policy path of the non-reference region onto the post-policy reference path. Thus, we apply the pre-policy normalization coefficients onto post-policy data in order to find a counterfactual path and identify the policy effects.

Our work also relates to other policy evaluation approaches like synthetic control methods (SCM) (Abadie and Gardeazabal, 2003; Abadie et al., 2010). The SCM approach essentially constructs a counterfactual time path based on a carefully weighted average across untreated (control) regions.[8] Two main differences stand out. First, analogous to other empirical strategies, SCM requires the existence of a set of untreated regions to construct the synthetic control group for identification. In contrast, SBI relies on cross-regional heterogeneity in the stage—not time—at which the policy is implementated. For this reason we can apply SBI to a nationwide policy occurring at the same time across all regions, unlike with SCM—or other methods for that matter. Second, our method does not require the use of observable potential determinants of cross-regional differences in the outcome path in order to generate the counterfactual. Instead, the counterfactual in our method is constructed using solely the time paths of the outcome of interest. Also, similar to SBI, the changes-in-changes method in Athey and Imbens (2006)

[6] See also the recent discussion in Goodman-Bacon (2021).

[7] In this context, there is a growing discussion on how to identify effects when parallel trends do not exactly hold. Abadie (2005) conditions the parallel trends to a set of observables using propensity scores (Heckman et al., 1998). This idea is extended to staggered rollout policy in Callaway and Sant'Anna (2021). Recently, Rambachan and Roth (2021) discuss how much the trends before policy implementation can differ from the trends after policy while still being able to identify causal effects.

[8] Doudchenko and Imbens (2017) use a joint framework for difference-in-difference and synthetic control groups.

features a mapping of an outcome variable from one region (group) to another. Their focus lies on capturing heterogeneity of the policy effect over the cross-sectional distribution of an individual level outcome. To this end, they map the cumulative cross-sectional distribution of one region on the cumulative cross-sectional distribution of the other region pre-policy and use this to construct the counterfactual distribution in the treated region.[9] Instead of cross-sectional distributions, we map pre-policy time paths of region-level outcomes. At the same time, the main difference described above remains, namely that the identification in changes-in-changes is also based on heterogeneity in the time of policy implementation across regions, whereas SBI does not require that heterogeneity.

Finally, in studies of the demographic transition (e.g., Greenwood et al., 2005) and structural transformation (e.g., Galor and Weil, 2000; Herrendorf et al., 2014; Cervellati and Sunde, 2015), the level of income per capita typically summarizes the "stage" of the outcome of interest (e.g. population growth, agricultural share, etc.) in cross-country comparisons. In contrast, rather than replacing time for an observable (here: income per capita), SBI conducts a normalization of the time path of the outcome of interest (possibly income per capita itself: see our evaluation of the effect of the German reunification on income per capita in West Germany). This also implies that the level of the outcome of interest (e.g., income per capita) is not a sufficient statistic to define the stage of a region in our approach. In the same manner, earlier work in Iorio and Santaeulàlia-Llopis (2010, 2016) also conducts a normalization of coordinates mapping country-specific HIV epidemic paths onto a reference path in order to define stages of the epidemic in a comparable manner across time and countries. We depart from that work in that we use our normalization to a reference path as base for identifying the effects of policies that aim to alter the path of the outcome of interest. For this reason, our normalization coefficients are obtained by strictly using pre-policy outcome paths.

The rest of the paper is structured as follows. We discuss our identification strategy in Section 2. We assess the performance of our method using model generated data in alternative policy contexts in Section 3. We conduct empirical applications in Section 4. We discuss heterogenous policy effects by stage, non-nationwide policy and spill-over effects in Section 5. Section 6 concludes.

[9] Their analysis is closely linked to Altonji and Blank (1999), who consider a decomposition of relative wage changes across groups into changes of the distribution of skills and the payoff for those skills.

2 A Stage-Based Method to Identify Policy Effects

To contextualize our contribution, we first briefly discuss how standard empirical strategies identify policy effects.[10] Consider a scenario in which, absent any policy intervention, the time path of an outcome $y_r(t)$ is identical across two regions $r \in \{\mathcal{C}, \mathcal{T}\}$.[11] Now assume that a policy is implemented only in region $\mathcal{T}$ at some date t_p which affects the outcome path in that region thereafter. Illustratively, we plot an outcome path of a treated region $y_{\mathcal{T}}(t)$ before policy implementation (solid red) and after policy implementation (dashed red) in panel (a) of Figure 2. We also show an outcome path for a region where the policy is not implemented, $y_{\mathcal{C}}(t)$ (solid blue). This scenario is ideal for the estimation of policy effects because the pre-policy outcome paths are identical across regions warranting the use of region $\mathcal{C}$ as control for region $\mathcal{T}$. That is, the outcome path $y_{\mathcal{C}}(t)$ provides a useful no-policy counterfactual to assess the effects of policy on $y_{\mathcal{T}}(t)$ after t_p. The effects of policy are captured by the difference between $y_{\mathcal{C}}(t)$ and $y_{\mathcal{T}}(t)$ in the interval (t_p, ∞). We can further add the implementation of the same policy to region $\mathcal{C}$ at some later date $t_p + \Delta$ with $\Delta > 0$; see panel (b) of Figure 2. Under this staggered rollout of the policy, the effects of policy on region $\mathcal{T}$ are identified using region $\mathcal{C}$ as counterfactual within the interval $(t_p, t_p + \Delta]$. In that interval, region $\mathcal{T}$ is subject to the policy whereas region $\mathcal{C}$ is not.

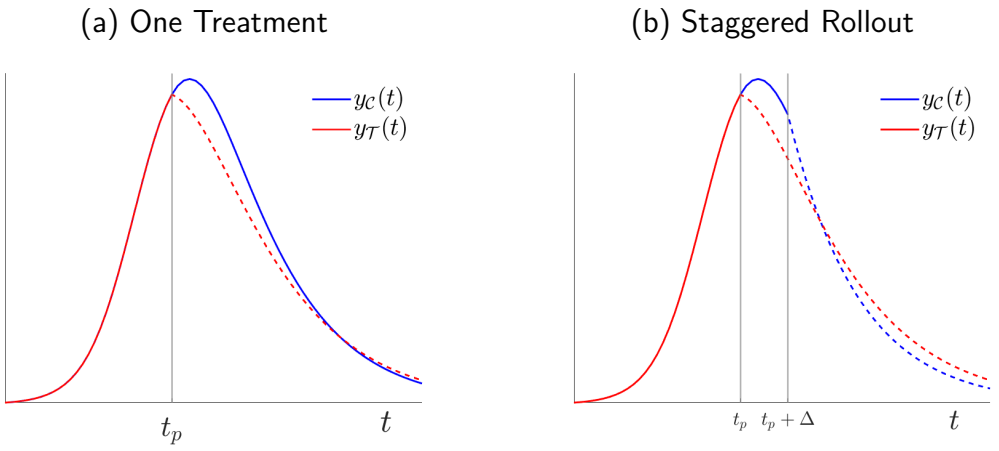
The standard identification strategies of policy effects just described fundamentally rely on two principles. First, the behavior of the outcome path before policy implementation must be credibly similar (the so-called parallel trends) across regions. Second, there must be variation in the time of policy implementation across regions which serves as source of identification. However, many policy contexts violate these conditions: First, the regional paths of the outcome variable before policy is implemented often differ across regions. In particular, outcome paths can differ by starting date, evolve at different speed and have different magnitude. Second, a large set of policies are implemented nationwide—i.e., carried out to *all regions* at the *same time*, which eliminates the source of identification used in standard strategies. We illustrate these two challenges in panel (a) of Figure 3 where a nationwide policy is implemented in a context where the outcome path in region $\mathcal{C}$ starts earlier, evolves at a faster speed and reaches a larger magnitude than in region $\mathcal{T}$.

Our strategy addresses these challenges in two steps. First, a normalization of regional outcome paths, and second, an identification based on the normalized paths.

[10] See comprehensive dicussions in, for example, Imbens and Rubin (2015) and Card (2022).

[11] We use *regions* in the description of the method due to the applications presented below. *Region* can be used interchangeably with *group* or *unit* throughout.

Figure 2: Ideal Policy Scenarios with Two Regions: Standard Identification Strategies



Notes: Denote with $y_{\mathcal{C}}(t)$ and $y_{\mathcal{T}}(t)$ the outcome paths of, respectively, region $\mathcal{C}$ and $\mathcal{T}$. Solid lines depict paths before policy implementation and dashed lines after policy. The identified policy effects are $\frac{\int_{\to t_p}^{h}(y_{\mathcal{C}}(t)-y_{\mathcal{T}}(t))dt}{\int_{\to t_p}^{h} y_{\mathcal{T}}(t)dt}$ with $h = \infty$ in the one-treatment case and $h = t_p + \Delta$ in the staggered rollout.

2.1 Normalization Procedure

Again, consider two regional outcome paths $y_r(t)$ with $r \in \{\mathcal{C}, \mathcal{T}\}$. We define the stage of a non-reference region as its location on the support of a reference path, which is the outcome path of a reference region. For this reference region, the stage is defined as time. For the non-reference region, the stage is the result of a normalization that maps its outcome time path onto the reference time path using only pre-policy data and uncovers stage variation at the time of policy implementation. We now describe our normalization—of the time and level of an outcome of interest—and provide a formal definition of stages afterwards.

The normalization starts with postulating the existence of the composite function,

$$\widetilde{y}_r(s) = (f_r \circ y_r \circ t_r)(s) = f_r(y_r(t_r(s))), \tag{1}$$

where $t_r(s) : S \to T$ is a stage-to-time transformation mapping stages $s \in S = \mathbb{R}$ into time $t \in T = \mathbb{R}$; $y_r(t) : T \to Y$ maps time into outcomes $y \in Y = \mathbb{R}$; and $f_r(y) : Y \to \widetilde{Y}$ maps outcomes into normalized outcomes $\widetilde{y}_r \in \widetilde{Y} = \mathbb{R}$. Thus, the composite function $\widetilde{y}_r(s) : S \to \widetilde{Y}$ defined in (1) maps stages s—i.e. normalized time—into normalized outcomes $\widetilde{y}$ for region r.

Without loss of generality, we treat the outcome path of region $\mathcal{T}$ as the reference path and that of region $\mathcal{C}$ as the non-reference path.[12] For the reference region, we set s to be a fixed point of $t_{\mathcal{T}}(.)$ for all s (i.e. $t = t_{\mathcal{T}}(s) = s$) and y to be a fixed point of $f_{\mathcal{T}}(.)$ for all y (i.e. $\widetilde{y} = f_{\mathcal{T}}(y) = y$) which implies that $\widetilde{y}_{\mathcal{T}}(s) = y_{\mathcal{T}}(s) = y_{\mathcal{T}}(t)$ always. Instead, for the non-reference region, we approximate $t_{\mathcal{C}}(.)$ and $f_{\mathcal{C}}(.)$ with $t_{\mathcal{C}}(.) \approx t(.;\boldsymbol{\psi}) = \sum_{k=0}^{K} \psi_k B_k^t(.)$ and $f_{\mathcal{C}}(.) \approx f(.;\boldsymbol{\omega}) = \sum_{m=0}^{M} \omega_m B_m^f(.)$, respectively. $\{B^f(.), B^t(.)\} \in \mathcal{B}^2$ are known basis functions in the space of continuous and differentiable functions. We denote the set of $M + K + 2$ unknown normalization coefficients by $\boldsymbol{\phi} = \{\boldsymbol{\psi}, \boldsymbol{\omega}\}$. This gives the composite function $\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}) = (f_{\mathcal{C}}(.;\boldsymbol{\omega}) \circ y_{\mathcal{C}} \circ t_{\mathcal{C}}(.;\boldsymbol{\psi}))(s) = f_{\mathcal{C}}(y_{\mathcal{C}}(t_{\mathcal{C}}(s;\boldsymbol{\psi}));\boldsymbol{\omega})$ by which we approximate $\widetilde{y}_{\mathcal{C}}(s)$:

$$\widetilde{y}_{\mathcal{C}}(s) \approx \widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}) = \sum_{m=0}^{M} \omega_m B_m^f \left(\left(y_{\mathcal{C}} \left(\sum_{k=0}^{K} \psi_k B_k^t(s) \right) \right) \right). \tag{2}$$

Then, conversely to the stage-to-time transformation, the stages in each region are defined as,

$$s = s_r(t;\boldsymbol{\psi}) = \begin{cases} t & \text{if } r = \mathcal{T} \\ t_{\mathcal{C}}^{-1}(s;\boldsymbol{\psi}) & \text{if } r = \mathcal{C}, \end{cases} \tag{3}$$

where for the reference region (here, $r = \mathcal{T}$), the stage at time t is the time itself (i.e. $s_{\mathcal{T}}(t;\boldsymbol{\psi}^*) = t = s$), whereas for the non-reference region (here, $r = \mathcal{C}$) the stage is the time in which region $\mathcal{C}$ is at the same stage than region $r = \mathcal{T}$ at time t (i.e. $s_{\mathcal{C}}(t;\boldsymbol{\psi}^*) = t_{\mathcal{C}}^{-1}(s_{\mathcal{T}}(t;\boldsymbol{\psi}^*);\boldsymbol{\psi}^*) = t_{\mathcal{C}}^{-1}(s;\boldsymbol{\psi}^*)$).

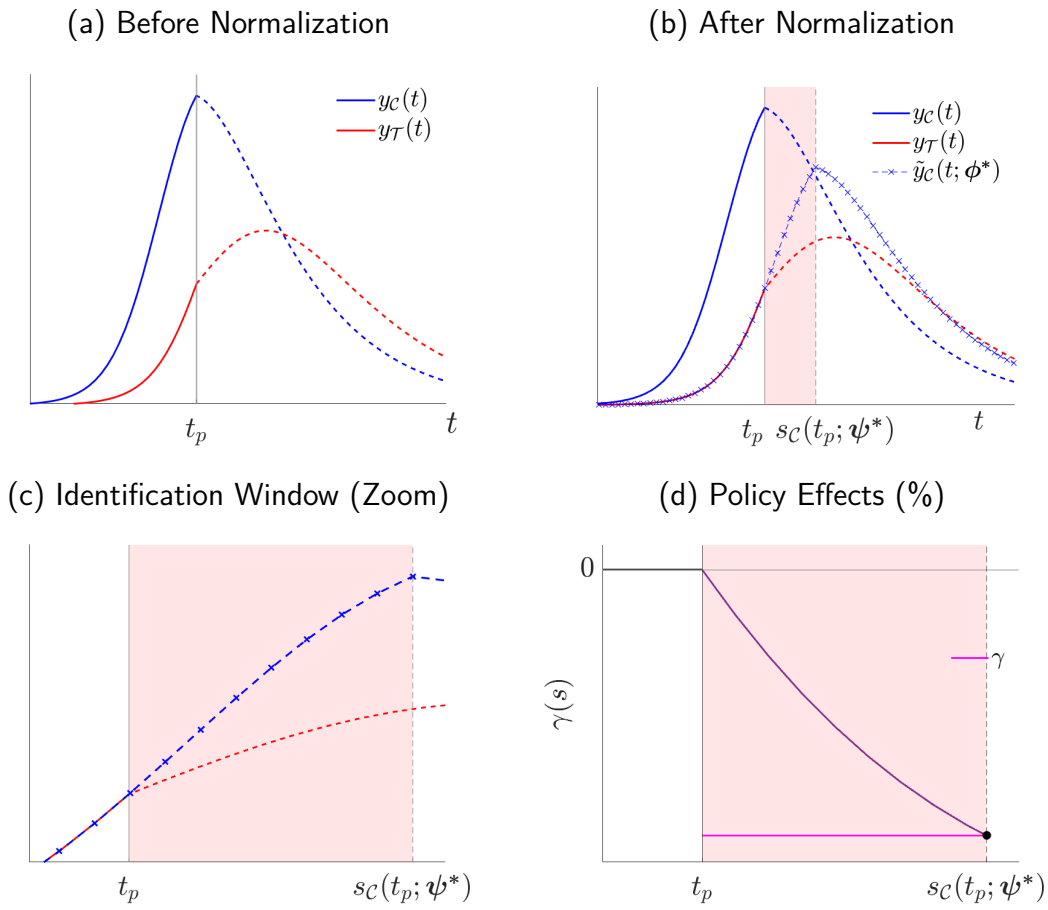
We choose monomials as benchmark for the basis functions $B^f(.)$ and $B^t(.)$ in (2). Then, the approximated normalized path of the non-reference region is,

$$\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}) = \sum_{m=0}^{M} \omega_m \left(y_{\mathcal{C}} \left(\sum_{k=0}^{K} \psi_k s^k \right) \right)^m. \tag{4}$$

A nice feature of the monomial basis is that it delivers a straightforward interpretation of the coefficients behind the stage-to-time transformation, $t_{\mathcal{C}}(s;\boldsymbol{\psi}) = \sum_{k=0}^{K} \psi_k s^k$. The parameter ψ_0 shifts the entire outcome path of region $\mathcal{C}$ forward (with $\psi_0 > 0$) or backwards (with $\psi_0 < 0$) in time, adjusting for different start dates. The parameter ψ_1 adjusts the speed in a constant way across periods. If $\psi_1 < 1$, then the outcome time-path of region $\mathcal{C}$ (in time) expands, whereas with $\psi_1 > 1$ it contracts. That is, if $\psi_1 < 1$, then region $\mathcal{C}$ is permanently faster (in time) than region $\mathcal{T}$—in one time-period region $\mathcal{C}$ advances by more than one stage, and vice versa for $\psi_1 > 1$. Further, allowing for the stage-to-time transformation to be quadratic (i.e. $\psi_2 \neq 0$) captures the notion that the relative speed across the regions can change over time: for example, the outcome path of region $\mathcal{C}$ might initially be slower than region $\mathcal{T}$, then catch up, and eventually move

[12] The choice of the reference region is innocuous, see our discussion in Section 2.2.

Figure 3: Stage-Based Identification of Policy Effects: A Nationwide Policy



Notes: In panel (d), we report the policy effects γ together with the interim cumulative effects of policy, $\gamma(s)$, as defined in Section 2.2.

faster. Throughout our analysis and applications, we typically set $K = 1$ and $M = 1$ in (4) and, hence, the normalized path of the non-reference region is $\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}) = \omega_0 + \omega_1 y_{\mathcal{C}}(\psi_0 + \psi_1 s)$.

Given observed time paths for all regions, i.e., $y_r(t)$ for $r \in \{\mathcal{C}, \mathcal{T}\}$, we determine the unknown coefficients $\boldsymbol{\phi} = \{\boldsymbol{\psi}, \boldsymbol{\omega}\}$ by minimizing the difference between the normalized path of the non-reference region, $\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi})$, and the outcome path of the reference region, $y_{\mathcal{T}}(s)$, that is:

$$\min_{\{\boldsymbol{\phi}\}} \|\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}) - y_{\mathcal{T}}(s)\|_{\mathbb{C}(s)}\,, \tag{5}$$

where $\|\cdot\|$ is the squared Euclidean distance defined on the interval of stages,

$$\mathbb{C}(s) = \left[s_{\overline{r}}(t_0;\boldsymbol{\psi}), s_{\underline{r}}(t_p;\boldsymbol{\psi})\right] \tag{6}$$

where $s_{\overline{r}}(t_0;\boldsymbol{\psi}) = \max\{s_r(t_0;\boldsymbol{\psi})\}$ and $s_{\underline{r}}(t_p;\boldsymbol{\psi}) = \min\{s_r(t_p;\boldsymbol{\psi})\}$ for $r = \{\mathcal{C},\mathcal{T}\}$. That is, the interval $\mathbb{C}(s)$ ensures that the minimization (5) only uses the outcome paths up to the stage s in which the policy is implemented first across regions, i.e. $s_{\underline{r}}(t_p;\boldsymbol{\psi})$. Note that the interval $\mathbb{C}(s)$ is determined endogenously during the minimization procedure. Now, we can define the stages of an outcome $y_r(t)$.

Definition 1. *The stage of an outcome $y_r(t)$ of region r at time t is $s_r(t;\boldsymbol{\psi}^*)$ where $\boldsymbol{\phi}^* \supset \boldsymbol{\psi}^*$ is the solution to the minimization of (5) subject to (4) and (6).*

In this way, the stages formally emerge as the result of our normalization procedure that maps the outcome path of a non-reference region onto the outcome path of the reference region before policy is implemented. To gain some intuition, we exemplify our method using a nationwide policy that affects the outcome paths of two regions, $y_{\mathcal{C}}(t)$ and $y_{\mathcal{T}}(t)$, in Figure 3. Before policy implementation at time t_p, the outcome path of region $\mathcal{C}$ (solid red) differs from region $\mathcal{T}$ (solid blue) in that it starts earlier, grows faster and is larger; see panel (a), which also shows the outcome paths after policy implementation for the two regions (dashed lines).[13]

The normalization procedure—i.e., the minimization of (5) subject to (4) and (6)—achieves two goals. First, it generates a normalized outcome path for the non-reference region in the stage domain, $\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}^*)$ (cross-dashed blue), that maps—up to minimization error—onto the outcome path of the reference region before the earliest stage in which policy is implemented across regions, $s_{\underline{r}}(t_p;\boldsymbol{\psi}^*)$; see panel (b) of Figure 3.[14] Second, since $s_{\underline{r}}(t_p;\boldsymbol{\psi}^*)$ is endogenous to $\boldsymbol{\psi}^*$, the normalization uncovers heterogeneity in the stage of policy implementation across regions.[15] For example, in our illustration, policy is implemented earlier—in stages—in region $\mathcal{T}$ than in region $\mathcal{C}$, i.e. $s_{\underline{r}}(t_p;\boldsymbol{\psi}) = s_{\mathcal{T}}(t_p;\boldsymbol{\psi}) < s_{\mathcal{C}}(t_p;\boldsymbol{\psi}) = s_{-\underline{r}}(t_p;\boldsymbol{\psi}^*)$ with $\underline{r} = \mathcal{T}$. Also, since we picked $\mathcal{T}$ to be the reference region, we obtain $s_{\underline{r}}(t_p;\boldsymbol{\psi}) = t_p$.

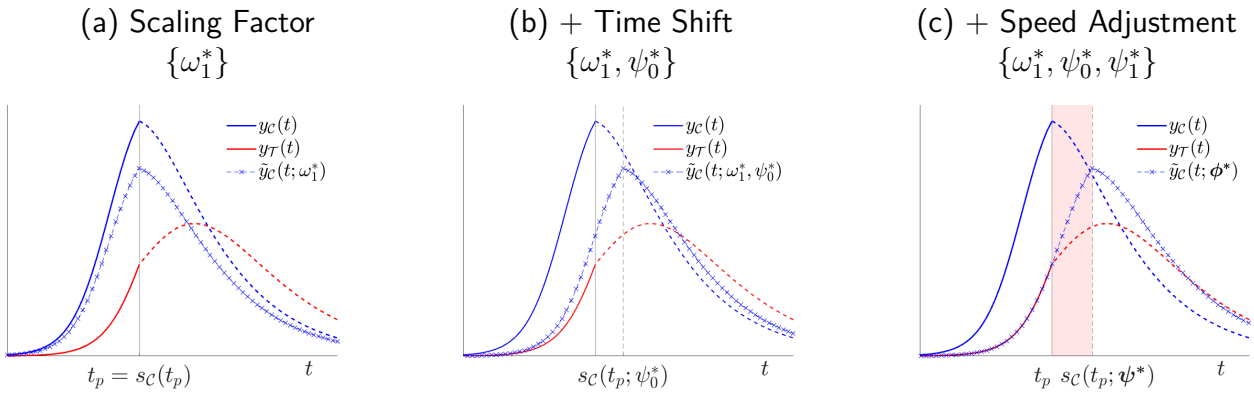
We further decompose the effects of each of the normalization coefficients $\{\phi_i\} \in \boldsymbol{\phi}^*$ on the path of the non-reference region in Figure 4. Since these coefficients are jointly determined in our minimization, we provide a non-orthogonal decomposition where we sequentially add the effects of each parameter. Note that in our illustration $y_{\mathcal{C}}(t_0) = y_{\mathcal{T}}(t_0) = 0$ and $\lim_{t\to\infty} y_{\mathcal{C}}(t) = \lim_{t\to\infty} y_{\mathcal{T}}(t) = 0$ and thus we focus on the role of the proportional level shifter ω_1 together with the stage-to-time transformation parameters ψ_0 and ψ_1—setting the constant level shifter

[13] In Appendix A, we provide additional examples.

[14] Note that outcome variables are typically observed on discrete dates. This is the case in all our applications. In these instances, since the mapping can generate dates $t_{\mathcal{C}}(s;\boldsymbol{\psi}^*)$ that are non-integer values—i.e., non-discrete dates—we interpolate between $y_{\mathcal{C}}(fl(t_{\mathcal{C}}(s;\boldsymbol{\psi}^*);\boldsymbol{\omega}^*)$ and $y_{\mathcal{C}}(cl(t_{\mathcal{C}}(s;\boldsymbol{\psi}^*);\boldsymbol{\omega}^*)$, where $fl(\cdot)$ and $cl(\cdot)$ denote the integer floor or integer ceiling, respectively.

[15] More generally, with more than two regions, $-\underline{r}$ refers to the complement set of $\underline{r}$, i.e. $-\underline{r} = \underline{r}^C$

Figure 4: Decomposition by Normalization Coefficient



Notes: We sequentially add the normalization parameters $\{\boldsymbol{\phi}\}$ to the non-reference path $y_{\mathcal{C}}(t)$ one-by-one.

to zero, $\omega_0 = 0$. In panel (a), we show that the coefficient $\omega_1^* < 1$ proportionally shifts down the entire outcome path of the non-reference region $\mathcal{C}$ throughout its support. In panel (b), the additional time shifter, $\psi_0^* > 0$, moves the outcome path to the right delaying the outcome's take off. In panel (c), adding the speed adjustment, $\psi_1^* < 1$, decreases the pace of the normalized outcome.[16]

2.2 Measuring the Policy Effects

In order to identify the policy effects, we exploit the fact that our normalization uncovers heterogeneity of the stage at the time of policy implementation, i.e. $s_{\underline{r}}(t_p;\boldsymbol{\psi}^*) < s_{-\underline{r}}(t_p;\boldsymbol{\psi}^*)$. In particular, we use the fact that inside a window (interval) of stages,

$$\mathbb{W}(s;\boldsymbol{\psi}^*) = \left[s_{\underline{r}}(t_p;\boldsymbol{\psi}^*), s_{-\underline{r}}(t_p;\boldsymbol{\psi}^*)\right], \tag{7}$$

region $\underline{r}$, i.e., the region where the policy is implemented first in stages, is subject to policy whereas region $-\underline{r}$ is not. In this context, we propose the following identification strategy:

Identification Assumption 1. *The normalization parameters $\boldsymbol{\phi}^*$ that solve the minimization of (5) subject to (4) and (6) are unaffected by policy.*

[16] Note that in the realm of standard empirical strategies one can partly address the time shift, ψ_0. This requires a choice by the researcher to fix the region-specific start dates of the outcome path of interest. For example, for the analysis of a Covid-19 containment policy this has been suggested in an event study design by Liu et al. (2021) and Glogowsky et al. (2021). In contrast, our method endogenously finds the appropriate time shifter (ψ_0) together with a speed adjustment (ψ_1).

That is, our identification assumes that, absent policy in region $\underline{r}$, the normalized path of the non-reference region obtained using ϕ^* and evaluated on stages $s > s_{\underline{r}}(t_p; \boldsymbol{\psi}^*)$ would yield a path identical to that of the reference region for all $s \in \mathbb{W}(s; \boldsymbol{\psi}^*)$.

Here, note that there is no ex-ante assignment to treatment or control for either reference or non-reference regions. Instead, the assignment of regions to treatment or control is determined endogenously (with $\boldsymbol{\psi}^*$) by the fact that policy arrives to the regions at different stages. We refer to the region that is at a more advanced (later) stage at the policy date as the stage-leading region. This region is then endogenously assigned to be the control region.[17] In the illustration, the stage-leading (control) region is $-\underline{r} = \mathcal{C}$, which is untreated inside $\mathbb{W}(s; \boldsymbol{\psi}^*) = [t_p, s_{\mathcal{C}}(t_p; \boldsymbol{\psi}^*)]$ and, hence, serves as no-policy counterfactual for the stage-lagging (treated) region $\underline{r} = \mathcal{T}$ inside that window; see panel (b) in Figure 4. The opposite roles (of reference and non-reference regions) would emerge if we picked $\underline{r} = \mathcal{C}$.[18]

Policy effect. Following our illustration, where the control region is $-\underline{r} = \mathcal{C}$ and the treated region is $\underline{r} = \mathcal{T}$, we measure the policy effect for the treated region as,

$$\gamma = \frac{\int_{\mathbb{W}(s;\boldsymbol{\psi}^*)} \left(y_{\mathcal{T}}(s) - \widetilde{y}_{\mathcal{C}}(s; \boldsymbol{\phi}^*)\right) ds}{\int_{\mathbb{W}(s;\boldsymbol{\psi}^*)} \widetilde{y}_{\mathcal{C}}(s; \boldsymbol{\phi}^*) ds}, \tag{8}$$

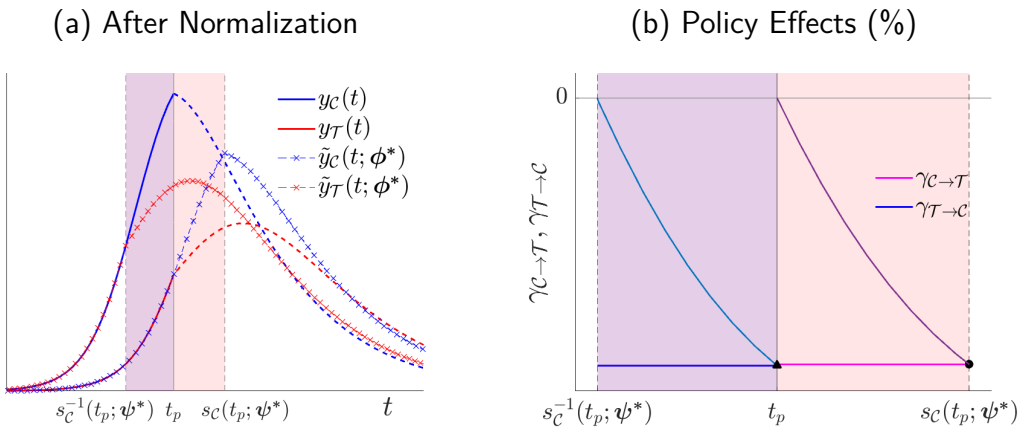
which measures the cumulative effect of policy relative to the scenario without policy in the treated region inside $\mathbb{W}(s; \boldsymbol{\psi}^*)$; see panel (c), Figure 3. The numerator is the area between the actual outcome path subject to policy of the treated region, i.e. $y_{\mathcal{T}}(s)$ (dashed red), and the no-policy counterfactual path for the treated region, i.e. $\widetilde{y}_{\mathcal{C}}(s; \boldsymbol{\phi}^*)$ (cross-dashed blue). The denominator captures the entire area below the no-policy counterfactual path for the treated region. In panel (d) of Figure 3, we zoom in on the identification window to show the policy effect γ together with the interim cumulative effects, $\gamma(s)$. Precisely, $\forall s \in \mathbb{W}(s; \boldsymbol{\psi}^*)$, we define $\gamma(s) = \frac{\int_{s_{\underline{r}}(t_p;\boldsymbol{\psi}^*)}^{s} \left(y_{\mathcal{T}}(\mathfrak{s}) - \widetilde{y}_{\mathcal{C}}(\mathfrak{s};\boldsymbol{\phi}^*)\right) d\mathfrak{s}}{\int_{s_{\underline{r}}(t_p;\boldsymbol{\psi}^*)}^{s} \widetilde{y}_{\mathcal{C}}(\mathfrak{s};\boldsymbol{\phi}^*) d\mathfrak{s}}$ for $\mathfrak{s} \in [s_{\underline{r}}(t_p; \boldsymbol{\psi}^*), s]$ where if $s = s_{-\underline{r}}(t_p; \boldsymbol{\psi}^*)$, then $\gamma(s) = \gamma$.

So far, we have used region $\mathcal{T}$ as reference. Since—as the normalization reveals—, $y_{\underline{r}}(s) = y_{\mathcal{T}}(s)$ is the treated path for all $s \in \mathbb{W}(s; \boldsymbol{\psi}^*) = [t_p, s_{\mathcal{C}}(t_p; \boldsymbol{\psi}^*)]$, the policy effect (8) measures the impact of policy on region $\mathcal{T}$ using as no-policy counterfactual the normalized path of region $\mathcal{C}$. Instead, reversing the reference region to $\mathcal{C}$—i.e., mapping the outcome path $y_{\mathcal{T}}(t)$ onto $y_{\mathcal{C}}(t)$ using pre-policy data and, consequently, redefining the stages for region $\mathcal{C}$ as $\boldsymbol{s} = \boldsymbol{t}_{\mathcal{C}}(\boldsymbol{s}) = t$

[17] We rule out the case of $s_{\underline{r}}(t_p; \boldsymbol{\psi}^*) = s_{-\underline{r}}(t_p; \boldsymbol{\psi}^*)$, where the identification window is empty and thus no identification is possible. Note that this is the case where the pre-policy time paths of the reference and non-reference regions are identical before normalization.

[18] We assess our identification assumption in the context of an analytical examples with exact identification in Section 2.3 and also through a placebo test with model-generated data without exact identification in Section 3.3.1.

Figure 5: Policy Effects with Alternative Reference Region



Notes: In panel (b), we report the policy effects γ together with the interim cumulative effects of policy, $\gamma(s)$, as defined in Section 2.2.

together with $\tilde{\boldsymbol{y}}_{\mathcal{C}}(\boldsymbol{s}) = y_{\mathcal{C}}(t = \boldsymbol{s})$ (and analogously for region $\mathcal{T}$ using (2) and (3))—implies that the normalized path of $\mathcal{T}$, $\tilde{\boldsymbol{y}}_{\mathcal{T}}(\boldsymbol{s})$, is the treated path for all $\boldsymbol{s} \in \mathbb{W}(\boldsymbol{s};\boldsymbol{\psi}^*) = [\boldsymbol{s}_{\mathcal{T}}(t_p;\boldsymbol{\psi}^*), t_p]$, where $\boldsymbol{\psi}^*$ $(\neq \boldsymbol{\psi}^*)$ is the coefficient vector that is obtained when choosing $\mathcal{C}$ as reference region. The associated policy effect measures the impact that the policy would have had on region $\mathcal{C}$ had it been treated at an earlier stage $\boldsymbol{s}_{\mathcal{T}}(t_p;\boldsymbol{\psi}^*)$, where we now compare the observed path of region $\mathcal{C}$ (which is untreated) to the normalized path $\tilde{\boldsymbol{y}}_{\mathcal{T}}(\boldsymbol{s},\boldsymbol{\psi}^*)$ (which is treated). Further, note that the two mappings are explicitly linked because $s_{\mathcal{C}}^{-1}(\mathfrak{t};\boldsymbol{\psi}^*) = \boldsymbol{s}_{\mathcal{T}}(\mathfrak{t};\boldsymbol{\psi}^*)$ for any $\mathfrak{t} \in T$ (e.g. $\mathfrak{t} = t_p$)—that is, the mapping $\mathcal{T}$ to $\mathcal{C}$ undoes the mapping $\mathcal{C}$ to $\mathcal{T}$, and vice versa.[19]

Going back to our example, we show the relationship between mappings in panel (a) of Figure 5 where using region $\mathcal{T}$ as reference implies that the normalization yields the identification window $\mathbb{W}(s;\boldsymbol{\psi}^*) = [t_p, s_{\mathcal{C}}(t_p;\boldsymbol{\psi}^*)]$ (pink shaded area) whereas, when using region $\mathcal{C}$ as reference, the normalization yields the identification window $\mathbb{W}(\boldsymbol{s};\boldsymbol{\psi}^*) = [s_{\mathcal{T}}(t_p;\boldsymbol{\psi}^*), t_p]$, which corresponds to $\left[s_{\mathcal{C}}^{-1}(t_p;\boldsymbol{\psi}^*), t_p\right]$ (purple shaded area). In panel (b) of Figure 5, we show the policy effects for both mappings. In the context of our illustrative example, the policy effects are identical across mappings because there are no level differences across regions, i.e., $\omega_0 = 0$. More generally, in instances where the reference region determines the units of the policy effect (i.e. $\omega_0 \neq 0$), a slight modification of the policy effects defined in (8) that explicitly takes into account the

[19] To see this, note that for any $\mathfrak{t} \in T$ (e.g. $\mathfrak{t} = t_p$), the stage function in the mapping $\mathcal{C}$ to $\mathcal{T}$, $s_{\mathcal{C}}(t;\boldsymbol{\psi}^*) = s = \mathfrak{t}$, injects t into $\mathfrak{t}$ whereas in the mapping $\mathcal{T}$ to $\mathcal{C}$ the function $\boldsymbol{s}_{\mathcal{T}}(\mathfrak{t};\boldsymbol{\psi}^*) = \boldsymbol{s} = t$ injects $\mathfrak{t}$ into t.

reference units makes irrelevant the choice of the reference region.[20] Nevertheless, our preferred measure of policy effects is (8) in which, as discussed above, the effects are interpreted differently across mappings.

2.3 Exact Identification of Policy Effects: Some Analytical Examples

We now discuss a setting in which we can explicitly express the normalization of the non-reference region in terms of the structural parameters of the data generating process. Note that this serves to illustrate the method, and to provide some guidance for interpretation of the normalization step. Indeed, if the data generating process were actually known, there would be no need to apply SBI. Our method operates under the proposition that if there exists a composite function (1) such that

$$\widetilde{y}_{\mathcal{C}}(s) = y_{\mathcal{T}}(s), \tag{10}$$

then our normalization procedure—the minimization of (5) subject to (2) and (6)—recovers the coefficients $\boldsymbol{\phi} = \{\boldsymbol{\psi}, \boldsymbol{\omega}\}$ up to a minimization error by approximating the functions $t_{\mathcal{C}}(.) \approx t_{\mathcal{C}}(.;\boldsymbol{\psi})$, $f_{\mathcal{C}}(.) \approx f_{\mathcal{C}}(.;\boldsymbol{\omega})$ and, hence, $\widetilde{y}_{\mathcal{C}}(.) \approx \widetilde{y}_{\mathcal{C}}(.;\boldsymbol{\phi})$ for all $s \in \mathbb{C}(s)$. Thus, under our identification assumption, we can identify the policy effects for all $s \in \mathbb{W}(s;\boldsymbol{\psi}^*)$.

In this context, here, we are interested in cases where (10) holds and (2) holds with equality and, hence, analytical solutions for the normalization coefficients $\boldsymbol{\phi}$ potentially exist for all $s \in \mathbb{C}(s)$. In that pursuit, consider a scenario in which the outcome path of a region r is,

$$y_r(t) = \left(1 - \gamma_{r,t}\mathbf{1}_{t \geq t_p}\right) g(t;\Theta_r), \qquad \text{for } t \in \{0, \ldots, t_p, \ldots, T\} \tag{11}$$

where Θ_r is a set of region-specific structural parameters that determine the behavior of the outcome path in that region and $\gamma_{r,t}$ captures the effect of policy that emerges after its nationwide implementation at t_p in regions $r = \{\mathcal{C}, \mathcal{T}\}$.

To identify the effects of the implemented policy, we apply SBI. We pick a region (e.g. $\mathcal{T}$) for the reference outcome path, $y_{\mathcal{T}}(t)$, and postulate a composite function (1) for the outcome

[20] Precisely, for a reference unit $y(\overline{s})$, the adjusted policy effects are:

$$\gamma\left(y(\overline{s})\right) = \frac{\int_{\mathbb{W}(s;\boldsymbol{\psi}^*)} \left(y_{\mathcal{T}}(s) - \widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}^*)\right) ds}{\int_{\mathbb{W}(s;\boldsymbol{\psi}^*)} \left(\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}^*) - y(\overline{s})\right) ds}. \tag{9}$$

E.g., if $y(\overline{s}) = y_r(s_{\underline{r}}(t_p;\boldsymbol{\psi}^*))$—where $y_{\mathcal{C}}(s_{\underline{r}}(t_p;\boldsymbol{\psi}^*)) = y_{\mathcal{T}}(t_p)$, then the adjusted policy effects (9) are measured with respect to the outcome of the reference region at the time of policy implementation. Note that although these adjusted policy effects are identical across mappings, now they depend on the choice of units, $y(\overline{s})$.

path of the non-reference region, $\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi})$. Here, we are interested in cases where (10) holds and (2) holds with equality because then $\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi})$ and $y_{\mathcal{T}}(s)$ share exactly the same functional form before policy is implemented first in the stage domain, i.e., for all $s \in \mathbb{C}(s)$ and, hence, we can recover—by the method of undetermined coefficients—the set of normalization coefficients by solving for ϕ in,

$$\Theta_{\mathcal{T}} = \widetilde{\Theta}_{\mathcal{C}}\left(\boldsymbol{\phi};\Theta_{\mathcal{C}}\right) \qquad \forall s \in \mathbb{C}(s), \tag{12}$$

which is a (potentially nonlinear) system with n equations—where n is the number of structural parameters—and with p unknowns—where $p = M + K + 2$ is the number of normalization coefficients in $\boldsymbol{\phi}$.[21] An interpretation of system (12)—and, hence, of our normalization procedure—is that the normalization coefficients $\boldsymbol{\phi}$ reshape the structural parameters of the non-reference, $\Theta_{\mathcal{C}}$, region into those of the reference region, i.e. $\Theta_{\mathcal{T}} = \widetilde{\Theta}_{\mathcal{C}}$ before policy implementation.

Theorem 2. *If there exists a composite function* (1) *such that* (10) *holds and* (2) *holds with equality for the regional outcome paths,* $y_r(t)$*, in* (11)*—i.e., if there exists a solution* $\boldsymbol{\phi}^*$ *for the system* (12)*—then stage-based identification (SBI) exactly and uniquely identifies the true policy effects for all* $s \in \mathbb{W}(s;\boldsymbol{\psi}^*)$*.*

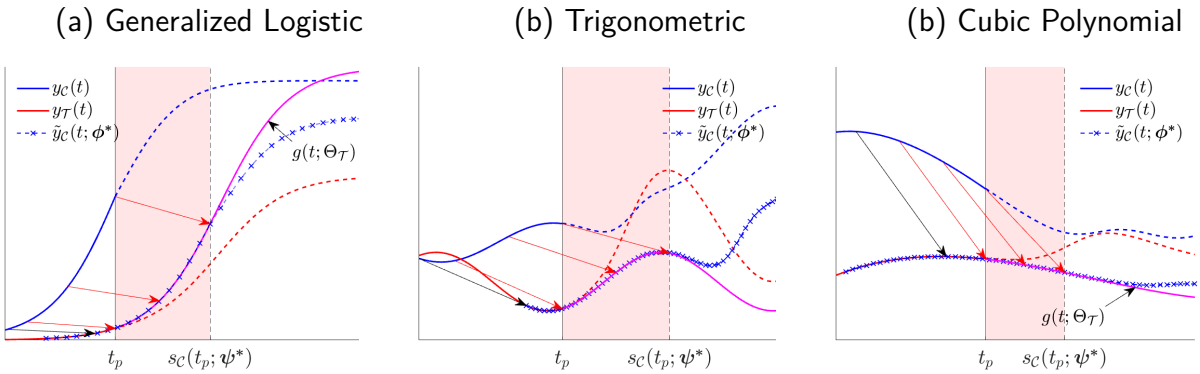
Proof. If the system (12) holds—i.e., there is an exact solution for $\boldsymbol{\phi}^*$—then the normalized outcome path of the non-reference region, $\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}^*)$, is exactly identical to the reference path, $y_{\mathcal{T}}(s)$ for all $s \in \mathbb{C}(s) = [s_{\overline{r}}(t_0,\boldsymbol{\psi}^*), s_{\underline{r}}(t;\boldsymbol{\psi}^*)]$. Since the outcome paths follow (11)—i.e., policy affects the path $y_r(t \geq t_p)$ but not the shape of $g(t;\Theta_r)$, then $\boldsymbol{\phi}^*$ is also a solution for the complement stage domain, i.e., for all $s \notin \mathbb{C}(s)$, in particular for $\mathbb{W}(s) = [s_{\underline{r}}(t_p;\boldsymbol{\psi}^*), s_{-\underline{r}}(t_p;\boldsymbol{\psi}^*)]$. This implies that $\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}^*)$ is exactly identical to $g(s;\Theta_T)$, that is, the true no-policy counterfactual of the reference region for all $s \in \mathbb{W}(s)$. Hence, SBI exactly and uniquely identifies the true policy effects for all $s \in \mathbb{W}(s;\boldsymbol{\psi}^*)$. □

Remark 1. Note that uniqueness of the normalization coefficients $\boldsymbol{\phi}^*$ is not necessary to recover unique policy effects.[22] To see this, note that although the presence of multiple solutions of $\boldsymbol{\phi}$ implies that there are multiple shapes for $f(.;\boldsymbol{\omega})$ and $t(.;\boldsymbol{\psi})$ that satisfy (10), the implied solution $\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi})$ for (10) is unique and, hence, so is the identified policy effect.

[21] That is, here, the minimization step in the normalization (in Section 2.1) is the solution to the system (12) emerging from the undetermined coefficients approach.

[22] Note that if (10) holds, (2) holds with equality and the inverse function $\boldsymbol{\phi} = \widetilde{\Theta}_{\mathcal{C}}^{-1}\left(\Theta_{\mathcal{T}};\Theta_{\mathcal{C}}\right)$ exists, then there exists a unique solution $\boldsymbol{\phi}^*$ for the system (12). This sufficiency for existence and uniqueness of $\boldsymbol{\phi}^*$ coincides with the Rouché–Frobenius Theorem in the cases where the system (12) is linear.

Figure 6: SBI with Exact Identification: Three examples



Notes: These panels show exact identification when $g(t;\Theta_r)$ is: (a) the first derivative of the general logistic function $\frac{\theta_{0,r}}{[1+\exp(-\theta_{2,r}t+\theta_{1,r})]^{\frac{1}{\theta_{3,r}}}}$; (b) a trigonometric function $\theta_{0,r}+\theta_{1,r}sin(\theta_{2,r}+\theta_{3,r}t)$;and (c) a cubic polynomial $\sum_{j=0}^{3}\theta_{j,r}t^j$, respectively. The analytical derivations are in Appendix B.

We now discuss some functional forms for the outcome paths, $y_r(t)$, for which SBI yields analytical solutions for ϕ using the approach just described. We start with generalized logistic functions. Assume that the regional outcome paths $y_r(t)$ are determined by (11) and that, absent policy, these paths are determined by $g(t;\Theta_r)$ with,

$$g(t;\Theta_r)=\frac{\theta_{0,r}}{\left[1+\exp\left(-\theta_{2,r}t+\theta_{1,r}\right)\right]^{\frac{1}{\theta_{3,r}}}} \tag{13}$$

where $\Theta_r=\{\theta_{0,r},\theta_{1,r},\theta_{2,r},\theta_{3,r}\}$ is a set of region-specific structural parameters that determine the behavior of the outcome paths for regions $r=\{\mathcal{C},\mathcal{T}\}$. We show an illustration of these paths for region $\mathcal{C}$ and $\mathcal{T}$ in panel (a) of Figure 6.[23] To identify the policy effects, we apply SBI to the log of (13), i.e., we let the outcome variable in region r be $y_{t,r}=\ln\left(g(t;\Theta_r)\right)$. We next pick a region (e.g. $\mathcal{T}$) for the reference path and postulate a composite for the non-reference region, $\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi})=\omega_1 y_{\mathcal{C}}(\psi_0+\psi_1 s)+\omega_0$. Then , we solve for the normalization coefficients $\boldsymbol{\phi}=\{\psi_0,\psi_1,\omega_0,\omega_1\}$ in (10) holding (2) with equality for all $s\in\mathbb{C}(s)$, that is,

$$\begin{aligned}\widetilde{y}_C &= \omega_0+\omega_1 y_C\left(\psi_0+\psi_1 s\right)\\ &= \underbrace{\omega_0+\omega_1\ln\left(\theta_{0,\mathcal{C}}\right)}_{=\ln\left(\theta_{0,\mathcal{T}}\right)}-\underbrace{\omega_1\frac{1}{\theta_{3,\mathcal{C}}}}_{=\frac{1}{\theta_{3,\mathcal{T}}}}\ln\left(1+\exp\left(-\underbrace{\theta_{2,\mathcal{C}}\phi_1}_{=\theta_{2,\mathcal{T}}}s+\underbrace{\theta_{1,\mathcal{C}}-\phi_0\theta_{2,\mathcal{C}}}_{=\theta_{1,\mathcal{T}}}\right)\right)=y_{\mathcal{T}}(s)\end{aligned}$$

[23] We assume that the outcome path for region $\mathcal{C}$ takes off earlier, $-\frac{\theta_{1,\mathcal{C}}}{\theta_{2,\mathcal{C}}}>-\frac{\theta_{1,\mathcal{T}}}{\theta_{2,\mathcal{T}}}$, grows faster $\theta_{2,\mathcal{C}}>\theta_{2,\mathcal{T}}$, reaches a higher maximum $\theta_{0,\mathcal{C}}>\theta_{0,\mathcal{T}}$ and has larger asymmetry $\theta_{3,\mathcal{C}}>\theta_{3,\mathcal{T}}$.

and, thus, by the method of undetermined coefficients we find $\boldsymbol{\phi} = \{\psi_0, \psi_1, \omega_0, \omega_1\}$ as the analytical solution of $\boldsymbol{\phi} = \widetilde{\Theta}_{\mathcal{C}}^{-1}(\Theta_{\mathcal{T}}; \Theta_{\mathcal{C}})$ given by

$$\omega_1 = \frac{\theta_{3,\mathcal{C}}}{\theta_{3,\mathcal{T}}}, \qquad \omega_0 = \ln(\theta_{0,\mathcal{T}}) - \frac{\theta_{3,\mathcal{C}}}{\theta_{3,\mathcal{T}}} \ln(\theta_{0,\mathcal{C}})$$
$$\phi_0 = \frac{\theta_{1,\mathcal{C}} - \theta_{1,\mathcal{T}}}{\theta_{2,\mathcal{C}}}, \qquad \phi_1 = \frac{\theta_{2,\mathcal{T}}}{\theta_{2,\mathcal{C}}}.$$

The normalization uncovers cross-regional stage heterogeneity at the time of policy implementation: in our illustration, the non-reference region is at a more advanced stage than the reference region at t_p, i.e. $s_{\underline{r}}(t_p; \boldsymbol{\psi}) = t_p$. This opens a window in the stage domain in which region $\mathcal{T}$ is subject to policy whereas region $\mathcal{C}$ is not, i.e., $\mathbb{W}(s; \boldsymbol{\psi}^*) = \left[t_p, s_{\mathcal{C}}(t_p, \boldsymbol{\psi}^*) = \frac{\theta_{1,\mathcal{T}} - \theta_{1,\mathcal{C}}}{\theta_{2,\mathcal{T}}} + \frac{\theta_{2,\mathcal{C}}}{\theta_{2,\mathcal{T}}} t_p\right]$. Then, under our identification assumption, the normalized outcome path of the non-reference region, $\widetilde{y}_{\mathcal{C}}(s; \boldsymbol{\phi})$, serves as no-policy counterfactual for the reference region for all $s \in \mathbb{W}(s)$. Indeed, since the outcome paths follow (11), $\boldsymbol{\phi}$ is also an analytical solution for the complement stage domain, i.e., for all $s \notin \mathbb{C}(s)$. That is, for all $s \in \mathbb{W}(s)$, the normalized outcome path of the non-reference region, i.e. $\widetilde{y}_{\mathcal{C}}(s; \boldsymbol{\phi})$, is identical to the reference path in the no-policy scenario, i.e. $g(t; \Theta_{\mathcal{T}})$ (solid magenta line); see panel (a), Figure 6. Thus, the identified policy effect is unique and identical to the true policy effect.

We repeat this analysis for outcome paths that follow trigonometric functions and polynomials. In panel (b) of Figure 6, we show the results of applying SBI for the case where $g(t; \Theta_r) = \theta_{1,t} \sin(\theta_{3,r} t + \theta_{2,t}) + \theta_{0,t}$. In panel (c) of Figure 6, we show the results of applying SBI for the case where $g(t; \Theta_r) = \sum_{i=0}^{I} \theta_i t^{i=1}$ with $I = 3$. In both cases, there is a unique solution for $\boldsymbol{\phi}^*$; see our derivations in Appendix B. In both cases, the identified policy effect is unique and identical to the true policy effect.

3 Method Performance

More generally, we are interested in analyzing policy in contexts where the data generating process is unknown. Here, we assess whether SBI is able to recover the true policy effects that emerge from theoretical models in such scenarios. First, we implement SBI on model-generated data (without using any knowledge about the model) and compare the identified effects with the true effects in Section 3.1. Second, we conduct a Monte Carlo analysis that provides bounds to the performance of our method in Section 3.2. There, we further assess how our method fares with time-varying heterogeneity, confounding policy, and endogenous policy. Third, we conduct inference in Section 3.3.

3.1 Does SBI Identify the True Policy Effects?

To address this question, we use three alternative policy contexts: a public health policy against a pandemic using a model where economic activity in the form of hours worked shapes and is shaped by a pandemic; the effects of the approval of the pill in a model of women career and fertility choices; and an economic growth policy using a model of structural transformation.

3.1.1 Public Health Policy Against a Pandemic

Here, we pose an economic model in the context of an epidemic where labor supply generates infections (and deaths). Then, we assess whether SBI recovers the model-generated effects of a nationwide lockdown that restricts labor supply after some period t_p.

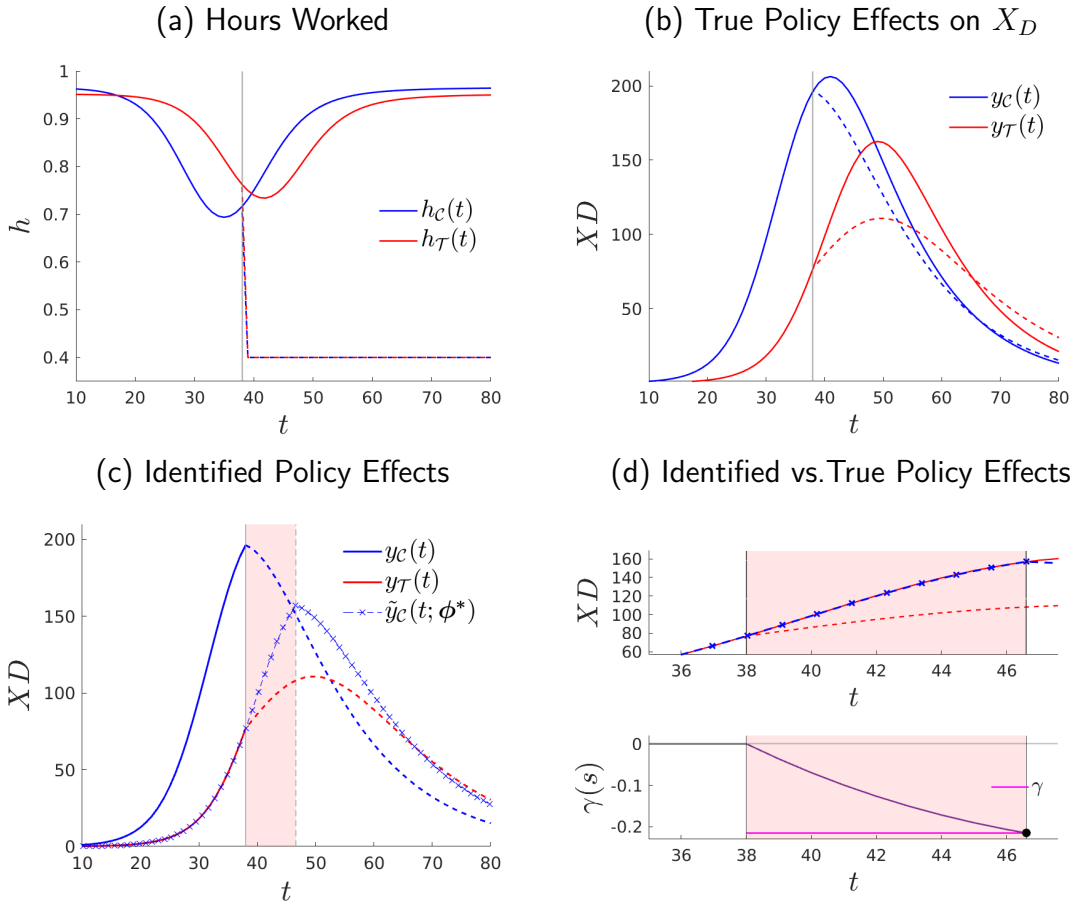
At the beginning of each period $t \in \{0, 1, \ldots\}$, total population N_t is composed of a stock of susceptible population S_t, infected individuals I_t and recovered individuals R_t, with $N_t = S_t + I_t + R_t$ and the normalization $N_0 = 1$. An epidemic starts with an initial number of infected $I_1 > 0$ in period $t = 1$. For pre-pandemic periods $t < 1$, the population is constant with $N_0 = S_0$ and $I_0 = R_0 = 0$. The probability that a susceptible individual meets an infected individual is given by $\beta \frac{I_t}{N_t}$, for $\beta \in (0, 1)$.[24] We assume that conditional on meeting there exists an objective probability $\lambda_{\mathcal{O}}(h_t)$ of getting infected which depends on economic activity here reflected by the average hours worked h_t. Further, with probability μ infected individuals in a given period t recover or die from the disease where the conditional probability of death in turn is denoted by ζ. New infections transit to death in the same period t, i.e. h_t has an immediate effect on the survival rate between t and $t + 1$.[25]

In this context, we consider the problem of a social planner that is constrained in that she has imperfect knowledge about the infection process. In particular, the planner's beliefs of the infection probability are $\lambda_{\mathcal{P}}(h_t)$, which may differ from the objective probability. Specifically, let $\lambda_b(h_t) = \xi_b h_t^{\alpha}$, $\xi_b > 0$ and $\alpha \in (0, 1)$, for beliefs $b \in \{\mathcal{O}, \mathcal{P}\}$, where $\mathcal{O}$ stands for objective and $\mathcal{P}$ for perceived. Thus, if $\xi_{\mathcal{P}} < \xi_{\mathcal{O}}$ then the constrained planner underestimates the actual effects of average hours worked h_t on infections and vice versa if $\xi_{\mathcal{P}} > \xi_{\mathcal{O}}$.

At every period t, before making plans for all future periods $z \geq t$, the planner receives an unanticipated knowledge shock that reveals the actual state of the economy $G_{\mathcal{O},t}$ for $G = (S, I, R, D)$, which potentially differs from the perceived state $G_{\mathcal{P},t}$. We assume that the planner updates the perceived survival probability accordingly and before choosing labor supply. Precisely,

[24] Parameter β captures features like density, health or pollution (among others) which can differ across regions.
[25] This innocuous assumption eases the exposition of the trade-off between economic activity and public health.

Figure 7: Stage-Based Identification of Model-Generated Policy Effects: A Nationwide Public Health Policy Against a Pandemic



Notes: We assume that $u(c_z, h_z) = \ln(c_z) - \kappa \frac{h_z^{1+\frac{1}{\nu}}}{1+\frac{1}{\nu}} + \chi$ for value of life parameter χ. Some parameters differ across regions: $\Theta_{\mathcal{C}} = \{\beta = 0.509, \zeta = 0.0010, \kappa = 1.05, \xi = 0.20, I_0 = 1\}$ and $\Theta_{\mathcal{T}} = \{\beta = 0.501, \zeta = 0.0008, \kappa = 1.07, \xi = 0.19, I_0 = 6\}$. The rest of the model parameters are identical across regions, $\{\delta = 0.95, \chi = 560400, z = 64, \beta = 0.501, \alpha = 0.65\}$. The parameters associated to the policy are $\bar{h} = 0.4, t_p = 38, t_f = 250$.

letting $X_{G,b,t} = G_{b,t+1} - G_{b,t}$, the planner's perceived survival probability is revised at the beginning of every period t to $\phi_{\mathcal{P}}(h_t) = 1 - \frac{X_{D,\mathcal{P},t}}{\tilde{N}_{b,t}}$ with $\tilde{N}_{b,t} = N_{\mathcal{O},t}$ for $t = z$ and $\tilde{N}_{b,t} = N_{\mathcal{P},t}$ if $t > z$.[26] Note that although the knowledge shock allows the planner to update the state of the economy at the beginning of every period t, however, since these shocks are unanticipated, the planner is unable to correct future forecast errors, i.e. $G_{\mathcal{O},z} - G_{\mathcal{P},z}\,|t$ for periods $z > t$.[27]

[26] Note that without subjective beliefs, the population evolves essentially as in, for example, Atkeson (2020).

[27] The forecast errors $\varepsilon_{G,z} = (G_{\mathcal{O},z} - G_{\mathcal{P},z}\,|t)$ can be reduced asymptotically with learning (Adam et al., 2017). For example, there could be learning about the odds of infection as in Aleman et al. (2022).

After updating the perceived survival probability, the planner maximizes the present-discounted stream of per period utilities for all periods $z \geq t$ with discount factor δ times the perceived unconditional probability to survive from any period t to the future, $\prod_{j=t+1}^{z} \phi_{\mathcal{P}}(h_{j-1})$. Importantly, since the perceived survival probability is revised at the beginning of every period t, the nature of the discounting process changes each period t and, hence, the planner needs to re-optimize—at each period t—the decision plans for all periods $z \geq t$. The per period utility, $u(c_z, h_z; \chi)$ is assumed strictly concave in consumption $c_z \geq 0$ and leisure $1 - h_z \in [0, 1]$ for a value of life parameter χ. Collecting elements, at each period t the constrained social planner solves,

$$\max_{\{c_z \geq 0, h_z \in [0,1]\}_{z=t}^{\infty}} \sum_{z=t}^{\infty} \delta^{z-t} \prod_{j=t+1}^{z} \phi_{\mathcal{P}}(h_{j-1}) u(c_z, h_z; \chi), \tag{14}$$

subject to an aggregate resource constraint $N_{\mathcal{P},z} c_z = w h_z N_{\mathcal{P},z}$ where w is the implicit price (marginal product) of labor using technology $Y_z = a h_z N_{\mathcal{P},z}$.

Then, given t, the amount of economic activity h_z is determined by the following condition,

$$\underbrace{\frac{\partial u(c_z, h_z; \chi)}{\partial c_z} w}_{\substack{\text{Marginal Benefit of Working:} \\ \text{Consumption Gain}}} - \underbrace{\frac{\partial u(c_z, h_z; \chi)}{\partial h_z}}_{\substack{\text{Marginal Cost of Working:} \\ \text{Loss of Leisure}}} = \underbrace{\delta \frac{\partial \phi_{\mathcal{P}}(h_z)}{\partial h_z} u(c_{z+1}, h_{z+1}; \chi)}_{\substack{\text{Marginal Cost of Working:} \\ \text{Loss of Lives}}} \qquad \forall z \geq t, \tag{15}$$

stating that the marginal benefit of working (more consumption) equates its marginal costs consisting of an intratemporal component (disutility from working) and an intertemporal component (loss of lives). Since the Euler equation (15) is a first-order difference equation in h_z, we can easily solve for the optimal labor path during the epidemic using standard techniques; see Appendix C.1.

True (model-generated) policy effects. We solve the model for two regions that differ in the underlying parameter values for $\Theta = \{\delta, \chi, a, \beta, \mu, \zeta, \kappa, \nu, \{\xi_i\}_{i \in \{\mathcal{O},\mathcal{P}\}}, \alpha, I_1\}$. In particular, we assume that the planner in region $\mathcal{C}$ underestimates the effect that economic activity has on infections by less than the planner in region $\mathcal{T}$. Consequently, hours are reduced earlier and also by a larger amount in region $\mathcal{C}$ than region $\mathcal{T}$ in response to the epidemic. The equilibrium response of hours without policy intervention for region $\mathcal{C}$ (solid blue) and region $\mathcal{T}$ (solid red) are shown in panel (a) of Figure 7. The earlier and stronger response in terms of hours of region $\mathcal{C}$ affects our outcome of interest, i.e., the epidemic path of deaths, by reducing the peak of deaths and flattening the curve in region $\mathcal{C}$ relative to region $\mathcal{T}$; see panel (b) in Figure 7. We also assume that region $\mathcal{C}$ has higher odds of encountering infected individuals at work (i.e., higher β) which advances and increases the peak of deaths for region $\mathcal{C}$ relative to region $\mathcal{T}$. Further, we assume

that region $\mathcal{C}$ has a lower disutility of work κ which implies a larger pre- and post-pandemic level of hours worked for region $\mathcal{C}$ than region $\mathcal{T}$.

In this scenario, we now introduce a nationwide public health policy that imposes an upper bound on hours worked, $h < \overline{h} = 0.5$, from $t_p = 38$ to $t_f = 250$. Since, without policy, households in both regions would work more hours than $\overline{h}$, the policy is binding in both $\mathcal{C}$ and $\mathcal{T}$—see the respective dashed lines that emerge after t_p in panel (a) of Figure 7. The lower economic activity imposed by the policy has consequences for the flow of deaths. With policy, the flow of deaths peaks earlier and by a lower magnitude in both $\mathcal{C}$ and $\mathcal{T}$—see the respective dashed lines that emerge after t_p in panel (b) of Figure 7. The difference between the flow of deaths with policy (dashed lines) and the flow of deaths without policy (solid lines) after t_p captures the true effects of policy generated from the model. However, the counterfactual paths of the flow of deaths without policy after policy implementation (i.e. the solid lines after t_p) are not available outside of the model. That is, from the perspective of an evaluator that wants to assess the policy effects, the data available for policy evaluation consists of the path without policy (solid lines) for all periods up to t_p along with the path with policy (dashed lines) for all periods after t_p. We now apply SBI on this data.

Stage-Based identified policy effects. The policy effects identified in this manner are shown in panel (c) of Figure 7. In particular, we map the path of the flow of deaths in region $\mathcal{C}$ (solid blue line) onto the path of region $\mathcal{T}$ (solid red line) using *only* pre-policy data; as described in Section 2. The result of SBI is a candidate no-policy counterfactual $\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}^*)$ (blue line with cross markers) for region $\mathcal{T}$ in the identification window $\mathbb{W}(s;\boldsymbol{\psi}^*) = [t_p, s_{\mathcal{C}}(t_p;\boldsymbol{\psi}^*)]$ (shaded pink area). In order to assess whether the identified policy effects recover the true policy effects generated by the model, we zoom in on the identification window in panel (d) of Figure 7 and compare our candidate counterfactual $\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}^*)$ with the true counterfactual (solid red line). The main result is that the identified policy effects are not significantly different from the true effects. The identified total number of lives saved is $\int_{\mathbb{W}(s;\boldsymbol{\psi}^*)} (\widetilde{y}_{\mathcal{C}}(s,\boldsymbol{\phi}^*) - y_{\mathcal{T}}(s))\, ds = 248.545$ in a window of $s_{\mathcal{C}}(t_p;\boldsymbol{\psi}^*) - t_p = 8.601$ days, whereas the true policy effects are 250.728 lives saved. Therefore, the policy prevented $\gamma = -21.496\%$ of the total deaths that would have occurred had the policy not been implemented, whereas the true effect is $\gamma_{\text{true}} = -21.644\%$. This implies that the margin of error is $\varepsilon(\gamma) = \left|\left(\frac{\gamma}{\gamma_{\text{true}}} - 1\right) \times 100\right| = 0.683\%$.

3.1.2 Oral Contraceptives and Women's Choices

We now pose a model in which the introduction of oral contraceptives (the pill) has effects on women's human capital, sexual and fertility choices. The pill provides access to a technology that

reduces unwanted pregnancies at the time (age) where human capital decisions are taken. We use this framework to assess whether SBI recovers the model-generated effects of the pill.

We assume that each cohort t of women derives utility from their choices on consumption $c \geq 0$, children $n \geq 0$ and sexual intercourse $x \geq 0$ and experiences disutility from pill usage o—e.g. a social norm. In addition, a woman chooses human capital investment paying q (tuition fees or job training) per unit of human capital. Earnings feature two components, a wage level w, and an endogenous human capital wage premium $z_t e(h)$ with the two components technology level z_t and a complementarity factor $e(h) \in [0,1]$[28] with $e_h(h) > 0, e_{hh}(h) < 0$ so that earnings per unit of time are $w(1 + z_t e(h))$. We model skill-biased technical change (SBTC) with a cohort-t specific growth factor γ_t so that $z_t = z_0 \prod_{\tau=1}^{\tau=t}(1+\lambda_\tau)$, where $z_0 > 0$ and $\lambda_t > 0$. We further assume that raising children bears a time cost of $\tau(n) \in [0,1]$ with $\tau_n(n) > 0, \tau_{nn}(n) < 0$ so that earnings are $(1-\tau(n))w(1+z_t e(h))$. Sexual intercourse increases the probability of pregnancy $\phi(x) \in [0,1]$ where $\phi_x(x) > 0, \phi_{xx}(x) < 0$ and we assume that successful pregnancies result in children. If women have access to the pill—which we model through policy dummy $\mathbf{1}_{t_p}$ that is equal to zero if a cohort t does not have access to the pill, and equal to one otherwise—, then the probability of pregnancy is adjusted downward by the pill effectiveness in preventing pregnancy, $g(o) \in [0,1]$. We assume that larger use of the pill—e.g., better adherence to follow protocol—increases the effectiveness of the pill. That is, $g_o(o) > 0$ with $g_{oo}(o) < 0$.[29]

Collecting elements, a woman solves

$$\max_{\{h,o,x\}} c + \kappa n + \zeta x - \iota o \tag{16}$$

subject to the budget constraint (17) and the children production technology (18):

$$c + qh = (1-\tau(n))w(1+z_t e(h)), \tag{17}$$

$$n = \phi(x)[1 - \mathbf{1}_{t_p} g(o)] \tag{18}$$

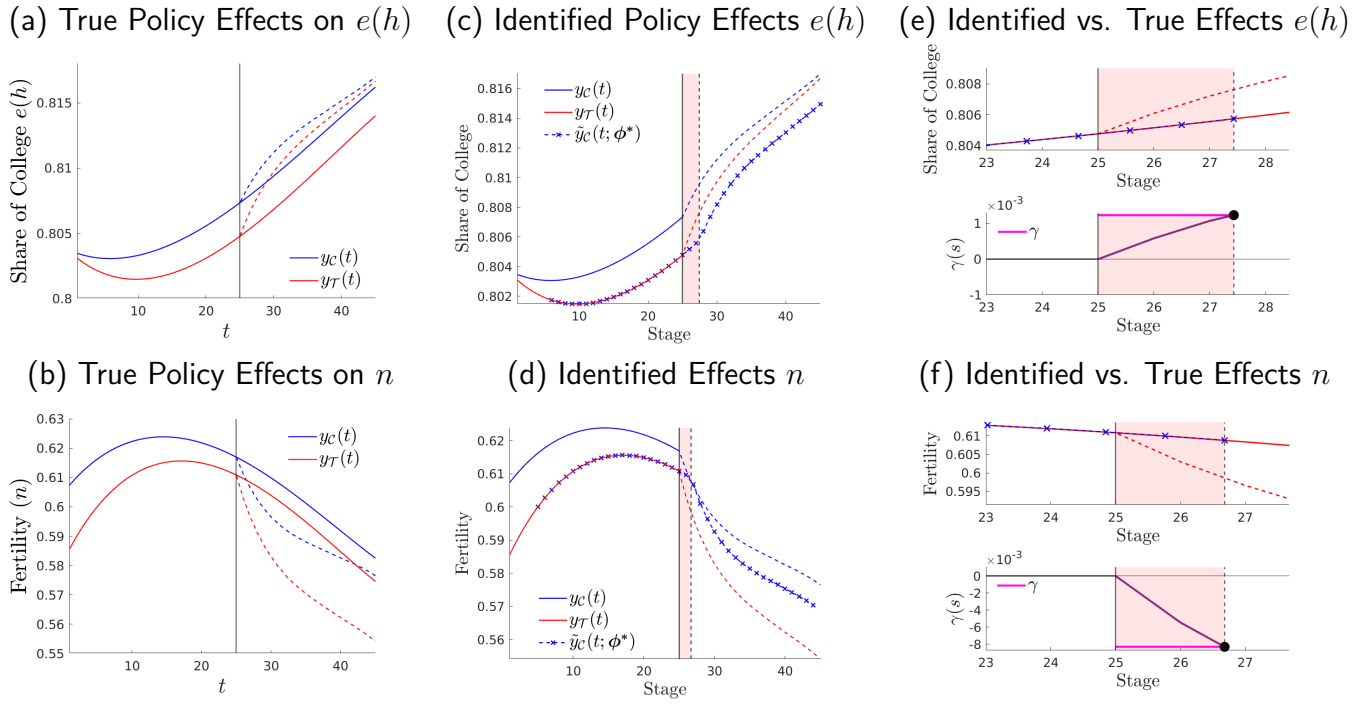
Plugging (17) and (18) into (16), the first order condition (FOC) of h is,

$$FOC(h): \quad \underbrace{q}_{\text{Marginal Cost of Human Capital}} = \underbrace{(1-\tau(n))w z_t e_h(h)}_{\text{Marginal Benefit of Human Capital}}, \tag{19}$$

[28] The mapping of the outcome variable from h to $e(h)$ is innocuous. In particular, since we model $e(h)$ as a rate we can interpret it as the fraction of educated women (e.g. college degree completion) in the population.

[29] Lawful access to the pill does not suffice to determine use which is also likely affected by social norms Goldin and Katz (2002). Further, the pill can—at the same time—shape social norms (Fernández-Villaverde et al., 2014).

Figure 8: Stage-Based Identification of Model-Generated Policy Effects: Introduction of the Pill



Notes: We plot outcomes for region $\mathcal{C}$ (red) and $\mathcal{T}$ (blue) without policy (solid lines) and with policy (dashed lines). The policy is the introduction of the pill for all periods $t_p \geq 25$. The parameter values that we choose for region $\mathcal{C}$ are $\Theta_{\mathcal{C}} = \{\xi = 8, q = 3.2, w = 64, z_0 = 1, \lambda = 0.1\%, \theta_x = 0.5, \theta_h = 0.4, \theta_o = 0.43, \iota_{t,\mathcal{C}}, \kappa_{t,\mathcal{C}}\}$ and for region $\mathcal{T}$ are $\Theta_{\mathcal{T}} = \{\xi = 8, q = 3.3, w = 63, z_0 = 1, \lambda = 0.1\%, \theta_x = 0.5, \theta_h = 0.4, \theta_o = 0.43, \iota_{t,\mathcal{T}}, \kappa_{t,\mathcal{T}}\}$.

where the price of human capital equates the marginal benefit consisting of a wage premium net of the costs of children. Since q is constant, the marginal benefit trades off n and h, i.e. a technology that reduces n enhances human capital. The FOC for sexual intercourse x is:

$$FOC(x): \quad \underbrace{\tau_n(n)\phi_x(x)(1-\mathbf{1}_{t_p}g(o))w(1+z_t e(h))}_{\text{Marginal Cost of Intercourse}} = \underbrace{\zeta + \kappa_t \phi_x(x)(1-\mathbf{1}_{t_p}g(o))}_{\text{Marginal Benefit of Intercourse}}, \quad (20)$$

where the marginal benefit considers the additional utility from sex and children. The marginal cost reflects the cost of children in terms of human capital. The FOC for pill use o is:

$$FOC(o): \quad \underbrace{\tau_n(n)\phi(x)\mathbf{1}_{t_p}g(o)w(1+z_t e(h))}_{\text{Marginal Benefit of Pill}} = \underbrace{\kappa_t \phi(x)\mathbf{1}_{t_p}g_o(o) + \iota}_{\text{Marginal Cost of Pill}} \quad (21)$$

where the marginal cost of the pill is a reduction of utility derived from children and the marginal benefit of the pill is a reduction in the price of human capital.

True (model-generated) policy effects. In Figure 8, we show the equilibrium path for women's schooling choices in panel (a) and fertility choices in panel (b). We show the model-generated paths in a scenario without the pill (solid lines) and in a scenario in which the government grants women legal access to the pill technology (dashed lines). We do this separately for region $\mathcal{C}$ (blue) and region $\mathcal{T}$ (red). Regions differ in the model parameters $\Theta = \{\kappa, \xi, q, w, z, \{\lambda_t\}_{t=1}^{T}, \theta_x, \theta_h, \theta_o\}$. In particular, we allow for the returns to human capital to be larger and grow faster in region $\mathcal{C}$ than in region $\mathcal{T}$ which explains the higher human capital in region $\mathcal{C}$ than in region $\mathcal{T}$. This also explains the lower fertility in region $\mathcal{C}$ than in region $\mathcal{T}$. Further, we exogenously shape the SBTC parameter γ such that the endogenous human capital path is S-shaped for both regions. We also choose an exogenous path for the relative utility derived from children, κ, in order for endogenous fertility to display a boom and bust.

Here, we assess the efffects of legalizing the pill permanently with $\mathbf{1}_{t_p} = 1$ for all cohorts of women $t_p \geq 25$. The policy endogenously reduces births (n) in both regions (dashed lines panel (b), Figure 8).[30] By reducing fertility, the pill reduces the cost of acquiring human capital which increases the share of women entering college ($e(h)$) (dashed lines in panel (a), Figure 8).

Stage-Based Identified policy effects. We apply SBI using region $\mathcal{T}$ as reference, hence, mapping the outcome path of region $\mathcal{C}$ (solid blue) onto that of region $\mathcal{T}$ (solid red) using only pre-policy data as in Section 2. Again, SBI delivers a candidate counterfactual $\widetilde{y}_{\mathcal{C}}(s; \boldsymbol{\phi}^*)$ (blue line with cross markers) for an identification window $\mathbb{W}(s; \boldsymbol{\psi}^*) = [t_p, s_{\mathcal{C}}(t_p; \boldsymbol{\psi}^*)]$ (shaded pink area); see panels (c) and (d) of Figure 8 for human capital and children, respectively. We zoom in the comparison between the identified and the true effects in panel (e) and (f) of Figure 8 for human capital and children, respectively. We find that the SBI policy effects capture well the true effects. The identified effect on human capital is an increase in the proportion of women going to college $e(h)$ by $\gamma = 0.122\%$, whereas the true policy effects are $\gamma_{\text{true}} = 0.123\%$. The identified effect on fertility is a reduction by $\gamma = 0.830\%$, whereas the true effect is $\gamma_{\text{true}} = 0.828\%$. The error $\varepsilon(\gamma)$ of the identified policy effects relative to the true policy effects is 0.182% for human capital and of 0.232% for fertility.

3.1.3 Growth Policy and Structural Transformation

We pose a structural transformation model assuming the presence of inefficient institutions in one economic sector (e.g. agriculture). Then, we assess whether SBI captures the model-generated effects on per capita income of a nationwide policy reform that reduces this inefficiency.

[30] The reduction in fertility follows an increase in the use of the new technology. The pill sustains a higher amount of sex with a lower amount of children and, hence, higher human capital.

There are two sectors in the model denoted by $i \in \{a, m\}$, for agriculture and manufacturing, respectively. A representative firm per sector faces competitive markets. The agricultural firm produces output y_a at relative price p_a (manufacturing is the numeraire good) employing labor n_a at wage rate w_a and land ℓ. We assume inefficient institutions in agriculture captured by a parameter τ that taxes revenue. Agricultural firms thus solve the problem,

$$\max_{n_{at}} \pi_t(\ell) = (1-\tau)p_{at}y_{at} - w_{at}n_{at} \quad \text{s.t} \quad y_{at} = z_{at}n_{at}^{\phi}\ell^{1-\phi},$$

where ϕ is the labor share in agriculture. Since land is fixed, the agricultural technology exhibits decreasing returns to scale.[31] Manufacturing firms produce output y_{mt} with labor n_{mt}—hired at wage w_{mt}—and capital k_t—rented at rate r_t—and solve the problem,

$$\max_{n_{mt},k_{t+1}} \quad y_{mt} - w_{mt}n_{mt} - r_t k_t \quad \text{s.t} \quad y_{mt} = z_{mt}n_{mt}^{\alpha}k_t^{1-\alpha},$$

where α is the labor share in manufacturing. Further, we assume that total factor productivity (TFP) differs by sector according to $z_{it} = z_{i,0}(1+\lambda_i)^t$ for $i = \{a, m\}$ with $\gamma_a < \gamma_m$.

An infinitely-lived representative agent discounts the future at factor $\beta \in (0, 1)$ and chooses sectoral allocations of consumption $\{c_{at}, c_{mt}\}_{t=0}^{\infty}$, labor $\{n_{at}, n_{mt}\}_{t=0}^{\infty}$, and next period capital $\{k_{t+1}\}_{t=0}^{\infty}$. The per period utility function from agricultural goods, $u(c_{at} - \bar{c}_a)$, features a non-homotheticity through a subsistence level, $\bar{c}_a$. Utility from manufacturing goods, $v(c_m)$, is additively separable. Both $u(\cdot)$ and $v(\cdot)$ are strictly concave. The household is endowed with one unit of time in each period, i.e. $n_{at} + n_{mt} = 1 \ \forall t$, that is allocated to either agriculture or manufacturing and receives wage rates $\{w_{at}, w_{mt}\}$. The household receives the rents $\pi(\ell)$ from inelastically supplying (renting) land to agricultural firms. Thus, the household maximizes,

$$\max_{\{c_{at},c_{mt},n_{at},n_{mt},k_{t+1}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \left(u(c_{at} - \bar{c}_a) + \kappa v(c_{mt})\right) \tag{22}$$

where $\kappa > 0$ is a relative utility parameter, subject to the budget constraint

$$p_{at}c_{at} + c_{mt} + k_{t+1} = \sum_{i \in \{a,m\}} w_{it}n_{it} + r_t k_t + (1-\delta)k_t + \pi_t(\ell). \tag{23}$$

[31] The structural change—from a decreasing returns to scale technology (Malthus) to a constant returns to scale (Solow) is studied in Hansen and Prescott (2002) in the context of a one-good economy. Below, we also introduce non-homothetic preferences as an additional mechanism for structural change (e.g. Gollin et al., 2002).

There are three first order conditions for the household problem.[32] First, an intratemporal condition governing the substitution across consumption goods:

$$FOC(c_{at}): \qquad u_{c_{at}}(c_{at})\frac{1}{p_{at}} = \kappa v_{c_{mt}}(c_{mt}) \tag{24}$$

Second, an intertemporal Euler condition for k_{t+1} governing the trade off between one additional unit of consumption today versus tomorrow's consumption,

$$FOC(k_{t+1}): \qquad u_{c_a}(c_{at})\frac{1}{p_{at}} = \beta u_{c_a}(c_{at+1})\frac{1}{p_{at+1}}\left(1 + r_{t+1} - \delta\right), \tag{25}$$

and note that we can rewrite this intertemporal condition in terms of c_m using (24). Third, an intratemporal condition for n_a equates wages across sectors,

$$FOC(n_{at}): \qquad u_{c_a}(c_{at})(w_{at} - w_{mt}) = 0 \tag{26}$$

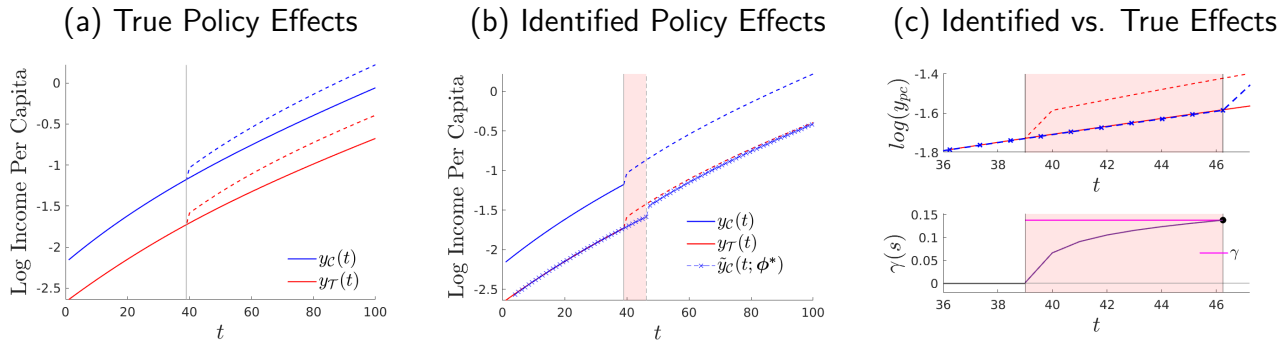
These allocations need to satisfy the marginal product conditions arising from the firms' problems in competitive markets, that is, $w_{at} = \phi\frac{p_{at}y_{at}}{n_{at}}$, $w_{mt} = \alpha\frac{y_{mt}}{n_{mt}}$ and $r_t = (1-\alpha)\frac{y_{mt}}{k_t}$.

True (model-generated) policy effects. We consider two regions that potentially differ in model parameters $\Theta = \{\beta, \bar{c}_a, \kappa, \delta, z_{a,0}, \lambda_a, z_{m,0}, \lambda_m, \phi, \alpha, \tau\}$. In particular, we allow for the total factor productivity in the manufacturing sector to be larger in region $\mathcal{C}$ than in region $\mathcal{T}$. The larger productivity of manufacturing in region $\mathcal{C}$ generates a larger amount of investment, lower agricultural share of labor and, ultimately, higher income per capita in region $\mathcal{C}$ than in region $\mathcal{T}$ at any point in time; see panel (a) in Figure 9.The model is able to generate an agricultural share that declines over time whereas, at the same time, capital and income per capita increase asymptotically reaching a balanced growth path with a trifling agricultural share. In this context, we introduce an unexpected nationwide growth policy that removes the institutional constraint τ in the agricultural sector in both regions; setting $\tau = 0$ after t_p in both regions. Removing the constraint in the agricultural sector accelerates investment (and capital) and the decline in agricultural sector. The reallocation to the non-agricultural sector increases income per capita in the eocnomy, see (dashed lines) in panel (a) of Figure 9.

Stage-Based Identified policy effects. The policy evaluator is not provided with the true counterfactual path without policy (solid lines for the periods after t_p). Under these same data constraints, we implement SBI mapping the outcome path in region $\mathcal{C}$ (solid blue line) onto

[32] Note that we can isolate c_m from (23) and plug it into (22) plus use $n_{mt} = 1 - n_{at}$. This implies that we can maximize the objective function in terms of the sequences of three unknowns $\{c_{at}, n_{at}, k_{t+1}\}_{t=0}^{\infty}$.

Figure 9: Stage-Based Identification of Model-Generated Policy Effects: Growth Policy



Notes: For region $\mathcal{T}$, we choose, $n_{a,0} = 0.45, z_{a,0} = 0.15, z_{m,0} = 0.17, \gamma_a = 0.007, \gamma_m = 0.0073$. For region $\mathcal{C}$, we choose, $n_{a,0} = 0.65, z_{a,0} = 0.145, z_{m,0} = 0.145, \lambda_a = 0.007, \lambda_m = 0.0072$. Common parameters between both regions are $\beta = 0.98, \alpha = 0.6, \phi = 0.8, \kappa = 2, \delta = 0.02$. Further, we assume that the felicity functions are logs, that is, $u(c_a - \bar{c}_a) = \ln(c_a - \bar{c}_a)$ and $v(c_m) = \ln c_m$.

the outcome path in region $\mathcal{T}$ (solid red line) using only pre-policy data. We plot the resulting counterfactual candidate $\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}^*)$ (blue line with cross markers) for the identification window between t_p and $s_C(t_p;\boldsymbol{\psi}^*)$ (shaded pink area); see panel (b), Figure 9. We zoom in on the identified counterfactual $\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}^*)$ and the true effects of policy in panel (c) of Figure 9. According to SBI, the growth policy increases income per capita by $\gamma = 13.781\%$ in the identification window whereas the true policy effect is $\gamma_{\text{true}} = 13.537\%$. That is, the identified policy effects catch the true policy effects with an error of $\varepsilon(\gamma) = 1.797\%$.

3.2 Bounds to Method Performance

The performance analysis in Section 3.1 shows that our identification strategy can recover the true policy effects. However, it is intuitive to assume that our strategy faces some boundaries. Here, we numerically characterize the bounds within which our method is able to recover the true effects of policy with a Monte Carlo experiment in Section 3.2.1. We further assess how our method fares in the presence of time-varying latent heterogeneity; confounding policy interventions; and endogenous policy in Section 3.2.2.

3.2.1 A Monte Carlo Analysis

We focus this analysis on the benchmark economic model with an endogenous pandemic described in Section 3.1.1. Specifically, we hold fixed the parameters of the non-reference region $\mathcal{C}$ and randomize a subset—$(\beta, \zeta, \kappa, t_o)$—of the structural parameters in that region in order to generate a large number of reference outcome paths $y_{\mathcal{T}}(m)$ for regions $m \in \mathcal{M} = \{1, ..., m, ..., M\}$.[33]

[33]We assume that the randomized parameters—β, ζ, κ and t_o—are uniformly and independently distributed. Then, we draw a total of $M = 381,000$ simulations (quadruplets).

In panel (a) of Figure 10, we show the epidemic path of our benchmark regions $\mathcal{C}$ and $\mathcal{T}$ as described in Section 3.1.1, together with one of the simulated reference regions that starts later, grows slower and reaches a lower magnitude than the benchmark reference region, $y_{\mathcal{T}}(t)$, and, therefore, is further away from the non-reference region, $y_{\mathcal{C}}(t)$.[34]

In this context, in order to assess the ability of SBI to identify the true policy effect we study the policy error across all simulations. For each simulation m, we apply SBI mapping the non-reference region, $y_{\mathcal{C}}(t)$, onto the simulated reference path, $y_{\mathcal{T}(m)}(t)$. This implies that we find a set of normalization coefficients $\boldsymbol{\phi}^*(m) = \{\boldsymbol{\psi}_0^*(m), \boldsymbol{\psi}_1^*(m), \boldsymbol{\omega}_1^*(m)\}$ per simulation $m \in \mathcal{M}$. Then, for each simulation, we measure the policy error as the (absolute) value of the policy effect identified by SBI relative to the (model-generated) policy effect; i.e. $\varepsilon(\gamma)(m) = \left|\left(\frac{\gamma(m)}{\gamma_{\text{true}}(m)} - 1\right) \times 100\right|$. In panel (b) of Figure 10, we plot the policy errors of each of our simulations $\boldsymbol{\phi}^*(m)$ that belong to the vector space $\Phi^q = \Psi_0^q \times \Psi_1^q \times \Omega_1^q = \{\psi_0(m) > 0.0\} \times \{\psi_1(m) > 1.0\} \times \{\omega_1(m) < 1.0\} \subset \Phi = \mathbb{R}^3$. We restrict the plot to the vector space $\{\psi_0(m) \in (0.000, 10.000)\} \times \{\psi_1(m) \in (1.000, 1.500)\} \times \{\omega_1(m) \in (0.350, 1.000)\} \subset \Phi^q$, which suffices to capture the policy error associated with the benchmark reference region $y_{\mathcal{T}}(t)$.[35]

Our main result is that the success of our method in identifying the true policy effects is bounded. To see this, first, note that the centroid in the vector space Φ, i.e. $\boldsymbol{\phi}_{\mathbf{c}}^* = (0.000, 1.000, 1.000)$, implies that the outcome path of the simulated reference region, $y_{\mathcal{T}(m)}(t)$, and that of the non-reference region, $y_{\mathcal{C}}(t)$, are identical.[36] Second, note that if the outcome path of a simulated reference region, $y_{\mathcal{T}(m)}(t)$, and the outcome path of the non-reference region, $y_{\mathcal{C}}(t)$, are similar—in that our identification strategy delivers a set of normalization coefficients that is in a neighborhood of the centroid $\mathcal{N}(\boldsymbol{\phi}_{\mathbf{c}}^*) \subset \Phi^q$—then the policy error is small; see panel (b) of Figure 10. To see this, note that policy errors with values of $\varepsilon(\gamma) \leq 5\%$ emerge in a bounded neighborhood (approximately) $\mathcal{N}(\boldsymbol{\phi}_{\mathbf{c}}^*) = \{0.000, 6.776\} \times \{1.000, 1.210\} \times \{0.764, 1.000\}$ which we depict (yellow area) around the centroid. Here, note that our benchmark reference outcome path $y_{\mathcal{T}}(t)$ falls in that neighborbood with a set of normalization coefficients $\boldsymbol{\phi}^* = \{6.592, 1.041, 0.803\}$ and a policy error $\varepsilon(\gamma) = 0.68\%$ (red marker). Third, moving away from the centroid increases the policy error. For example, the simulated reference outcome path $y_{\mathcal{T}(m)}(t)$ in panel (a) of Figure 10 implies a set of normalization coefficients $\boldsymbol{\phi}^*(m) = \{5.083, 1.200, 0.436\}$ that falls outside of the neighborhood $\mathcal{N}(\boldsymbol{\phi}_{\mathbf{c}}^*)$ and delivers a larger policy error of 36.04%. We further reconduct our exercise using an alternative measure of the policy error defined as $|\gamma - \gamma_{\text{true}}|$

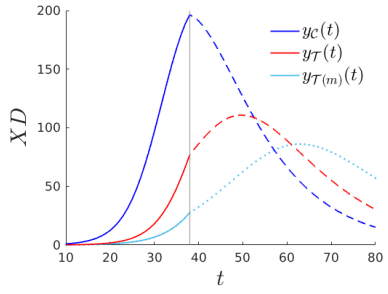
[34] Note that there can be cases where the simulated reference region flips control versus treatment assignment.

[35] Our insights do not change with alternative choices of the vector space.

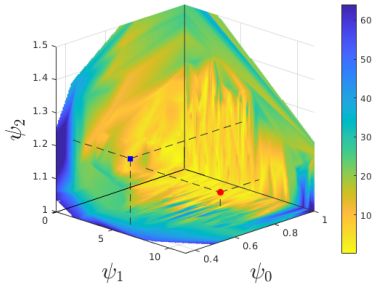
[36] Indeed, exactly at the centroid the policy effects are not identified because $y_{\mathcal{C}}(t) = y_{\mathcal{T}(m)}(t)$ and there is no heterogeneity in stages at the time of policy implementation.

Figure 10: Bounds to Method Performance: A Monte Carlo Analysis

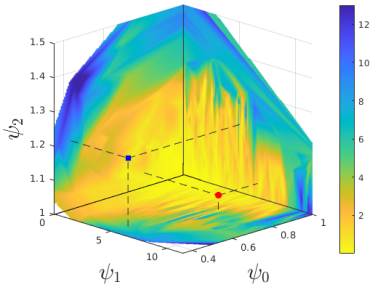
(a) A Simulation Example



(b) Policy Error: $|\epsilon(\gamma)|$



(c) Policy Error: $|\gamma_{true} - \gamma|$



Notes: When constructing the set $\{y_{\mathcal{T}(m)}(t)\}_m$, we assume that $\{\beta, \zeta, \kappa, t_0\}$ are uniformily and independnetly distributed. The simulations are drawn from the intervals $\left[\beta^{lb}, \beta^{ub}\right] \times \left[\zeta^{lb}, \zeta^{ub}\right] \times \left[\kappa^{lb}, \kappa^{ub}\right] \times \left[t_o^{lb}, t_o^{ub}\right] = [0.5, 0.9] \times [0.001, 0.008] \times [1.05, 1.89] \times [-10, 10]$ where the superindices lb and ub denote, respectively, the lower and upper bounds of each parameter space. We pick the bounds of the uniform distribution in a manner that our simulations generate sufficiently different outcome paths of the reference region in order to assess the performance of our method. We constructed a total of $M =$ 381,000 simulations though not all the simulations fall in the vector space $\boldsymbol{\psi}^*$ in panel (b). Precisely, the hyperplane (ψ_0, ω_1) has 3,698 simulations, the hyperplane (ψ_1, ω_1) has 17,504 simulations and the hyperplane (ψ_0, ψ_1) has 3,698 simulations. Panels (b) and (c) show values from an evenly spaced 200×200 grid on each hyperplane. We approximate the values on the grid through linear interpolation of the simulated data.

in panel (c) of Figure 10 reaching similar insights. Thus, as long as the regional outcome paths are similar enough, the method can successfully identify the policy effect.

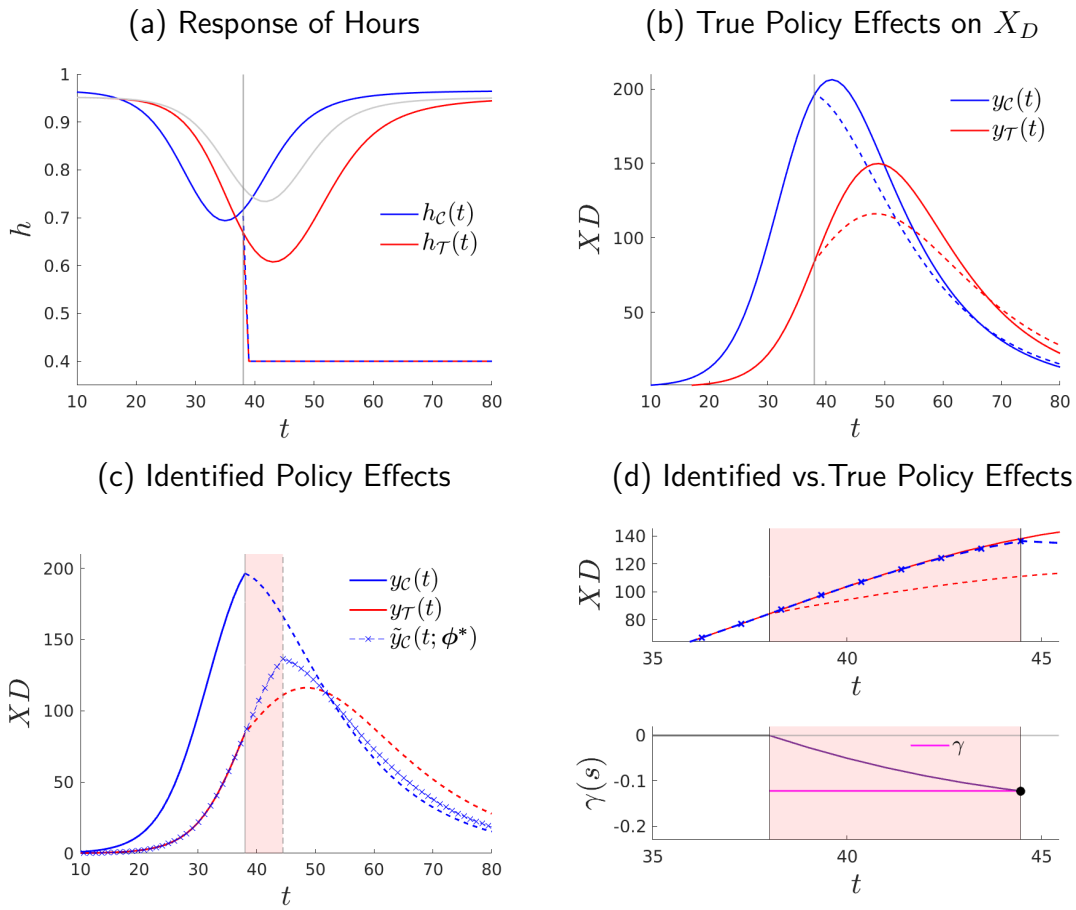
3.2.2 Time-Varying Latent Heterogeneity, Confounding Policy and Endogeneity

We assess how the presence of time-varying latent heterogeneity, confounding policy and endogenous policy affects the ability of SBI to recover the true effects of policy. Here, we would like to emphasize that the goal of our method is not to answer what would have been the effect of the policy under evaluation had the time-varying latent heterogeneity, confounding policy or endogenous policy (or other confounding factors for that matter) not been present. That is, we do not pursue the identification of a "pure" policy effect that nets out (controls for) the presence of confounding factors.[37]

Instead, we acknowledge that the effect of the same policy can naturally be different when it is applied in a different context—e.g., due to the presence of different confounding factors. SBI

[37] Therefore, from the perspective of SBI, the presence of confounding factors does not change neither the aim of the normalization step, which is to reduce the cross-regional differences in the pre-policy determinants—including potentially unobserved structural parameters—of the path of the outcome of interest, nor the identification step that assumes the normalization parameters are unaffected by the policy that is evaluated.

Figure 11: Stage-Based Identification of Policy Effects: Time-Varying Latent Heterogeneity



Notes: Where $\bar{h} = 0.4, t_p = 38, t_f = 250$, $\gamma = -12.242\%, \epsilon(\gamma) = -3.318\%$.

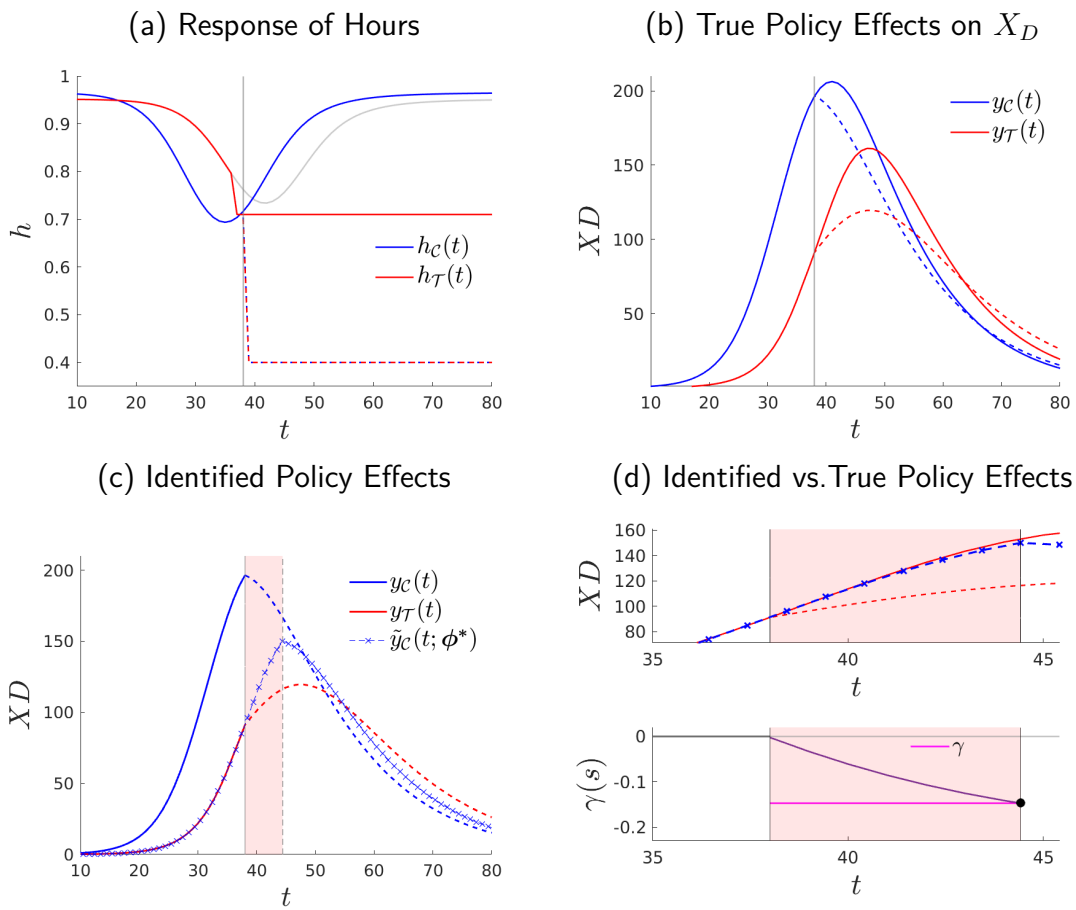
is designed to measure these conditional policy effects. Thus, we are interested in assessing how well SBI can recover the true policy effects that emerge given different confounding factors. To do so, we use as benchmark the econ-epi model described in Section 3.1.1 in which there were no confounding factors. Then, we introduce confounding factors (one by one) into the model in order to show how the true policy effects, which now explicitly depend on the specific set of confounding factors that are present, change and assess whether SBI can recover these true policy effects.

Time-Varying Latent Heterogeneity. Using the econ-epi model in Section 3.1.1, we formalize the time-varying latent heterogeneity across two regions as an underlying time-varying structural parameter present in one region but not the other. In particular, we consider a scenario in which one region, $\mathcal{T}$, learns about the process of infection before policy implementation. That is, we

allow for the beliefs on the infection process, $\xi_{\mathcal{P}}$, to exogenously and gradually move closer to the actual ξ in region $\mathcal{T}$ but not in region $\mathcal{C}$; see Figure 26 in the Appendix D. This path of beliefs induces pre-policy behavioral change relative to the scenario with a fixed $\xi_{\mathcal{P}}$. In region $\mathcal{T}$, there is now a larger behavioral response (reduction of hours) to the pandemic, see panel (a) in Figure 11, which also plots the path under fixed beliefs in gray. Note that in our illustration the pre-policy behavioral change is rather large, in the sense that now the drop in hours worked before policy implementation in region $\mathcal{T}$ becomes larger than that in region $\mathcal{C}$. We show the implied true policy effects on the flow of deaths in panel (b), the identified policy effects in panel (c) and a comparison between true and identified effects in panel (d) of Figure 11. The main finding is that our method can recover the true policy effects under time-varying unobserved heterogeneity. The estimated percentage of lives saved is $\gamma = 12.242\%$ which is close to the true effects, $\gamma_{\text{true}} = 12.663\%$. Hence, SBI can recover the policy effects in contexts where there is time-varying latent heterogeneity. However, analogously to our discussion in Section 3.2.1, the robustness of our method to time-varying heterogeneity is bounded by how far the time-varying component drives the outcome paths across regions away from each other.

Confounding Policies We now consider a scenario in which an additional confounding policy is introduced in region $\mathcal{T}$ right before the actual nationwide stay-home policy under evaluation is implemented in period t_p. In particular, we assume the existence of an unanticipated policy that imposes an additional (and weaker) constraint on hours worked, $\bar{h} = 0.71$, one period before the nationwide stay-home policy is put in place, see panel (a) of Figure 12. Then, we apply SBI to assess the nationwide stay-home policy introduced at t_p, while purposefully ignoring the presence of the additional policy introduced before t_p in region $\mathcal{T}$. Our method identifies the effects of policy to be a 14.708% of lives saved in the identification window which is close to the true effects in that window, 15.588%. This implies a policy error of 5.644%. We further re-conduct the exercise imposing less strict confounding policies in region $\mathcal{T}$ one period before t_p with similar insights. That is, SBI can recover the true policy effects in a context where there exists a confounding policy. At the same time, analogously to our discussion in Section 3.2.1, it is straightforward to show that the error by which SBI captures the true policy effect increases if the effect of the additional confounding policy makes the cross-regional outcome paths sufficiently dissimilar before t_p. That is, how much our identification strategy recovers of the true policy effects in a context where confounding policy is present is bounded by the strength of the confounding policy in making the cross-regional outcome paths differ from each other.

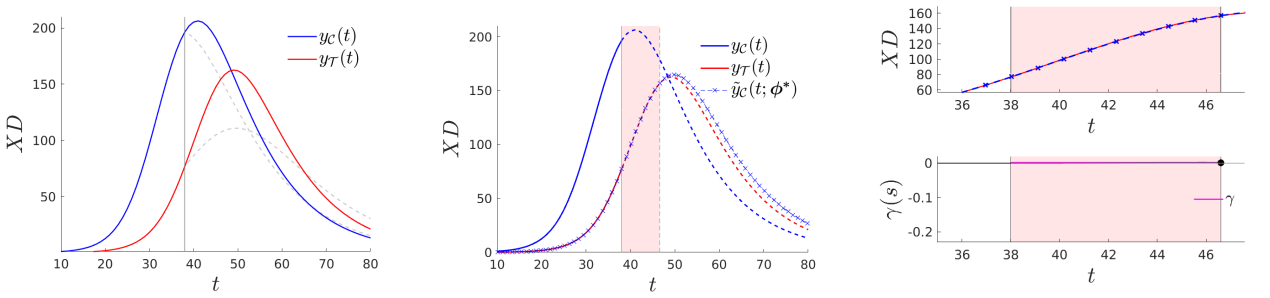
Figure 12: Stage-Based Identification of Policy Effects: With Confounding Policy in $\mathcal{T}$



Notes: Where $\bar{h}_1 = 0.71$ at $t = 37$ in $\mathcal{T}$, this policy is unobserved, $\bar{h}_2 = 0.4$ (lockdown) at $t_p = 38, t_f = 250$, $\gamma = -14.708\%, \epsilon(\gamma) = -5.644\%$.

Figure 13: Stage-Based Identification of Policy Effects: A Placebo Test

(a) True Policy Effects (b) Identified Policy Effects (b) Identified vs.True Effects



Notes: Where $\bar{h} = 0.4, t_p = 38, t_f = 250$, $\gamma = 0.188\%$.

3.3 Inference

We conduct inference in two ways. First, we conduct a placebo diagnosis in order to assess how our method evaluates inexistent policy effects. Second, we assess our method when the outcome path of interest is subject to a stochastic component.

3.3.1 Placebo Diagnosis

Here, we assess whether SBI identifies policy effects when the policy effects are non-existent. In such scenario, a successful diagnosis is one in which our method identifies the effects of policy to be nil, as they truly are. To conduct this assessment, we apply our method to model-generated data from models that are not subject to policy. We use as benchmark the econ-epi model from Section 3.1.1 with the relevant difference that we do not impose a policy at time t_p. Under such scenario, the paths for the flow of deaths in region $\mathcal{C}$ (solid blue) and region $\mathcal{T}$ (solid red) are as depicted in panel (a) of Figure 13. For reference, we also plot the path for deaths that would have occurred (dashed light gray) had the policy been implemented at t_p as we did in Section 3.1.1.

We now apply SBI as if there was a policy at some period t_p—when there is actually none. Given that the normalization uses only pre-policy data, we obtain the same identification window over stages as if there was an actual policy. We show the outcome paths for the two regions, $y_{\mathcal{T}}(t)$ and $y_{\mathcal{C}}(t)$, along with the obtained normalized path $\widetilde{y}_{\mathcal{C}}(s, \boldsymbol{\phi}^*)$ in panel (b) of Figure 13; panel (c) zooms in on the identification window. Note that the normalized outcome path $\widetilde{y}_{\mathcal{C}}(s, \boldsymbol{\phi}^*)$ is practically identical to the outcome path $y_{\mathcal{T}}(t)$ on the identification interval: the identified counterfactual matches the actual outcome path—which here is also the outcome path without policy. That is, SBI correctly identifies that in this scenario without policy the policy effects

are non-existent—or quantitatively negligible, $\gamma = 0.188\%$. We find similar insights after re-conducting this exercise for different values of t_p.

3.3.2 Stochastic Component

In empirical applications, the outcome path of interest often is subject to fluctuations due to the presence of a stochastic component, which can capture measurement error.[38] When facing such noisy data, we add a smoothing—or trend-extraction—step that precedes the normalization step of the SBI method. The goal of this smoothing step is to purge the observed pre-policy outcome paths of the stochastic fluctuations—of higher frequency than the object of interest—defined as deviations from some estimand. We then apply the normalization step of the SBI method on the smoothed pre-policy data. Furthermore, given that the available data does not exactly capture the true path, it is important to conduct statistical inference, for which we propose a bootstrap procedure that builds on the stochastic component around an extracted fitted value.

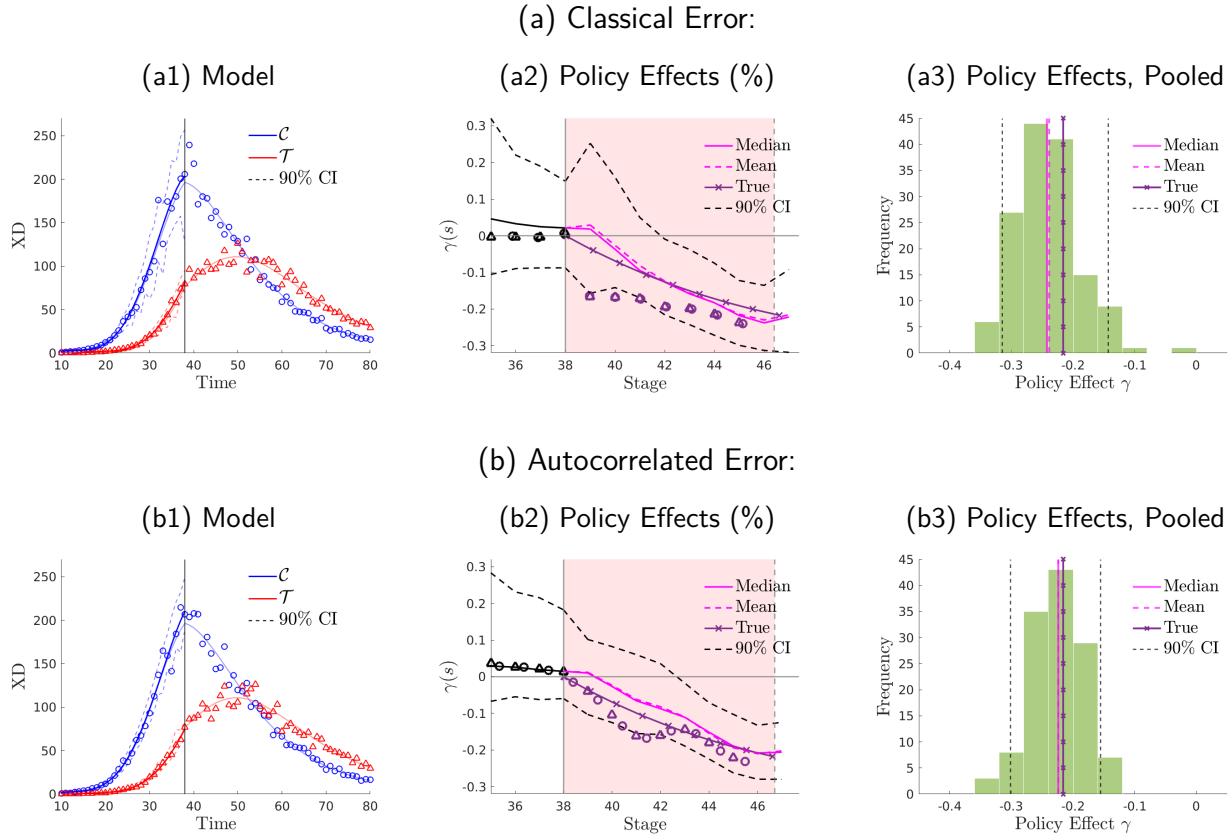
Consider first the presence of classical measurement error, and let the outcome paths be

$$\widehat{y}_r(t) = y_r(t) + u_r(t) \quad \text{with} \quad u_r(t) \sim N(0, \sigma^2_{u,r}), \tag{27}$$

for each region $r = \{\mathcal{C}, \mathcal{T}\}$, where $\widehat{y}_r(t)$ is the outcome path observable to the policy evaluator, $y_r(t)$ is the unobservable true outcome path and the innovations $u_r(t)$ capture measurement error that follows a Normal distribution with zero mean and variance σ^2_r. In the case where the stochastic component is autocorrelated, we replace the innovations in specification (27) with $u_r(t) = \rho u_r(t-1) + \upsilon_r(t)$ where $\upsilon_r(t) \sim N(0, \sigma^2_{\upsilon,r})$. In panel (a.1) of Figure 14, we show the (unobserved) true outcome paths $y_r(t)$ for the two regions as light blue and light red lines, together with the observed outcome paths $\widehat{y}_r(t)$ for one simulation of the errors in (27), which are indicated by the circle and triangle markers. Further, in what follows we denote by $\widehat{\widehat{y}}_r(t < t_p)$ the estimand of $y_r(t < t_p)$. To obtain this estimand, we fit Chebyshev polynomials to the observed data $\widehat{y}_r(t < t_p)$ and, hence, also recover the time-series of regional errors $u_r(t < t_p)$ as the deviations from the fitted values. Then, for each region, we construct $B = 1,000$ bootstrap draws, $\widehat{y}_{r,b}(t)$ with $b \in B$. For each bootstrap draw, we randomly draw a sequence of errors from the region-specific set of errors with replacement, which we add to the fitted values of the pre-policy path, $\widehat{\widehat{y}}_r(t < t_p)$. In panel (a1) of Figure 14, we show the median (solid lines) and 90% confidence intervals (dashed lines) of the bootstrap pre-policy paths $\widehat{y}_{r,b}(t < t_p)$.

[38] Alternatively, the stochastic fluctuations might be genuine but yet of higher frequency than the outcome path of interest—e.g., autocorrolated business cycles when the object of interest is the growth path. Also in this scenario it is important to extract the lower-frequency component, in order to then use SBI on it.

Figure 14: Stage-Based Identification of Model-Generated Policy Effects: Inference



Notes: We use the benchmark parameterization of Section 3.1.1. The top panels (a) add classical error in our model with $\{\sigma^2_{\mathcal{C}}, \sigma^2_{\mathcal{T}}\} = \{0.008, 0.008\}$. The bottom panels (b) add autocorrelated error with $\{\rho_{\mathcal{C}}, \rho_{\mathcal{T}}\} = \{0.13, 0.13\}$ and $\{\sigma^2_{\mathcal{C}}, \sigma^2_{\mathcal{T}}\} = \{0.008, 0.008\}$.

Now, we apply SBI to each bootstrap sample using the recovered estimands $\widehat{\widehat{y}}_{r,b}(t < t_p)$ to perform the normalization, i.e., the mapping of the non-referece region $\mathcal{C}$ onto the reference region $\mathcal{T}$. Then, on the obtained identification window, we use the data of region $\mathcal{T}$ and the normalized path for region $\mathcal{C}$ to measure the effect of policy effect for each bootstrap draw, γ_b. Importantly, the heterogeneity in γ_b across bootstrap draws arises from both, differences in the policy effect per stage during the identification window, and differences in the size of the identification window itself, stemming from the bootstrap-draw-specific stage of policy implementation in the non-reference region, $s_{\mathcal{C}}(t_p; \boldsymbol{\psi}^*_b)$. We thus split the reporting into two steps. First, we report confidence bands over the window of the observed original data, for which we focus on those bootstrap simulations that deliver roughly the same window length. Second, we report overall confidence bands of the policy effect, based on all bootstrap draws. In panel (a2) of Figure 14 we plot again the true cumulative policy effects $\gamma_b(s)$ from Figure 7. We then add to the figure the mean, median, and the 5th and 95th percentiles of the bootstrap draws that give

a window of about the same length—we select those bootstrap draws that fall into plus/minus 10% of the length of the average bootstrap draw. Two results emerge. First, the normalization generates outcome paths that are not significantly different before policy implementation; see the non-shaded area in panel (a2). Second, the identified policy effects—using data with measurement error—are not significantly different from the true (model-generated) policy effects without measurement error (purple line with crossed markers): the identified mean policy effect (dashed magenta line) is 22.58%—within a 90 percent confidence interval of [11.71,31.58]—which is not significantly different from the true (model-generated) policy effect without measurement error, i.e., 21.50%. The median policy effect (solid magenta line) is of similar size, 23.1%. However, unsurprisingly, the significance of the identified policy effect can be affected by the size of the measurement error. Further, we find that the identified mean policy effect is similar (23.89%) when we do not restrict our analysis to bootstrap draws of the same window size; see panel (a3) of Figure 14, which shows the distribution of the overall policy effects across all bootstrap draws. Naturally, the distribution is somewhat wider than within the restricted bootstrap subsample—however, the 90% confidence band is only mildly larger. We also conduct robustness of our methodology using a wider set of smoothers and find similar insights—our alternative smoothers include B-splines, cubic splines, moving averages and the Hodrick-Prescott filter.[39]

We re-conduct our analysis assuming that the stochastic component is autocorrelated, see panel (b1) of Figure 14. In this case, to keep the empirical autocorrelation structure of the error terms—including potentially temporal differences in the cross-sectional variance—we use a block bootstrap procedure that increses the sampling weight of preceding error terms in a pre-specified window (Carlstein, 1986).[40] With autocorrelated measurement error, we also find that our identification strategy is able to recover policy effects that are not significantly different from the true (model-generated) policy effects, see panel (b2) and (b3) of Figure 14.

Finally, in order to assess the role of the smoother on the identification of the true policy effects, we perform SBI directly on the observed data $\widehat{y}_r(t < t_p)$—i.e., the markers in panel (a1) of Figure 14. That is, we conduct the normalization by mapping directly the outcome path $\widehat{y}_{\mathcal{C}}(t)$ onto $\widehat{y}_{\mathcal{T}}(t)$ without the smoothing step.[41] There is a unique identified policy effect (purple markers) that we show in panel (a2) of Figure 14. The policy effect identified using the observed

[39] An altogether alternative way to conduct inference with the recovered estimates for the error terms $u_r(t)$ is to estimate the sample variance of the errors, i.e. $\widehat{\sigma}_r$. Then, under a normality assumption on the error term in (27), we simulate $Q = 1,000$ paths of errors (instead of drawing from the empirical distribution). The results under this different inference are in Appendix E. Overall, we find similar insights with an identified mean policy effect of 21.12% [14.41,28.23] that is not significantly different from the true (model-generated) policy effect.

[40] We select a block window of size 5

[41] To measure the distance between the normalized data of the non-reference region and the actual data of the reference region, we apply linear interpolation.

data $\widehat{y}_r(t)$ also replicates the true policy effect, which suggests that the smoothing step (which is necessary to conduct inference in the way described here) does not substantially affect the identified policy effect itself.

4 Applications

We use SBI to identify the policy effects in a set of empirical applications associated with nationwide policies. First, we assess the effects of stay-home policies on the flow of Covid-19 deaths in Spain in Section 4.1.[42] Second, we assess the effects of the approval of oral contraceptivies on fertility rates and women's college education in the United States in Section 4.2. Third, we study the effects of the German reunification on income per capita Section 4.3.

4.1 The Spanish *Confinamiento* Against Covid-19

In response to the Covid-19 pandemic, on March 14, 2020, the Spanish government announced a nationwide stay-at-home policy—enacted the following day—which locked down all non-essential workers in all regions of Spain. Indicative of its strictness, the public debate referred to the policy as confinement. The strictest measures were lifted on May 2 when the first wave of the epidemic flattened out. Here, we apply SBI to assess the effects of this policy intervention on the course of the pandemic. As outcome of interest, we focus on the daily flow of deaths attributed to Covid-19.[43] We use two Spanish regions to assess the nationwide policy: Madrid and an artificially created region Rest of Spain (RoSPA) which is composed of all Spanish regions without Madrid. We label Madrid as region $\mathcal{C}$ and RoSPA as region $\mathcal{T}$.[44]

We show the daily flow of Covid-19 deaths (per million inhabitants) for Madrid (blue circles) and the RoSPA (red triangles) in panel (a) of Figure 15; Instituto de Salud Carlos III. In order to mitigate potential measurement error on the reported deaths, we smooth the pre-policy data using as benchmark Chebyshev polynomials separately by region as described in Section 3.3.[45] Note that we add a lag parameter to the policy date, reflecting that a policy that aims at reducing

[42] The Covid-19 has generated lots of empirical work assessing public health policies against the pandemic; see, for example, Fang et al. (2020) for a careful study of the early mobility restrictions in China and Liu et al. (2021) for the provision of density forecasts with Bayesian techniques for a panel of countries and regions.

[43] Although daily deaths are potentially imperfectly measured, we regard these data as less prone to measurement error than infections data, especially during the onset of the pandemic, when testing was largely unavailable.

[44] Note that SBI can be conducted for all pairs of regions; see Section 5.1 for an analysis of the stay-home policy using the power set of all Spanish regions. However, in order to ease the exposition, we focus here on two regions (or groups of regions). To select these regional groups, we conduct the normalization in Section 2.1 by mapping the path of Covid-19 deaths of each Spanish region onto the aggregate path for Spain. This normalization uncovers that, at the time of policy implementation, Madrid is at the most advanced stage, i.e. Madrid $= \arg\max_r s_r(t_p; \boldsymbol{\psi}^*)$. For this reason, we focus on Madrid and RoSPA for our analysis.

[45] We use a Chebyshev polynomial of degree 6 and perform robustness on the choice of the smoother.

infections will show an effect on the flow of deaths with a delay. We choose a lag of 12 days, which implies that the policy is (effectively) implemented on March 27; The resulting smoothed daily flow of deaths for Madrid (solid blue) and RoSPA (solid red) are also in panel (a) of Figure 15.

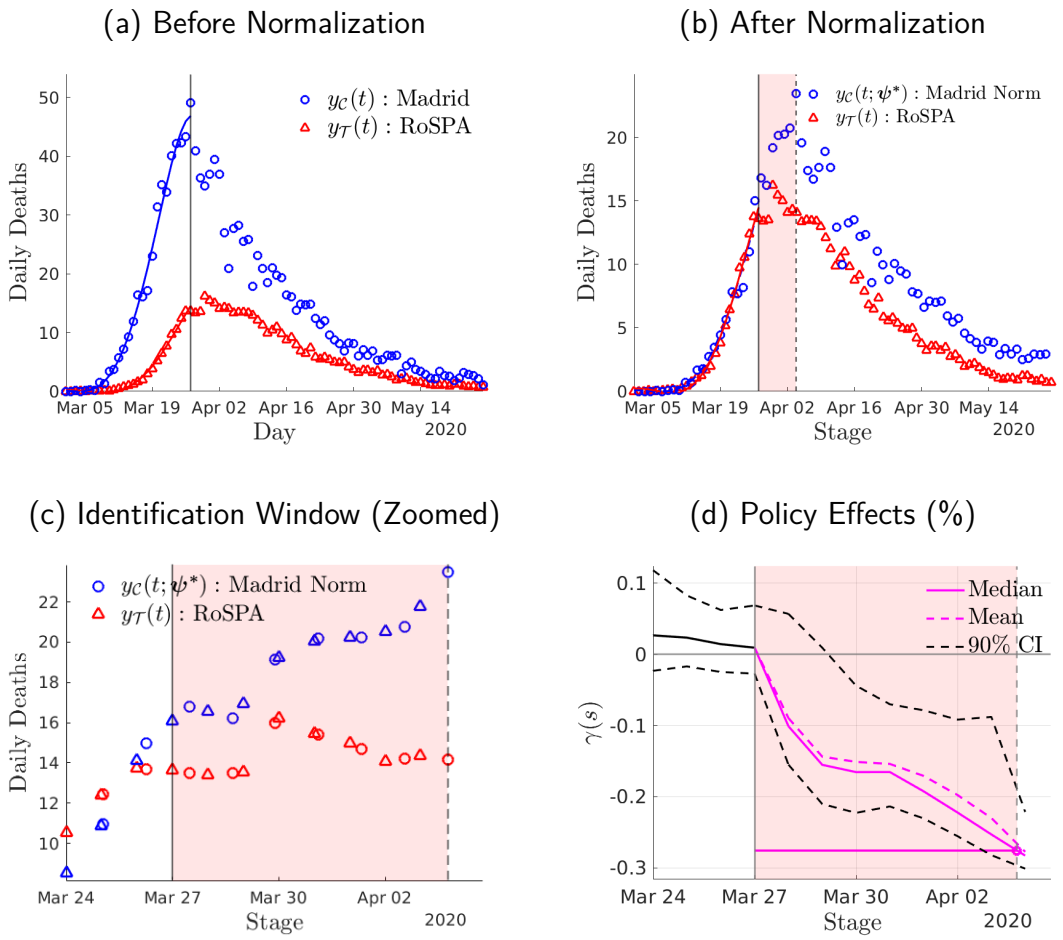
There are clear differences in the path of the flow of deaths between Madrid and the RoSPA. First, one death (per million inhabitants) is reached in March 08 for Madrid and March 14 for the RoSPA. Second, by March 14 the daily flow of deaths in Madrid is 9.3 deaths (per million inhabitants) whereas this figure is 1.2 for the RoSPA. Furthermore, at the (effective) time of policy implementation, the flow of deaths is reaching a peak in Madrid at 50 deaths (per million inhabitants), whereas the peak in the RoSPA is smaller at 16 deaths (per million inhabitants) and occurs about a week after that in Madrid. That is, the flow of deaths starts at an earlier date, it raises more rapidly and reaches a larger peak in Madrid than in the RoSPA.

Normalization. We now apply SBI following the normalization described in Section 2. Picking the RoSPA as reference region $\mathcal{T}$, we map the flow of deaths of the region Madrid ($y_{\mathcal{C}}(t)$, solid blue circles) onto the flow of deaths of RoSPA ($y_{\mathcal{T}}(t)$, solid red circles) using *only* pre-policy data. The normalization step delivers a normalized path for Madrid $\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}^*)$ that is not different—up to a minimization error—from that of RoSPA, $y_{\mathcal{T}}(t)$; see panel (b) of Figure 15. We find ψ_0 =-0.14 [-0.24,-0.04], ψ_1 =1.21 [1.16,1.24] and ω_1 =0.47 [0.39,0.53] which, respectively, delays the start, slows down the growth and lowers the peak of daily deaths in Madrid.A result of our normalization is that Madrid leads the epidemic in Spain. Precisely, the policy is implemented in Madrid at a later stage than in RoSPA, i.e. $s_{\mathcal{T}}(t_p;\boldsymbol{\psi}^*) = t_p < s_{\mathcal{C}}(t_p;\boldsymbol{\psi}^*)$. Hence, the normalization unveils a window in stages $\mathbb{W}(s;\boldsymbol{\psi}^*) = [t_p, s_{\mathcal{C}}(t_p;\boldsymbol{\psi}^*)]$ (shaded pink area) running from March 27 to April 03 in which the stage-leading region, Madrid, is not yet subject to policy whereas RoSPA is.[46] Therefore, under our identification assumption, the normalized path of the Madrid serves as no-policy counterfactual for RoSPA inside $\mathbb{W}(s;\boldsymbol{\psi}^*)$.

Policy Effect. We zoom in on the outcome paths $y_{\mathcal{C}}(s;\boldsymbol{\phi}^*)$ (normalized Madrid) and $y_{\mathcal{T}}(t)$ (RoSPA) inside the identification window in panel (c) of Figure 15. The implied policy effects are in panel (d), where we restrict the attention to the bootstrap simulations within the neighborhood of the median window size (plus/minus 10%). Across these bootstrap draws the (mean) identified total number of lives saved (per million inhabitants) is $\int_{\mathbb{W}(s;\boldsymbol{\psi}^*)} (\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}^*) - y_{\mathcal{T}}(s))\, ds$ = 36.92 within approximately one week after policy implementation, which corresponds to a total amount of lives saved by the policy in RoSPA of 1,734 during that week. That is, the stay-home policy prevented $\gamma = -24.71\%$ of the total deaths that would have occurred in the RoSPA had the

[46] Precisely, the window $\mathbb{W}(s;\boldsymbol{\psi}^*)$ runs from the effective policy date in RoSPA (t_p = March 27) to the effective policy date in the stage domain for Madrid, $s_{\mathcal{C}}(t_p;\boldsymbol{\psi}^*) = t_p + 7.7$ days (i.e., during April 03).

Figure 15: The Effects of the Spanish *Confinamiento* Against Covid-19



Notes: Panel (a) shows the daily Covid-19 deaths for Madrid, region $\mathcal{C}$, and for an artificial region $\mathcal{T}$ that aggregates the rest of Spain (RoSPA). We use a Chebyshev smoother (solid lines) of degree 6. Panel (b) shows the results of our normalization using region $\mathcal{T}$ as reference and mapping the pre-policy outcome paths of region $\mathcal{C}$ onto region $\mathcal{T}$. Panel (c) zooms the identification window. Panel (d) shows the policy effect where γ is defined in equation (8). We show the mean, median and 90% confidence interval bands from bootstrapped simulations constructed as described in Section 3.3. We estimate a significant auto-correlation coefficient for the residuals ($\rho_{\mathcal{C}} = 0.37$ $\rho_{\mathcal{T}} = 0.65$, respectively) and thus perform block-bootstrap with a block window of 5 days.

policy not been implemented. These effects are significant with a 90% confidence interval of [-29.71,-19.30]. The median effect is similar, -26.45%. Further, considering the policy effect across all bootstrap draws (i.e., without restricting the window size) we find similar significant policy effects with mean -22.11% and median -22.95%. Last, redoing our assessment without the smoothing step implies that the policy prevented 25.61% of the deaths in RoSPA during approximately the first week.

4.2 The 1960 FDA Approval of Oral Contraceptives in the U.S.

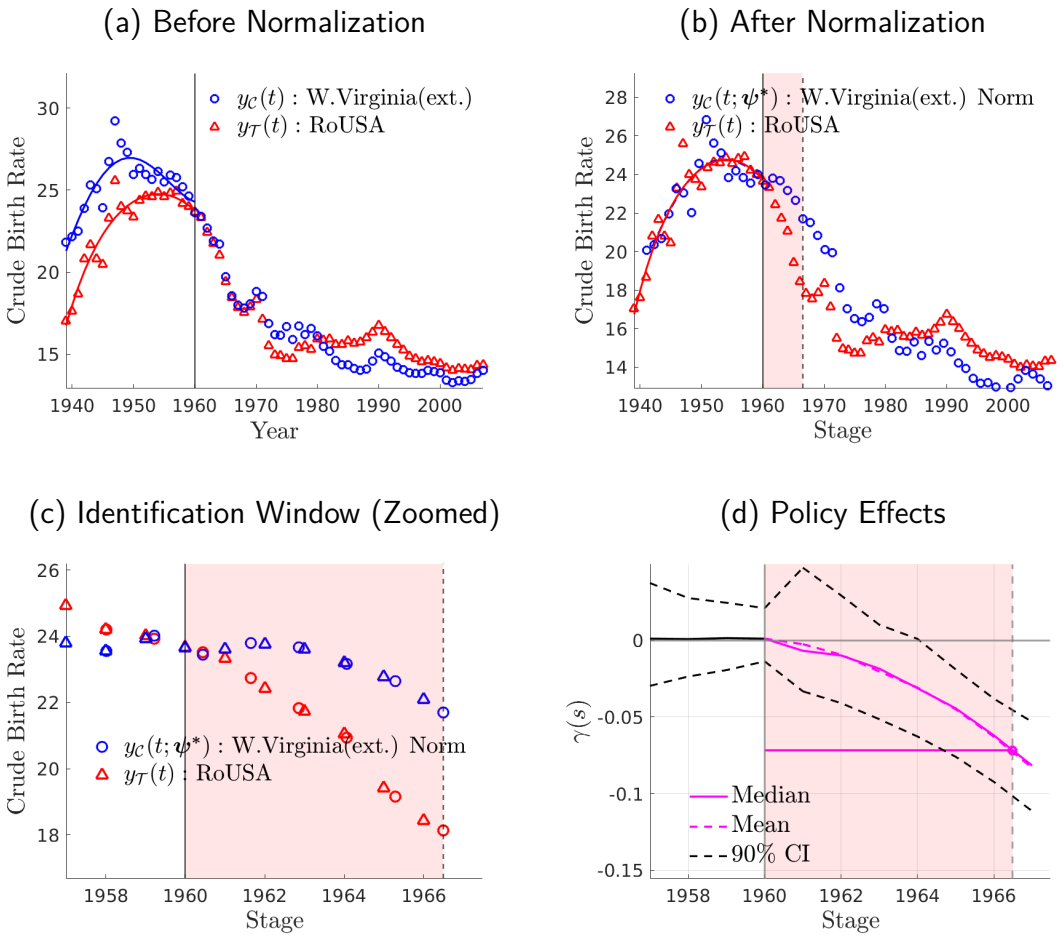
In 1960, the first hormonal birth control pill (oral contraceptive) was approved in the U.S. by the Food and Drug Administration (FDA). The use of pill was approved for use by women above the age of majority. In a seminal paper, Goldin and Katz (2002) use state-level variation in the age of majority in order to assess how women in that threshold change schooling and career choices.[47] Since SBI does not require non-nationwide policy for identification (e.g. state-time variation of the policy that determines the age of majority), we assess the effects of the nationwide (federal) approval of the pill on the entire population of adult women. We focus on two outcome variables. First, we study the effects of the pill on women's fertility choices—crude birth rates, using as regions the state of West Virginia (ext.) and the rest of the United States (RoUSA), where we label West Virginia (ext.) as region $\mathcal{C}$ and RoUSA as region $\mathcal{T}$.[48] Second, we study the effects of the pill on women's college choices—the share of women with completed college by age 25, using as regions the state of Washington D.C (ext.) and RoUSA, where we label Washington D.C. (ext.) as region $\mathcal{C}$ and RoUSA as region $\mathcal{T}$.

The crude birth rates shows a inverted-U shape pattern typically labeled as the baby boom and baby bust; see panel (a) in Figure 16. We find differential patterns across states. In particular, the birth rate in the region of West Virginia (ext.) peaks in the second half of the 1940s and in 1960 is already busting and close the 1940 levels. Instead, the birth rate in the RoUSA peaks in the second half of the 1950s at somewhat lower level and, on average, has barely started to decline by year 1960. In terms of women's college completion, the proportion of women of age 25 with completed college attainment has more than tripled over a span of twenty years raising from 8% in 1950 to 26% in 1970 in the leading states; see panel (a) of Figure 17. In the RoUSA, the proportion of women of age 25 with completed college attainment shows a larger relative increase from 2% in 1950 to 15% in 1970.

[47] Further, Bailey (2006) uses state-level variation in the age of majority to assess the effects of the pill on the timing of first births and women's labor force participation. Greenwood and Guner (2010) use an equilibrium matching model to assess the effects of oral contraceptives on premarital sex and how it is perceived in society.

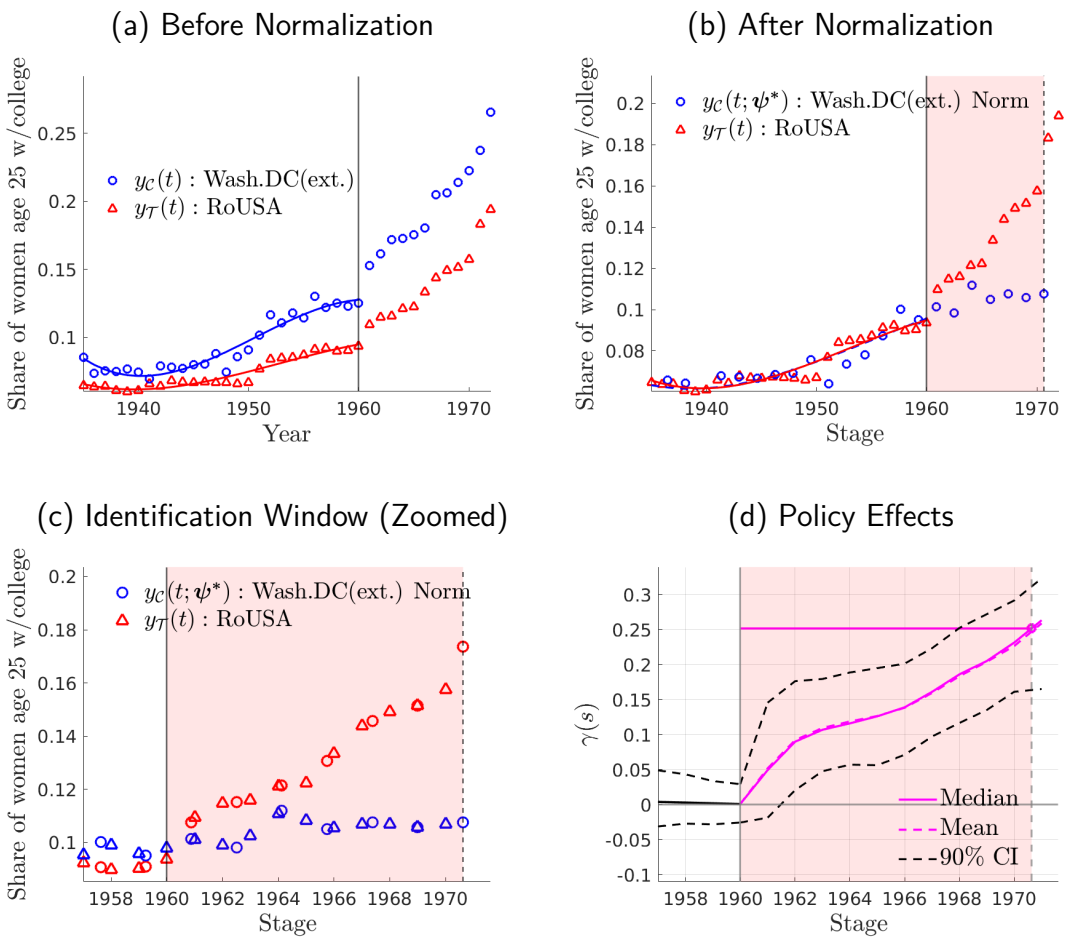
[48] Analogously to what we did in Section 4.1, we normalize the each U.S. state's time path of crude birth rates to the aggregate path of the United States. This normalization uncovers that West Virginia leads the rest of the United States (RoUSA) in that it shows the largest cross-regional stage at the time of policy implementation. In order to increase the sample size, we further add the next three leading states (Idaho, Nevada and Arkansas) to construct an artificial region as the population weighted average of these four regions, which we label West Virginia (ext.). Then, we also construct an artificial region that consists of the RoUSA. For the case of the share of women that at age 25 have completed college, the leading state is Washington D.C. Since this state has a relatively small population size, we construct an artificial region as the population weighted average that additionally includes the next three leading states in terms of women college completion (Massachusetts, Colorado and Connecticut). We name this artificial region as Washington D.C. (ext.).

Figure 16: The Effects of the 1960 FDA Approval of Oral Contraceptives: Crude Birth Rate



Notes: Panel (a) shows the crude birth rate for a region $\mathcal{C}$ which consists of a set of states leading the fertility bust (West Virginia, Idaho, Nevada and Arkansas) and a region $\mathcal{T}$ that aggregates the rest of the United States. We use a Chebyshev smoother (solid lines) of degree 5. Panel (b) shows the results of our normalization using region $\mathcal{T}$ as reference and mapping the pre-policy outcome paths of region $\mathcal{C}$ onto region $\mathcal{T}$. Panel (c) zooms the identification window. Panel (d) shows the policy effect $\gamma(s)$ as defined in equation (8). We show the mean, median and 90% confidence interval bands from bootstrapped simulations constructed as described in Section 3.3. We find a non-significant auto-correlation coefficient for the residuals, $\rho_{\mathcal{C}} = 0.31$ $\rho_{\mathcal{T}} = 0.18$, respectively.

Figure 17: The Effects of the 1960 FDA Approval of Oral Contraceptives: Women College



Notes: Panel (a) shows the proportion of women of age 25 that completed college for a region $\mathcal{C}$ which consists of a set of states leading women's college completion (Washington D.C., Massachusetts, Colorado and Connecticut) and a region $\mathcal{T}$ that aggregates the rest of the United States. We use a Chebyshev smoother (solid lines) of degree 4. Panel (b) shows the results of our normalization using region $\mathcal{T}$ as reference and mapping the pre-policy outcome paths of region $\mathcal{C}$ onto region $\mathcal{T}$. Panel (c) zooms the identification window. Panel (d) shows the policy effect $\gamma(s)$ as defined in equation (8). We show the mean, median and 90% confidence interval bands from bootstrapped simulations constructed as described in Section 3.3. We find a non-significant auto-correlation coefficient for the residuals, $\rho_{\mathcal{C}} = 0.21$ and $\rho_{\mathcal{T}} = 0.64$, respectively.

Normalization. In terms of crude birth rates, picking RoUSA as reference region $\mathcal{T}$, we apply our normalization by mapping the pre-policy birth rates of West Virginia (ext.) ($y_{\mathcal{C}}(t)$, solid blue circles) onto the pre-policy crude births rates of the RoUSA ($y_{\mathcal{T}}(t)$, solid red circles). This results in a normalized path for West Virginia (ext.) $\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}^*)$; see panel (b), Figure 16. The estimates are ψ_0 =1.85 [-0.53,9.96], ψ_1 =1.21 [0.41,1.58] and ω_1 =0.91 [0.89,0.94] which, respectively, delays the start, slows down the growth, and lowers the peak of the baby boom for the leading region in stages.[49] A result of the normalization, West Virginia (ext.) leads the crude births rate path in that it is in a more advanced stage than RoUSA at the time of policy implementation, i.e. $s_{\mathcal{T}}(t_p;\boldsymbol{\psi}^*) = t_p < s_{\mathcal{C}}(t_p;\boldsymbol{\psi}^*)$. Hence, the normalization unveils a window of stages $\mathbb{W}(s;\boldsymbol{\psi}^*) = [t_p, s_{\mathcal{C}}(t_p;\boldsymbol{\psi}^*)]$ (shaded pink area) in which West Virginia (ext.) is not subject to policy whereas RoUSA is. For the case of women's college completion, we choose RoUSA as reference region $\mathcal{T}$ and map the pre-policy path of Washington DC (ext.) ($y_{\mathcal{C}}(t)$, solid blue circles) onto that of RoUSA ($y_{\mathcal{T}}(t)$, solid red circles), which generates a normalized path for Washington DC (ext.) $y_{\mathcal{C}}(s;\boldsymbol{\phi}^*)$; see panel (b), Figure 17. The normalizing parameters are ψ_0 =-5.65 [-11.17,1.84], ψ_1 =1.62 [1.07,1.73] and ω_1 =0.85 [0.76,0.92] which results in Washington D.C. (ext.) as stage-leading region at t_p.

Policy Effect. Following our identification assumption, the stage-leading region in birth rates at the time policy implementation, West Virginia (ext.), serves as no-policy counterfactual for RoUSA inside $\mathbb{W}(s;\boldsymbol{\psi}^*)$; see the outcome paths $y_{\mathcal{C}}(s;\boldsymbol{\phi}^*)$ (normalized West Virginia (ext.)) and $y_{\mathcal{T}}(t)$ (RoUSA) inside the identification window in panel (c), Figure 16. The policy significantly reduced by γ = -8.36% the number of births (per 10,000 inhabitants) that would have otherwise occurred without the pill; panel (d), Figure 16. The median effects are similar: a -6.94% reduction. In the previous effects, the window size is restricted to be in the neighborhood of the median window size (plus/minus 10%). Not restricting the window size, we also find significant effects, γ = -7.53%. Analogously, for the the share of women that completed college education at age 25, the stage-leading region at the time policy implementation, Washington D.C. (ext.), serves as no-policy counterfactual for RoUSA. The FDA approval of oral contraceptives significantly increased the share of women with completed college at age 25 by γ = 24.69% during the decade that followed the policy compared to what would have occurred without the pill; panel (c) and

[49] The fact that the policy happens after peak of the crude birth rate can provide a role for an asymmetry parameter in the stage-to-time transformation, i.e. adding the monomial basis $\psi_2 s^2$. However, at the time of policy implementation the decline in the crude birth rate for the non-leading region has barely started and when we introduce an asymmetry parameter ψ_2 we find that is not significantly different from zero.

(d), Figure 17. The median effects are almost identical, $\gamma = 24.00\%$. Not restricting the window size also yields significant effects of $\gamma = 19.19\%$.[50]

4.3 The German Reunification

In 1990, after the fall of the Berlin wall in 1989, the German Democratic Republic was abolished and integrated fully into the Federal Republic of Germany. Given large differences between the West German states and the East German states, the political and economic integration came at some cost—the size of which is subject to debate. Abadie et al. (2014) study the consequences of the German reunification for West Germany and forming a counterfactual path for GDP per capita using a Synthetic Control Group (SCG) approach. Here, we apply SBI to the same context, and construct a counterfactual for the evolution of GDP per capita in West Germany had it not been for the reunification. In contrast with Abadie et al. (2014), our counterfactual is constructed using the GDP per capita paths of West German regions only.[51] To conduct our analysis, we focus Hessen and an artificially created region for Rest of West Germany (RoGER) which is composed of all West Germany regions excluding Hessen.[52] We label Hessen as region $\mathcal{C}$ and RoGER as region $\mathcal{T}$; see panel (a) of Figure 18.

Normalization. Picking RoGER as reference region, we apply our normalization by mapping the pre-policy GDP per capita of Hessen ($y_{\mathcal{C}}(t)$, solid blue circles) onto that of RoGER ($y_{\mathcal{T}}(t)$, solid red circles). This results in a normalized path for Hessen $\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\phi}^*)$; see panel (b) of Figure 16. The normalizing parameters are ψ_0 =1.98 [1.51,6.69], ψ_1 =1.24 [1.13,1.48] and ω_1 =1.00 [1.00,1.01]. A result of the normalization is that Hessen leads RoGER in stages at the time of policy implementation. Precisely, our normalization opens a window in stages, $\mathbb{W}(s;\boldsymbol{\psi}^*) = [t_p, s_{\mathcal{C}}(t_p;\boldsymbol{\psi}^*)]$ (shaded pink area), running from approximately seven years in which Hessen is not subject to the German reunification but RoGER is. Therefore, under our identification assumption, the normalized path for Hessen provices a no-policy counterfactual for RoGER inside that window.

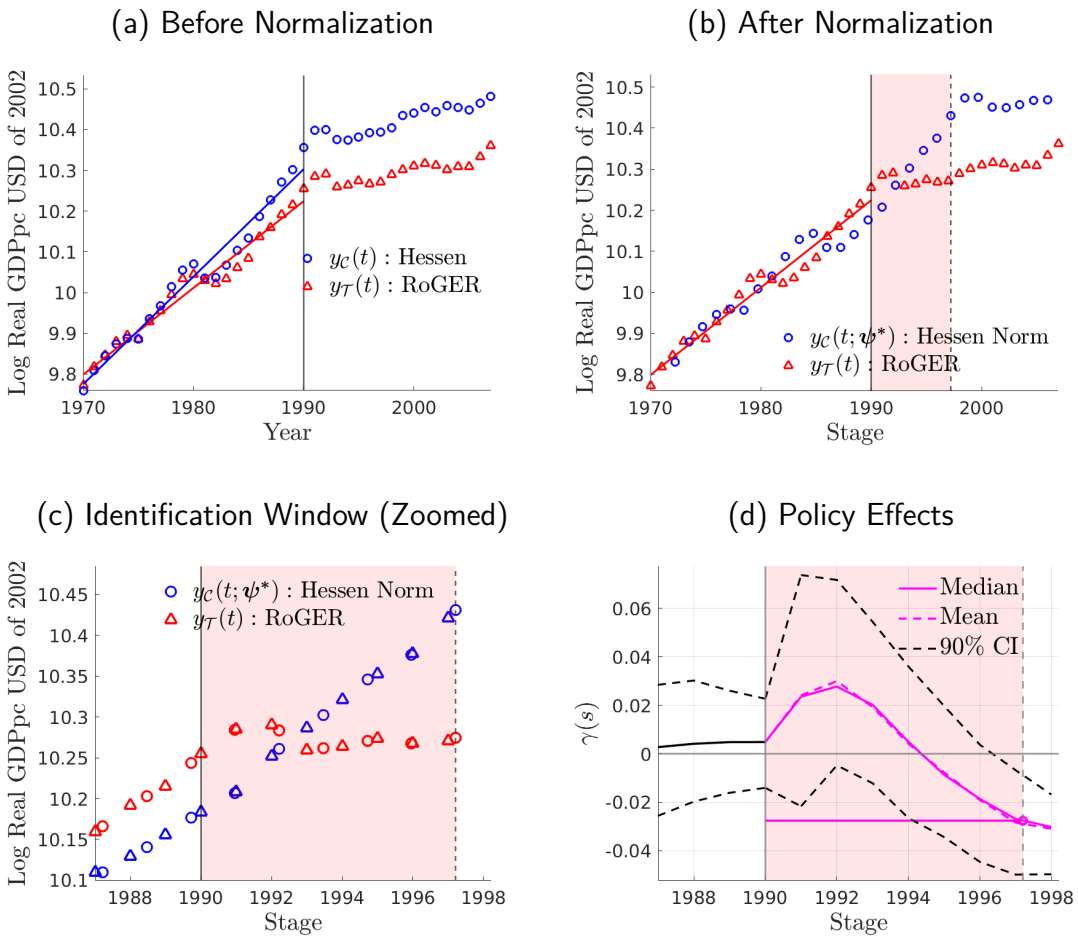
Policy Effect. We zoom in on the identification window for the GDP per capita in panel (c) and the associated policy effects in panel (d) of Figure 16. We find that the German Reunification significantly reduced the GDP per capita of RoGER by $\gamma = 3.39\%$ compared to the GDP that it

[50] Without the smoothing step, the FDA approval of oral contraceptives implies a reduction of -14.81% in the number of births and an increase in the share of women of age 25 that complete college education by 18.25%. These effects are not significantly different from the mean bootstrapped effects.

[51] In Section 5.2.1, we compare the two approaches SBI and SCG in more detail.

[52] Analogously to Section 4.1, we normalize the path of GDP per capita of West Germany states to the aggregate GDP per capita path of West Germany. This normalization uncovers that the stage at which Hessen is at the time of the German reuinification is the most advanced across regions. That is, Hessen leads the rest of Germany (RoGER) at the time of policy implementation. For this reason, we focus on Hessen and RoGER for our analysis.

Figure 18: The Effects of the German Reunification on GDP per capita



Notes: Panel (a) shows the GDP per capita of region $\mathcal{C}$, Hessen, that leads West Germany and a region $\mathcal{T}$ that aggregates the rest of West Germany. We use a Chebyshev smoother (solid lines) of degree 3. Panel (b) shows the results of our normalization using region $\mathcal{T}$ as reference and mapping the pre-policy outcome paths of region $\mathcal{C}$ onto region $\mathcal{T}$. Panel (c) zooms the identification window. Panel (d) shows the policy effect $\gamma(s)$ as defined in equation (8). We show the mean, median and 90% confidence interval bands from bootstrapped simulations constructed as described in Section 3.3. We find a significant auto-correlation coefficient for the residuals ($\rho_{\mathcal{C}} = 0.78$ and $\rho_{\mathcal{T}} = 0.74$, respectively) and thus perform block-bootstrap. We use a block window of 3 years.

would have otherwise attained without the Reunification. The median policy effects are similar, $\gamma = 3.30\%$. In the previous effects, the window size is restricted to be in the neighborhood of the median window size (plus/minus 10%). Again, not restricting the window size, we also find significant effects of similar size $\gamma = 3.29\%$. Further, without the smoothing step, the German reunification generates a reduction of 4.82% in the GDP per capita of RoGER, which is not significantly different from our mean bootstrapped effects.

5 Further Discussion

Here, we first discuss the heterogeneity of policy effects by stage at the time of policy implementation in Section 5.1. Second, we show that SBI can be applied to non-nationwide policy including cases where there are untreated regions or a staggered rollout of the policy in section 5.2. Interestingly, we show that in the context of a staggered rollout policy, SBI can endogenously uncover a flip between standardly defined control region and treatment region.

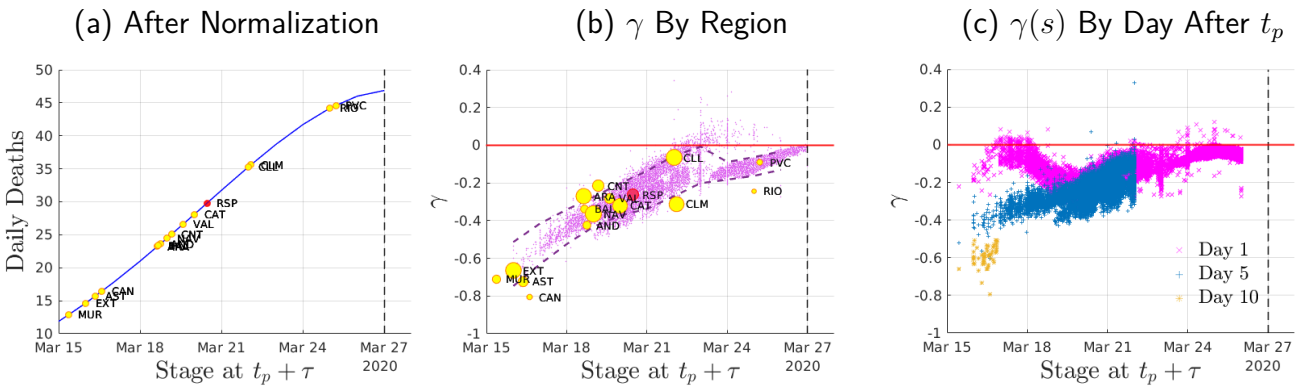
5.1 Heterogeneous and Aggregate Policy Effects

So far, we have only entertained the idea of policy evaluation with two regions (or sets of regions) in the models of Section 3 and the applications of Section 4. We now expand our analysis to assess the policy effects for multiple regions and stages, including an extension that builds the power set of treated regions.

To conduct our analysis, we focus on the stay-home nationwide policy implemented in the first wave of Covid-19 in Spain across all regions at the same time; see Section 4.1. The idea is to use the available paths of Covid-19 deaths of multiple regions—which might differ by stage at the time of policy implementation—in order to measure the heterogeneous effects of policy by stage. That is, we apply SBI by region. In particular, since Madrid leads all other regions we pick Madrid as reference and separately map each region to Madrid. Hence, the stage domain is the same or all regions. Madrid. The result of this mapping is in panel (a) of Figure 19, where since we picked Madrid as reference region, the calendar time is the stage for Madrid. Then, for example, for the region of Murcia (MUR) policy implementation occurs at a stage corresponding to approximately twelve days earlier than the stage at which Madrid received the policy. The closest to Madrid in terms of stages of the epidemic is the Basque Country (PVC) that lags Madrid for approximately two days.

Our main result is that there are heterogeneous policy effects by region that we plot (yellow markers) in panel (b) of Figure 19. Clearly, the effects of policy are largest for the regions when the stage at the time of policy implementation differs the most with respect to the stage in which Madrid enters the policy. For example, in the region of Murcia, the policy effect reaches $\gamma = -0.65$ and, hence, the policy prevented 65% of the deaths that would have otherwise occurred in Murcia in a scenario without policy. In contrast, in the Basque Country, which is closest to Madrid in terms of stages at the time of policy implementation, the policy effects are $\gamma = -0.12$ and, hence, the policy prevented 12% of the deaths that would have otherwise occurred in the Basque Country in a scenario without policy. To further complete our exploration, we construct hybrid regions from the power set of the treated regions, i.e. a total of $2^{16} - 1 = 131,072$ hybrid

Figure 19: Heterogeneous Policy Effects by Stage



Notes: We have a total of 17 region (comunidad autonoma) names: Andalucia (AND), Aragon (ARA), Asturias (AST), Baleares (BAL),Canarias (CAN), Cantabria (CNT), Castilla-La Mancha (CLM), Castilla y Leon (CLL), Catalunya (CAT), Ceuta(CEU), Valencia (VAL), Extremadura (EXT), Galicia (GAL), Madrid (MAD), Melilla (MEL), Murcia (MUR),Navarra (NAV), Pais Vasco (PVC), La Rioja (RIO). We exclude GAL from the analysis due to the fact that we find positive (yet, non-significant) effects of the policy on the flow of deaths. The size of the yellow is the stock of deaths per thousand inhabitants accumulated during the identification window. In panel (b), we report the policy effects γ (see Section 2.2) by region where the (yellow) marker size is the flow of deaths at the time of policy implementation. In addition to the policy effects by region, we also report the policy effects for each hybrid region constructed for each element in the the power set $2^{16} - 1$ of regions (tiny markers) in panel (b). In panel (b), the 90% CI's exclude the top 5% and bottom 5% of policy effects by stage in rolling windows of 2 stages/days. In panel (c) we show the interim effect $\gamma(s)$ (see Section 2.2) by stage for day 1, day 5 and day 10 after policy implementation.

regions, that we separately map using SBI to Madrid.[53] We report the policy effects (tiny purple markers) associated with each of these hybrid regions (with 90% confidence intervals) in panel (b) of Figure 19. We reach similar insights as the policy effects are largest in instances where the the stage at the time of policy implementation is farthest away from the stage at which Madrid implemented the policy.

What drives the differences in policy effects by stage? An obvious candidate to determine these differences is the size of the identification window—i.e. the closest a region is to Madrid in terms of stages at the time of policy implementation, the smaller is the identification window. At the same time, differences in policy effects can emerge within the same horizon into the policy within the identification window. To assess this question, we isolate the effects of policy by the number of stages within the identification window. Here, note that since we picked Madrid as reference, the stage for Madrid is the actual calendar time (i.e. days). In those terms, we find substantial heterogeneity across identification windows by stage. For example, one day into the policy (i.e inside the identification window) at a stage of approximately 10 days before Madrid

[53] Precisely, a hybrid path between region A and region B is constructed as the weighted sum of the flow of deaths per capita in each region.

enters policy (e.g. March 18) the policy effect is below 10%, whereas one day into the policy at a stage of approximately 7 days before Madrid enters policy (e.g. March 12) the policy effect is above 10%, and one day into the policy at a stage of approximately 3 days before Madrid enters policy (e.g. March 24) is again below 10%; see the magenta markers in panel (c) of Figure 19. We also show differences across stages in the policy effects for the cases of five days (blue markers) and ten days (yellow markers) into the policy. That is, not only the size of the identification window matters (i.e. how close in stages a given region is to Madrid at the time of policy implementation) but there are also differences in policy effects driven by the differential within-window policy effects across identification windows.

5.2 Non-Nationwide Policy

Here, we show that our identification strategy works in scenarios where there are regions that never receive the policy intervention in Section 5.2.1 and scenarios where there is a staggered rollout of the policy in Section 5.2.2.

5.2.1 Untreated Regions

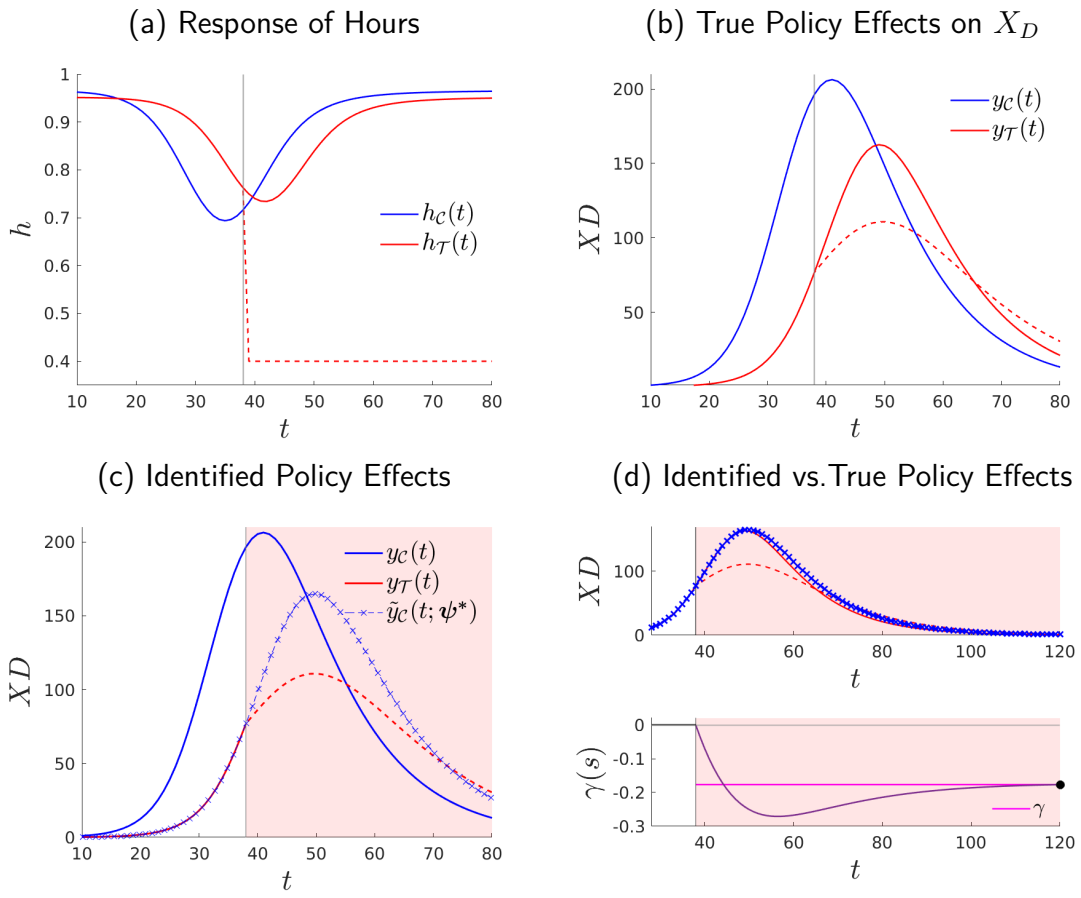
Consider a scenario with two regions where one region, e.g. $\mathcal{T}$, receives the policy intervention at period t_p and the other region, e.g. $\mathcal{C}$, is never treated. To illustrate this scenario we use our benchmark model with endogenous pandemics as described in Section 3.1.1. In that context, we introduce the stay-home policy that puts an upper bound on hours worked in region $\mathcal{T}$, but not in region $\mathcal{C}$; see panel (a) of Figure 20. The implications for the outcome of interest, the flow of deaths, is displayed in panel (b) of Figure 20. The policy has an impact on region $\mathcal{T}$, but not on region $\mathcal{C}$.

In order to identify the policy effects, we need to modify the set on which the normalization is conducted. In particular, picking region $\mathcal{T}$ as reference, the normalization parameters are the solution to the minimization of (5) subject to (3) and

$$\mathbb{C}(s) = \begin{cases} [t_0, t_p] & \text{if } \; r = \mathcal{T} \\ [s_{\mathcal{C}}(t_0; \boldsymbol{\psi}^*), s_{\mathcal{C}}(t_f; \boldsymbol{\psi}^*)] & \text{if } \; r = \mathcal{C} \end{cases} \tag{28}$$

for $r = \{\mathcal{C}, \mathcal{T}\}$ where t_0 denotes the first period of observed data and t_f the last.

Figure 20: Stage-Based Identification of Policy Effects: Untreated Regions



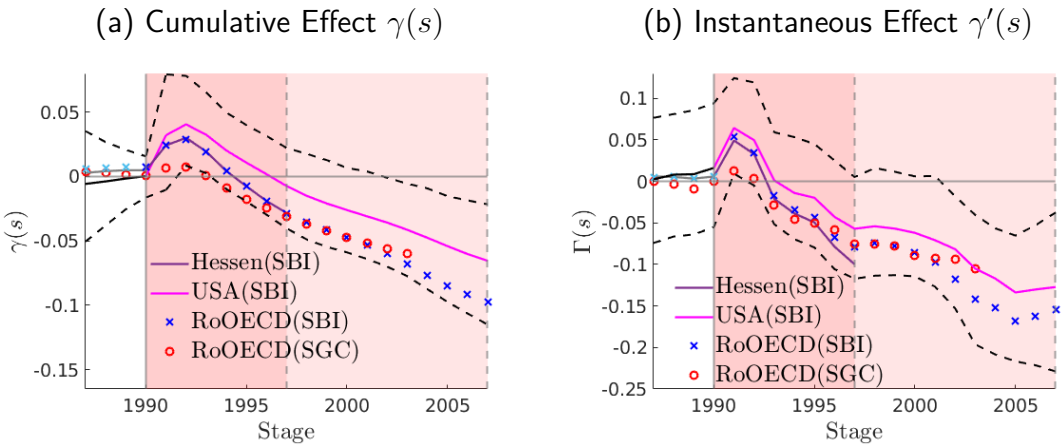
Notes: Where $\bar{h} = 0.4, t_p = 38, t_f = 250$, $\gamma = 17.66\%$ and $\epsilon(\gamma) = 58.19\%$

The results of the normalization are shown in panel (c) of Figure 20. Note that since only region $\mathcal{T}$ is treated, the identification window is,

$$\mathbb{W}(s;\boldsymbol{\psi}^*) = \begin{cases} [t_p, s_{\mathcal{C}}(t_f;\boldsymbol{\psi}^*)] & \text{if } \; s_{\mathcal{C}}(t_f;\boldsymbol{\psi}^*) > t_p \\ \varnothing & \text{if } \; s_{\mathcal{C}}(t_f;\boldsymbol{\psi}^*) < t_p \end{cases} \tag{29}$$

Hence, there is a policy effect if and only if the stage of normalized series evaluated at the last period of observed data $y_{\mathcal{C}}(s;\boldsymbol{\phi}^*)$ with $s = s_{\mathcal{C}}(t_f;\boldsymbol{\psi}^*)$ falls beyond the period of policy implementation t_p. Otherwise, the identification window is the empty set because the treated region leads throughout the entire sample. In our model-generated example, the normalization shows that region $\mathcal{C}$ covers stages beyond that of region $\mathcal{T}$ at the time of policy implementation which implies that $\mathbb{W}(s;\boldsymbol{\psi}^*) = [t_p, s_{\mathcal{C}}(t_f;\boldsymbol{\psi}^*)]$ and we can assess policy effects. We show the

Figure 21: The Effects of the German Reunification: Stage-Based Identification (SBI) and Synthetic Control Methods (SCM)



Notes: The outcome variable is real GDP per capita in USD of 2002. The plotted 90% confidence intervals correspond to the United States.

identified policy effects in panel (d) of Figure 20, for which we find that overlap with the true (model-generated) policy effects.

We further exemplify how to use stage-based identification in cases in which not all regions are treated by re-conducting our assessment of the German Reunification. Here, we take West Germany as the treated region and use as potential controls the United States and an aggregate consisting of the same sample of OECD countries (that excludes Germany) studied in Abadie et al. (2014). Hence, this exercise also serves as means for comparison between SBI and SCM (Abadie and Gardeazabal, 2003). In order to apply SBI, we pick West Germany as reference region. Then we conduct the normalization by mapping the GDP per capita path of the U.S. and the OECD aggregate onto the GDP per capita path of West Germany. We show the policy effects that emerge from SBI in panel (a) of Figure 21. To ease the comparison with Abadie et al. (2014), we also show in panel (b) of Figure 21 the instantaneous policy effects by stage defined as (abusing some notation), $\gamma'(s) = \frac{y_{\mathcal{T}}(s) - \widetilde{y}_{\mathcal{C}}(s)}{\widetilde{y}_{\mathcal{C}}(s)}$. That is, $\gamma'(s)$ measures the change in GDP per capita of region $\mathcal{T}$ (West Germany) relative to the counterfactual region $\mathcal{C}$ (e.g. the U.S. or the rest of the OECD) at any given stage s due to policy. We further show the results from using Hessen as leading region in the context of nationwide policy within Western Germany reported in Section 4.3. For inference, we show the 90% confidence intervals associated to the USA constructed as described in Section 3.3.

Our main finding is that the policy effects that emerge from using SBI either for the U.S.—or the OECD aggregate that excludes Germany—are not significantly different from those obtained

using SCM in Abadie et al. (2014). In particular, the instantaneous policy effects imply a loss of income per capita for West Germany due to the Reunification of 12.73% when compared to the United States and of 15.44% when compared to the rest of the OECD in 2007. These figures are, respectively, 10.51% and 14.13% in 2003 which are not significantly different from the effects of reunification of 10.04% obtained in Abadie et al. (2014). Further, we also find that within the shorter window that emerges when the counterfactual from the SBI strategy is Hessen as in Section 4.3, the results under the alternative counterfactuals are not significantly different from the results obtained with Hessen.

5.2.2 Staggered Rollout

Now, consider a scenario with two regions where one region, e.g. $\mathcal{T}$, receives the policy intervention at period $t_p^{\mathcal{T}} = t_p$ and the other region, e.g. $\mathcal{C}$, receives the policy at a later date $t_p^{\mathcal{C}} = t_p + \Delta$ with $\Delta > 0$. This scenario exemplifies the staggered rollout of policy.[54] In panel (a) of Figure 22, we introduce the same stay-home policy that we discussed in our benchmark example in Section 3.1.1, i.e. an upper bound on hours worked. We show the true (model-generated) policy effects of the rollout policy on the daily deaths by region in panel (b) of Figure 22.

We apply SBI in this context, and pick region $\mathcal{T}$ as reference to minimize (5) subject to (2) and

$$\mathbb{C}(s) = \left[s_{\overline{r}} \left(t_0^{\overline{r}}, \boldsymbol{\psi}^* \right), s_{\underline{r}} \left(t_p^{\underline{r}}, \boldsymbol{\psi}^* \right) \right] \tag{30}$$

for $r = \{\mathcal{C}, \mathcal{T}\}$ where $\overline{r} = \arg\max_r s_r \left(t_0^r; \boldsymbol{\psi}^* \right)$ is the region that at time t_0 is at the most advanced stage and $\underline{r} = \arg\min_r s_r \left(t_p^r; \boldsymbol{\psi}^* \right)$ is the region that at time t_p is at the earliest stage. That is, we generalize the set on which the normalization must be conducted in order to accommodate for differences in the time of policy implementation across regions.

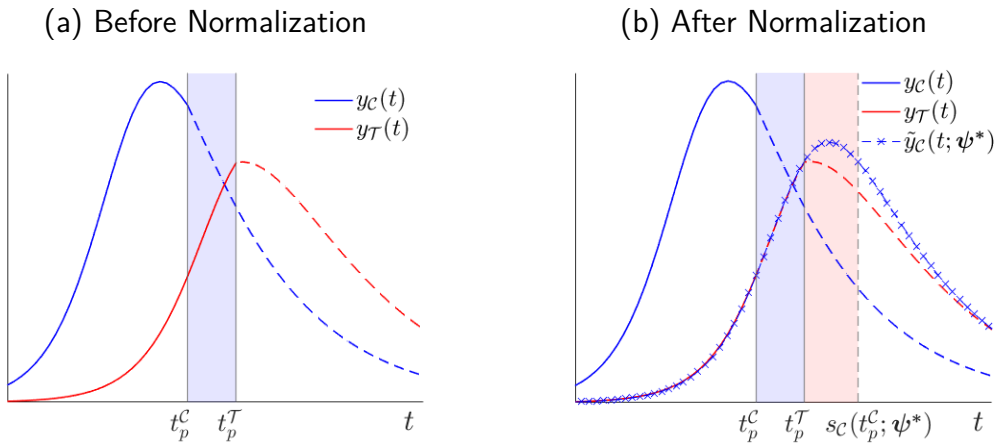
The result of the normalization is shown in panel (c) of Figure 22 which opens a window in stages (pink shaded area),

$$\mathbb{W}(s) = \left[s_{\underline{r}} \left(t_p^{\underline{r}}, \boldsymbol{\psi}^* \right), s_{\overline{r}} \left(t_p^{\overline{r}}, \boldsymbol{\psi}^* \right) \right] \tag{31}$$

where $\overline{r} = \arg\max_r s_r \left(t_p^r; \boldsymbol{\psi}^* \right)$ is the region that at time t_p is at the most advanced stage. In our illustration, the normalization delivers $t_p^{\mathcal{T}} \leq s_{\mathcal{C}}(t_p^{\mathcal{C}}; \boldsymbol{\psi}^*)$ and, therefore, $\mathbb{W}(s; \boldsymbol{\psi}^*) = \left[t_p^{\mathcal{T}}, s_{\mathcal{C}}(t_p^{\mathcal{C}}; \boldsymbol{\psi}^*) \right]$. That is, even though the policy was implemented earlier in time in region

[54] See the recent analysis in Goodman-Bacon (2021) for a careful assessment of DiD strategies that can be used under different timings and scenarios of staggered rollout policies.

Figure 22: Stage-Based Identification of Policy Effects: Staggered Rollout: Example with a Control and Treatment Flip



Notes: Panel (a) shows a settig in which policy is introduced first (in time) in region $\mathcal{C}$, then in region $\mathcal{T}$. Panel (b) shows the normalized paths, which reveal that in terms of stages, policy is introduced first in region $\mathcal{T}$, then in region $\mathcal{C}$.

$\mathcal{C}$ than in region $\mathcal{T}$, the stage at which the policy is implemented in region $\mathcal{C}$ is more advanced than the stage at which the policy is implemented in region $\mathcal{T}$. Then, following our identification assumption, we show that the identified policy effects that emerge from SBI capture the true (model-generated) policy effects; see panel (d) of Figure 22.

We emphasize two main differences of SBI with respect to the standard empirical methods for the assessment of rollout policies as implemented in the illustration in Figure 22. First, in standard strategies the identification of the effects of rollout policies relies on the heterogeneity in the time of policy implementation across regions. Precisely, in the model-generated example the standard identification strategy corresponds to the (difference-in-differences) window $\mathbb{DD}(t) = \left[t_p^{\mathcal{C}}, t_p^{\mathcal{T}}\right]$ (shaded purple area) where $t_p^{\mathcal{C}} < t_p^{\mathcal{T}}$ and, hence, region $\mathcal{T}$ serves as control (without policy) for region $\mathcal{C}$ in all periods $t \in \mathbb{DD}(t)$. Instead, our identification resides on the stage domain using heterogeneity in the stage at the time of policy implementation—including cases in which policy implementation occurs at different dates for different regions. Specifically, in the illustration our strategy shows that the control and treatment regions that emerge from SBI—which catch the true (model-generated) effects—actually reverts the control and treatment that would be used in standard empirical strategies. Second, the outcome paths of regions $\mathcal{C}$ and $\mathcal{T}$ can be far from showing the parallel trends inside the window $\mathbb{DD}(t)$, which is typically required in the context of standard empirical strategies to warrant the policy assessment. Indeed, within the realm of empirical strategies that use heterogeneity in the time of policy implementation, there is

an exciting debate regarding more flexible forms of parallel trend assumptions (e.g. Callaway and Sant'Anna, 2021; Rambachan and Roth, 2021). In this context, in our illustration, the outcome path of region $\mathcal{T}$ is monotonically increasing before $\mathbb{DD}(t)$, whereas the outcome path of region $\mathcal{C}$ shows a non-monotonicity by first rising, reaching a peak and then already declining before entering $\mathbb{DD}(t)$. This behavior of the outcome paths before policy implementation—in calendar time—makes standard strategies largely unworkable. However, as we showed, SBI—which aims to minimize regional differences of the pre-policy outcome paths in the stage domain—does not rely upon the parallel trends over the time domain in order to provide a policy assessment that is credible—in that it is able to recover the true model-generated effects.

6 Conclusion

We develop a novel empirical methodology—Stage-Based Identification—to evaluate the effects of nationwide policy, i.e., contexts where all regions experience the policy intervention at the same time. Our method consists of a normalization that maps the time-paths of regional outcomes onto a reference path—using only pre-policy data. Since the normalized regions can differ by *stage* at any point in time, the normalization uncovers heterogeneity in the *stage* at the time of policy implementation—even in instances where the implementation occurs at the same time across regions. We use this *stage* variation at the time of policy implementation to identify the policy effects: a *stage*-leading region serves as control for the other region(s) inside a window in which non-leading regions are subject to policy when the leading region is not. We showcase that the method identifies the true policy effects in various model simulations (in which the true effect is known), and illustrate the approach in several empirical applications: the effects of public health stay-home policies on the course of an epidemic (application: the national lockdown against Covid-19 in Spain), the effects of oral contraceptives on women's fertility and college education choices (application: the FDA nationwide approval of oral contraceptives in 1960 in the U.S.), and the effects of growth policy on GDP per capita (application: German Reunification). We also show that our method works in the evaluation of non-nationwide policy with untreated regions and with staggered rollouts.

References

Abadie, A. (2005). Semiparametric Difference-in-Differences Estimators. *Review of Economic Studies*, 72(1):1–19.

Abadie, A., Diamond, A., and Hainmueller, J. (2010). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association*, 105(490):493–505.

Abadie, A., Diamond, A., and Hainmueller, J. (2014). Comparative Politics and the Synthetic Control Method. *American Journal of Political Science*, 59(2):495–510.

Abadie, A. and Gardeazabal, J. (2003). The Economic Costs of Conflict: A Case Study of the Basque Country. *American Economic Review*, 93(1):113–132.

Adam, K., Marcet, A., and Beutel, J. (2017). Stock price booms and expected capital gains. *American Economic Review*, 107(8):2352–2408.

Aleman, C., Iorio, D., and Santaeulalia, R. (2022). A quantitative theory of the hiv epidemic: Education, risky sex and asymmetric learning. Working paper, Barcelona School of Economics.

Altonji, J. and Blank, R. (1999). Race and gender in the labor market. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 3, Part C, chapter 48, pages 3143–3259. Elsevier, 1 edition.

Angrist, J. D. and Krueger, A. B. (1999). Empirical strategies in labor economics. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 3 of *Handbook of Labor Economics*, chapter 23, pages 1277–1366. Elsevier.

Athey, S. and Imbens, G. W. (2006). Identification and inference in nonlinear difference-in-differences models. *Econometrica*, 74(2):431–497.

Athey, S. and Imbens, G. W. (2017). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives*, 31(2):3–32.

Athey, S. and Imbens, G. W. (2021). Design-based analysis in difference-in-differences settings with staggered adoption. *Journal of Econometrics*.

Atkeson, A. (2020). What will be the economic impact of covid-19 in the us? rough estimates of disease scenarios. Working Paper 26867, National Bureau of Economic Research.

Bailey, M. J. (2006). More Power to the Pill: The Impact of Contraceptive Freedom on Women's Life Cycle Labor Supply*. *The Quarterly Journal of Economics*, 121(1):289–320.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How Much Should We Trust Differences-In-Differences Estimates? *The Quarterly Journal of Economics*, 119(1):249–275.

Blundell, R. and Macurdy, T. (1999). Labor supply: A review of alternative approaches. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 3 of *Handbook of Labor Economics*, chapter 27, pages 1559–1695. Elsevier.

Borusyak, K., Jaravel, X., and Spiess, J. (2021). Revisiting Event Study Designs: Robust and Efficient Estimation. Working papers, University College London.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-Differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

Card, D. (1990). The impact of the mariel boatlift on the miami labor market. *ILR Review*, 43(2):245–257.

Card, D. (2022). Design-based research in empirical microeconomics. *American Economic Review*, 112(6):1773–81.

Card, D. and Krueger, A. B. (2000). Minimum wages and employment: A case study of the fast-food industry in new jersey and pennsylvania: Reply. *The American Economic Review*, 90(5):1397–1420.

Carlstein, E. (1986). The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *The Annals of Statistics*, 14(3):1171–1179.

Cervellati, M. and Sunde, U. (2015). The economic and demographic transition, mortality, and comparative development. *American Economic Journal: Macroeconomics*, 7(3):189–225.

Doudchenko, N. and Imbens, G. W. (2017). Balancing, regression, difference-in-differences and synthetic control methods: A synthesis.

Fang, H., Wang, L., and Yang, Y. (2020). Human mobility restrictions and the spread of the novel coronavirus (2019-ncov) in china. *Journal of Public Economics*, 191:104272.

Fernández-Villaverde, J., Greenwood, J., and Guner, N. (2014). From Shame To Game In One Hundred Years: An Economic Model Of The Rise In Premarital Sex And Its De-Stigmatization. *Journal of the European Economic Association*, 12(1):25–61.

Galor, O. and Weil, D. N. (2000). Population, technology, and growth: From malthusian stagnation to the demographic transition and beyond. *American Economic Review*, 90(4):806–828.

Glogowsky, U., Hansen, E., and Schächtele, S. (2021). How effective are social distancing policies? evidence on the fight against covid-19. *PLOS ONE*, 16(9):1–12.

Goldin, C. and Katz, L. F. (2002). The Power of the Pill: Oral Contraceptives and Women's Career and Marriage Decisions. *Journal of Political Economy*, 110(4):730–770.

Gollin, D., Parente, S., and Rogerson, R. (2002). The role of agriculture in development. *American Economic Review*, 92(2):160–164.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277.

Greenwood, J. and Guner, N. (2010). Social change: The sexual revolution. *International Economic Review*, 51(4):893–923.

Greenwood, J., Seshadri, A., and Vandenbroucke, G. (2005). The Baby Boom and Baby Bust. *American Economic Review*, 95(1):183–207.

Hansen, G. D. and Prescott, E. C. (2002). Malthus to solow. *American Economic Review*, 92(4):1205–1217.

Heckman, J. J., Ichimura, H., and Todd, P. (1998). Matching As An Econometric Evaluation Estimator. *Review of Economic Studies*, 65(2):261–294.

Herrendorf, B., Rogerson, R., and Valentinyi, A. (2014). Growth and structural transformation. In *Handbook of Economic Growth*, volume 2, chapter 06, pages 855–941. Elsevier, 1 edition.

Imbens, G. W. and Rubin, D. B. (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press.

Iorio, D. and Santaeulàlia-Llopis, R. (2010). Education, HIV Status, and Risky Sexual Behavior: How Much Does the Stage of the HIV Epidemic Matter? Working papers, washington university in st. louis.

Iorio, D. and Santaeulàlia-Llopis, R. (2016). Education, HIV Status, and Risky Sexual Behavior: How Much Does the Stage of the HIV Epidemic Matter? Working papers, Barcelona Graduate School of Economics.

Liu, L., Moon, H. R., and Schorfheide, F. (2021). Panel forecasts of country-level Covid-19 infections. *Journal of Econometrics*, 220(1):2–22.

Rambachan, A. and Roth, J. (2021). A More Credible Approach to Parallel Trends. Technical report.

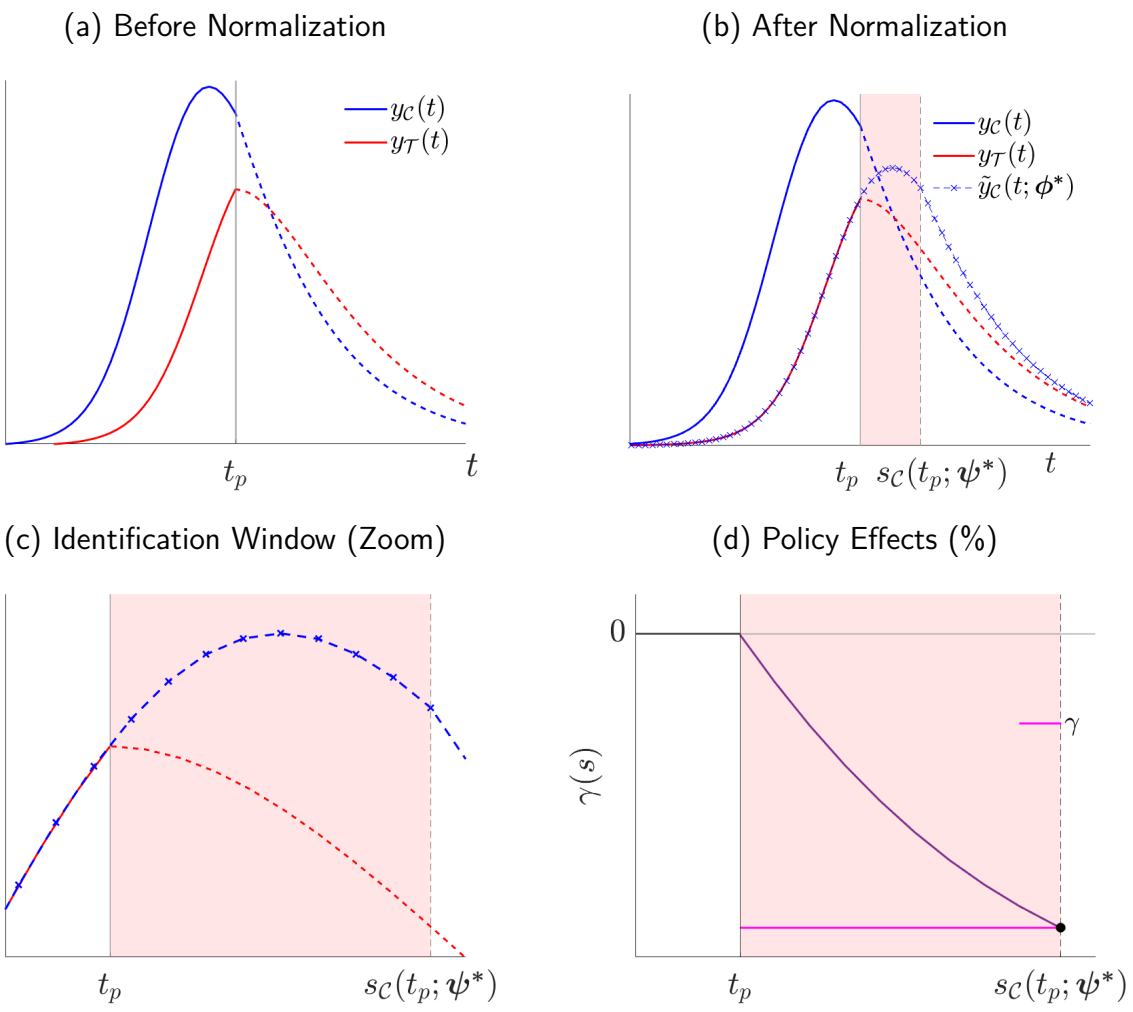
A Further Examples: Policy After the Peak

Here, we study cases in which the policy is implemented after the peak in addition to our before-the-peak benchmark in Section 2.

A.1 Policy After the Peak: One Region

Here we discuss a scenario in which the nationwide policy is implemented before the peak in region $\mathcal{T}$ and after the peak in region $\mathcal{C}$; see panel (a) of Figure 23. Note that in this case the outcome paths of region $\mathcal{T}$ and region $\mathcal{C}$ show clearly differentiated (not parallell) trends with region $\mathcal{C}$ monotonically increasing and region $\mathcal{T}$ displaying a non-monotic path. This makes standard empirical strategies such unworkable. In contrast, applying our method mapping region $\mathcal{C}$ onto region $\mathcal{T}$ using pre-policy data only generates the normalized outcome path $\widetilde{y}_{\mathcal{C}}(t;\boldsymbol{\psi}^*)$; see panel (b) in Figure 23. Hence, out methodology opens a window in stages between t_p and $s_{\mathcal{C}}(t;\boldsymbol{\psi}^*)$ in which the effects of policy are identified. We zoom the identification window in panel (c) of Figure 23 and the associated policy effects in panel (d) of Figure 23. In Figure **??**, we further unpack the contribution of each normalization coefficient in generating the normalized path.

Figure 23: A Stage-Based Identification of Policy Effects: A Nationwide Policy



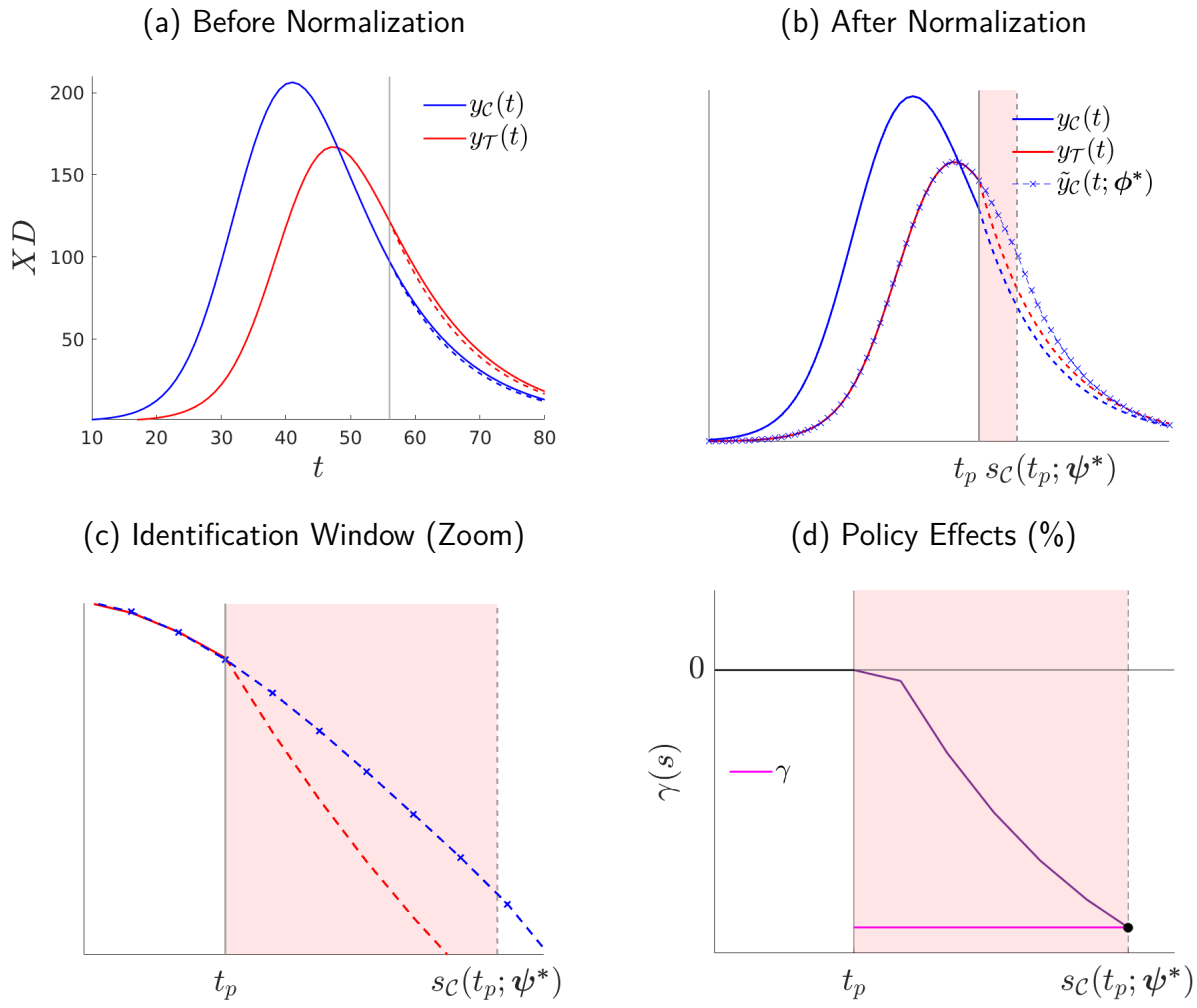
Notes: See the notes in Figure 3.

A.2 Policy After the Peak: Two Regions

Here we discuss a scenario in which the nationwide policy is implemented before the peak in region $\mathcal{T}$ and also inregion $\mathcal{C}$; see panel (a) of Figure 24. Applying our method mapping region $\mathcal{C}$ onto region $\mathcal{T}$ using pre-policy data only generates the normalized outcome path $\widetilde{y}_{\mathcal{C}}(t;\boldsymbol{\psi}^*)$; see panel (b) in Figure 24. Hence, out methodology opens a window in stages between t_p and $s_{\mathcal{C}}(t;\boldsymbol{\psi}^*)$ in which the effects of policy are identified. We zoom the identification window in panel (c) of Figure 24 and the associated policy effects in panel (d) of Figure 24.

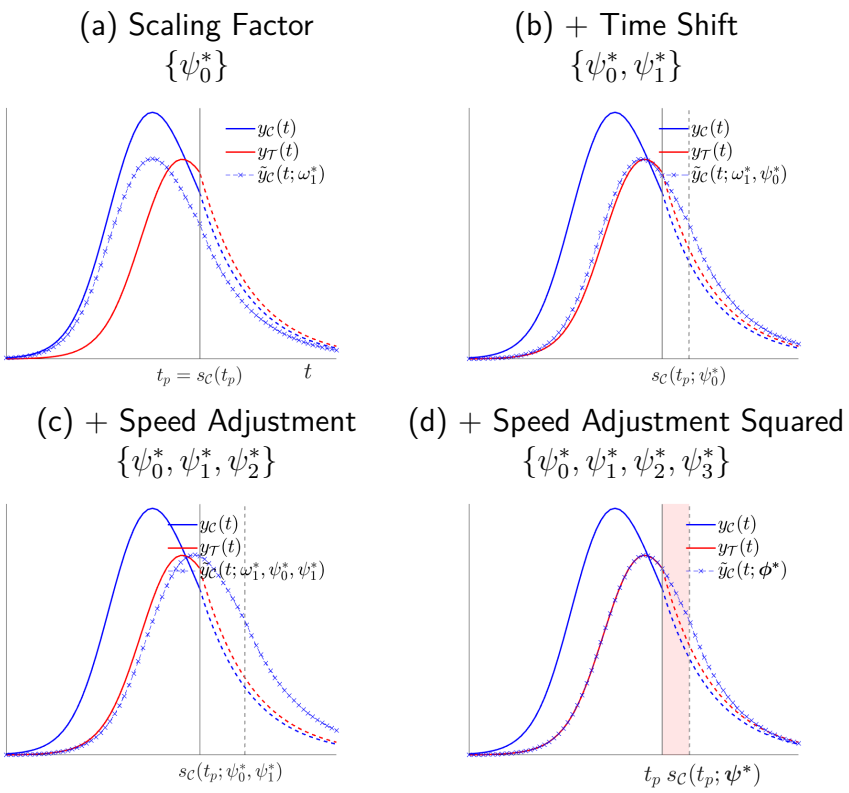
In Figure 25, we further unpack the contribution of each normalization coefficient in generating the normalized path. Note that in this scenario, since the outcome paths of the two regions is already affected by the potential asymmetry in which the outcome paths increase and decrease, the additional parameter ψ_3 that asymmetrically shapes time into stages plays a role.

Figure 24: A Stage-Based Identification of Policy Effects: A Nationwide Policy



Notes: See the notes in Figure 3.

Figure 25: Unpacking the Normalization



Notes: See the notes in Figure 4.

B Analytical Derivations

We show an illustration of these paths for region $\mathcal{C}$ and $\mathcal{T}$ in panel (a) of Figure 6.

B.1 Trigonometric Functions

We choose $y_{t,r} = g(t;\Theta_r) = \theta_{1,r}\sin(\theta_{3,r}t + \theta_{2,r}) + \theta_{0,r}$ with $\Theta_{\mathcal{C}} = \{1, 5, 4.1, 1\}$ and $\Theta_{\mathcal{T}} = \{1.5, 3, 1.1, 1\}$ then,

$$\begin{aligned}
\widetilde{y}_{\mathcal{C}}(s;\boldsymbol{\psi}) &= \omega_1 y_{\mathcal{C}}(\psi_0 + \psi_1 s) + \omega_0 \\
&= \omega_1\left(\theta_{1,\mathcal{C}}\sin(\theta_{3,\mathcal{C}}(\psi_0 + \psi_1 s) + \theta_{2,\mathcal{C}}) + \theta_{0,\mathcal{C}}\right) + \omega_0 \\
&= \underbrace{\omega_1\theta_{1,\mathcal{C}}}_{\theta_{1,\mathcal{T}}}\sin\left(\underbrace{\theta_{3,\mathcal{C}}\psi_1}_{\theta_{3,\mathcal{T}}}s + \underbrace{(\theta_{3,\mathcal{C}}\psi_0 + \theta_{2,\mathcal{C}})}_{\theta_{2,\mathcal{T}}}\right) + \underbrace{\omega_1\theta_{0,\mathcal{C}} + \omega_0}_{\theta_{0,\mathcal{T}}} = y_{\mathcal{T}}(s)
\end{aligned}$$

and hence through undetermined coefficients,

$$\begin{aligned}
\omega_0 &= \theta_{0,\mathcal{T}} - \frac{\theta_{1,\mathcal{T}}}{\theta_{1,\mathcal{C}}}\theta_{0,\mathcal{C}}, \qquad \omega_1 = \frac{\theta_{1,\mathcal{T}}}{\theta_{1,\mathcal{C}}} \\
\psi_0 &= \frac{\theta_{2,\mathcal{T}} - \theta_{2,\mathcal{C}}}{\theta_{3,\mathcal{C}}}, \qquad \psi_1 = \frac{\theta_{3,\mathcal{T}}}{\theta_{3,\mathcal{C}}}
\end{aligned}$$

With $\boldsymbol{\phi} = \{\psi_0, \psi_1, \omega_0, \omega_1\} = \{-4.500, 1.500, -3.000, 1.000\}$

B.2 Polynomial Functions, The Cubic Case

We choose $y_{t,r} = g(t;\Theta_r) = \theta_{0,r} + \theta_{1,r}t + \theta_{2,r}t^2 + \theta_{3,r}t^3$ then,

$$\begin{aligned}
\widetilde{y}_{\mathcal{C}}(t;\boldsymbol{\psi}) &= \omega_1 y_{\mathcal{C}}(\psi_1 + \psi_2 t) + \omega_0 \\
&= \omega_1\left(\alpha_0 + \alpha_1(\psi_1 + \psi_2 t) + \alpha_2(\psi_1 + \psi_2 t)^2 + \alpha_3(\psi_1 + \psi_2 t)^3\right) + \omega_0 \\
&= \omega_0 + \omega_1\left(\alpha_0 + \alpha_1\psi_1 + \alpha_2\psi_1^2 + \alpha_3\psi_1^3\right) \\
&\quad + \omega_1\left(\alpha_1\psi_2 + 2\alpha_2\psi_1\psi_2 + 3\alpha_3\psi_1^2\psi_2\right)t \\
&\quad + \omega_1\left(\alpha_2\psi_2^2 + 3\alpha_3\psi_1\psi_2^2\right)t^2 \\
&\quad + \omega_1\alpha_3\psi_2^3 t^3 = y_{\mathcal{T}}(s)
\end{aligned}$$

This delivers the following 4-equation system with four unknowns ($\boldsymbol{\psi}$):

$$\beta_0 = \beta_0(\omega_0, \omega_1\psi_1) = \omega_0 + \omega_1\left(\alpha_0 + \alpha_1\psi_1 + \alpha_2\psi_1^2 + \alpha_3\psi_1^3\right) \tag{32}$$

$$\beta_1 = \beta_1(\omega_1, \psi_1, \psi_2) = \omega_1\left(\alpha_1\psi_2 + 2\alpha_2\psi_1\psi_2 + 3\alpha_3\psi_1^2\psi_2\right) \tag{33}$$

$$\beta_2 = \beta_2(\omega_1, \psi_1, \psi_2) = \omega_1\left(\alpha_2\psi_2^2 + 3\alpha_3\psi_1\psi_2^2\right) \tag{34}$$

$$\beta_3 = \beta_3(\omega_1, \psi_2) = \omega_1\alpha_3\psi_2^3 \tag{35}$$

Next, we show that this system exactly identified with a unique solution (there is a requirement on the shape of the cubic polynomials for the existence of a solution).

We can solve for $\boldsymbol{\psi}$ in this system in several steps.

STEP 1. Isolate ω_1 in $\beta_3(\omega_1, \psi_2)$,

$$\omega_1 = \omega_1(\psi_2) = \frac{\beta_3}{\alpha_3 \psi_2^3}$$

STEP 2. Plug $\omega_1(\psi_2)$ in $\beta_3(\omega_1, \psi_1, \psi_2)$ and isolate ψ_1:

$$\begin{aligned}
\beta_2 &= \omega_1 \left(\alpha_2 \psi_2^2 + 3\alpha_3 \psi_1 \psi_2^2\right) \\
&= \frac{\beta_3}{\alpha_3 \psi_2^3} \left(\alpha_2 \psi_2^2 + 3\alpha_3 \psi_1 \psi_2^2\right) \\
&= \frac{\beta_3}{\alpha_3 \psi_2} \left(\alpha_2 + 3\alpha_3 \psi_1\right)
\end{aligned}$$

hence,

$$\psi_1(\psi_2) = \frac{1}{3}\frac{\beta_2}{\beta_3}\psi_2 - \frac{1}{3}\frac{\alpha_2}{\alpha_3}$$

STEP 3. Plug $\omega_1(\psi_2)$ and $\psi_1(\psi_2)$ in $\beta_1(\omega_1, \psi_1, \psi_2)$ and isolate ψ_2:

$$\begin{aligned}
\beta_1 &= \omega_1 \left(\alpha_1 \psi_2 + 2\alpha_2 \psi_1 \psi_2 + 3\alpha_3 \psi_1^2 \psi_2\right) \\
&= \frac{\beta_3}{\alpha_3 \psi_2^3} \left(\alpha_1 \psi_2 + 2\alpha_2 \psi_1 \psi_2 + 3\alpha_3 \psi_1^2 \psi_2\right) \\
&= \frac{\beta_3}{\alpha_3 \psi_2^2} \left(\alpha_1 + 2\alpha_2 \psi_1 + 3\alpha_3 \psi_1^2\right) \\
&= \frac{\beta_3}{\alpha_3 \psi_2^2} \left(\alpha_2 + 2\alpha_3 \left(\underbrace{\frac{1}{3}\frac{\beta_3}{\beta_3}\psi_2 - \frac{1}{3}\frac{\alpha_3}{\alpha_3}}_{\psi_1(\psi_2)}\right) + 3\alpha_3 \left(\underbrace{\frac{1}{3}\frac{\beta_3}{\beta_3}\psi_2 - \frac{1}{3}\frac{\alpha_3}{\alpha_3}}_{\psi_1(\psi_2)}\right)^2\right) \\
&= \frac{\beta_3}{\alpha_3 \psi_2^2} \left(\alpha_1 + 2\alpha_2 \left(\frac{1}{3}\frac{\beta_2}{\beta_3}\psi_2 - \frac{1}{3}\frac{\alpha_2}{\alpha_3}\right) + 3\alpha_3 \left(\frac{1}{9}\left(\frac{\beta_2}{\beta_3}\right)^2 \psi_2^2 + \frac{1}{9}\left(\frac{\alpha_2}{\alpha_3}\right)^2 - 2\frac{1}{9}\frac{\beta_2}{\beta_3}\frac{\alpha_2}{\alpha_3}\psi_2\right)\right) \\
&= \frac{1}{3}\frac{\beta_2^2}{\beta_3} + \frac{2}{3}\left(\frac{\alpha_2}{\beta_3}\beta_2 - \frac{\beta_2}{\beta_3}\alpha_2\right)\frac{1}{\psi_2} + \frac{\beta_3}{\alpha_3}\left(\alpha_1 - \frac{1}{3}\frac{\alpha_2^2}{\alpha_3}\right)\frac{1}{\psi_2^2}
\end{aligned}$$

That is,

$$\underbrace{\frac{\beta_3}{\alpha_3}\left(\alpha_1 - \frac{1}{3}\frac{\alpha_2^2}{\alpha_3}\right)}_{a}\frac{1}{\psi_2^2} + \underbrace{\frac{2}{3}\left(\frac{\alpha_2}{\beta_3}\beta_2 - \frac{\beta_2}{\beta_3}\alpha_2\right)}_{b=0}\frac{1}{\psi_2} + \underbrace{\frac{1}{3}\frac{\beta_2^2}{\beta_3} - \beta_1}_{c} = 0$$

where note that $b = 0$. Thus, we can isolate $\psi_2 = \sqrt{\frac{a}{-c}}$ which delivers a unique solution for ψ_2 if $\frac{a}{-c} > 0$ (and no solution otherwise).[55]

STEP 4. Plug ψ_2 into $\omega_1(\psi_2)$ and $\psi_1(\psi_2)$ in order to recover ω_1 and ψ_1.

STEP 5. Plug ω_1, ψ_1 and ψ_2 into $\beta_0(\omega_0, \omega_1, \psi_1)$ and isolate (recover) ω_0.

The solution for the case where $\Theta_{\mathcal{C}} = \{12, -4.5, 0.3, 0.2\}$ and $\Theta_{\mathcal{T}} = \{7.84, -1.28, -0.17, 0.07\}$ the is $\boldsymbol{\phi} = \{\psi_0, \psi_1, \omega_0, \omega_1\} = \{2.788, 0.278, -1.350, 1.093\}$

[55] One can further elaborate this to show that $\frac{a}{-c} = \frac{(3\alpha_1\alpha_3 - \alpha_2^2)/\alpha_3^2}{(3\beta_1\beta_3 - \beta_2^2)/\beta_3^2}$ and hence the solution to the system (32)-(35) emerging from the cubic time paths exists as long as the term in the numerator $(3\alpha_1\alpha_3 - \alpha_2^2)$ and the term in the denominator $(3\beta_1\beta_3 - \beta_2^2)$ have the same sign.

C Solution Algorithms for Theoretical Frameworks

C.1 The Epi-Econ Model

We separately solve for the pre-pandemic equilibrium at $t = 0$ (actually, for any $t \leq 0$) before the unexpected arrival of the pandemic at $t = 1$. In this pre-pandemic era there are no infections and, hence, $\phi_i(h_0) = 1$. That is, the equilibrium labor supply sets the right-hand side of the Euler equation (15) to zero in which case h_0 simply solves an intra-temporal trade-off. The same equilibrium emerges after the pandemic at some large $t = T$ which delivers a terminal condition $h_T = h_0$.

Step 1. Solve for hours worked in the pre-pandemic steady state ($\bar{h}$).

Step 2. Select the number of periods to simulate $\mathfrak{T}$. Pick a large number T. Set $h_0 = h_T = \bar{h}$.

Step 3. Given parameters Θ, guess a sequence $\{h_t\}_{t=0}^{T}$.

Step 4. With $\{h_t\}_{t=0}^{T}$, compute sequences for S_t, I_t, R_t, D_t, N_t.

Step 5. Use the above sequences to back out a new sequence for $\{h_t\}_{t=0}^{T}$ using (15). Solve backwards.

Step 6. If the new sequence $\{h_t\}_{t=0}^{T}$ is different than the guess in Step 3, update the guess and go back to Step 4.

Step 7. Store the second value of the sequence, namely h_1, and set $h_0 = h_1$. Go back to Step 3. Repeat this step $\mathfrak{T}$ times. Construct a solution sequence $\{h_{sol,t}\}_{t=1}^{\mathfrak{T}}$ using all values stored. Simulate the underlying epidemic dynamics associated to the solution sequence.

Step 8. To obtain the series with the effect of policy: set all values of the solution sequence after t_p equal to h_{pol}, and simulate the epidemic dynamics.

Special case, endogenous policy: In this case, the policy hits when the cumulative number of deaths reaches a certain number $\overline{D}$. In Step 5 solve backwards only from the date in which policy is implemented, continue with the rest of the steps as described above.
To obtain the series without policy, use as initial values those immediately before the policy hits, go to Step 3 and solve for a new sequence of h assuming the policy constrain is never binding.

C.2 Growth Policy and Structural Transformation

We solve the economy by guessing the sequences of factor prices $\{w, r\}_{t=0}^{\infty}$ with $w_t = w_{at} = w_{mt}$. Given these prices, we find the allocations c_{at}, k_{t+1} and n_{at} that solve the set of first order conditions (24)-(26) with $p_{at} = \frac{w_t}{\phi z_{at}} \left(\frac{n_{at}}{\ell}\right)^{1-\phi}$.[56] There is market clearing in labor and capital, and aggregate consistency. Note that the interemporal Euler condition (25) is a second order different equation in $\{k_t, k_{t+1}, k_{t+2}\}$ at every period t. We use as initial and terminal conditions the corresponding stationarized economies at $t = 0$ and at a large T with negligible agricultural share of labor.

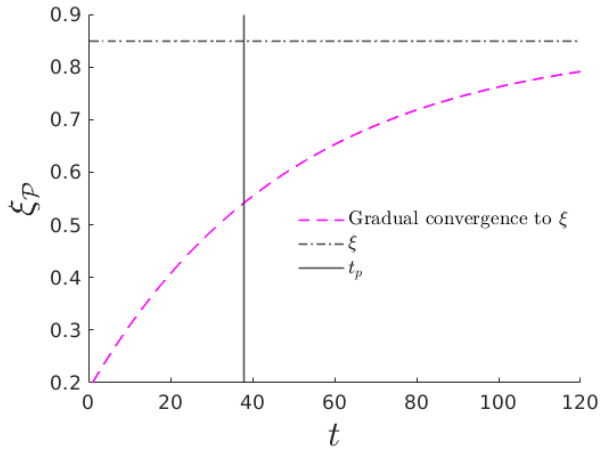
[56] Note that without the distortion τ, if $\phi = \alpha$ and if we had the same factor inputs in the production of both goods, then the equality of the ratio of factor input prices across sectors would imply that the ratio of factor inputs must be identical across sectors. In turn, this would imply a standard result for the pricing of agricultural goods, $p_{at} = \frac{z_{mt}}{z_{at}}$, which renders the price of agricultural good as exogenous. The fact that we allow for ϕ to differ from α and that we have different factor inputs differ across sectors both prevents the standard result. Indeed, in our case, the price of agricultural goods depends endogenously on n_{at}.

D More on Confounding Factors

D.1 More on Time-Varying Latent Heterogeneity

In Figure 26, we show the assumed path for the beliefs on $\xi_{\mathcal{P}}$ in region $\mathcal{T}$ that generates the time-varying latent heterogeneity studied in Section **??**. That is, we assumed that beliefs on the probability of infection conditional on hours work $\xi_{\mathcal{P}}$ (magenta dashed line) converge from below to the true probabilities ξ (dashed gray line). That is, a structural parameter $\xi_{\mathcal{P}}$ that is unobserved to the policy evaluator evolves over time before and after policy implementation.

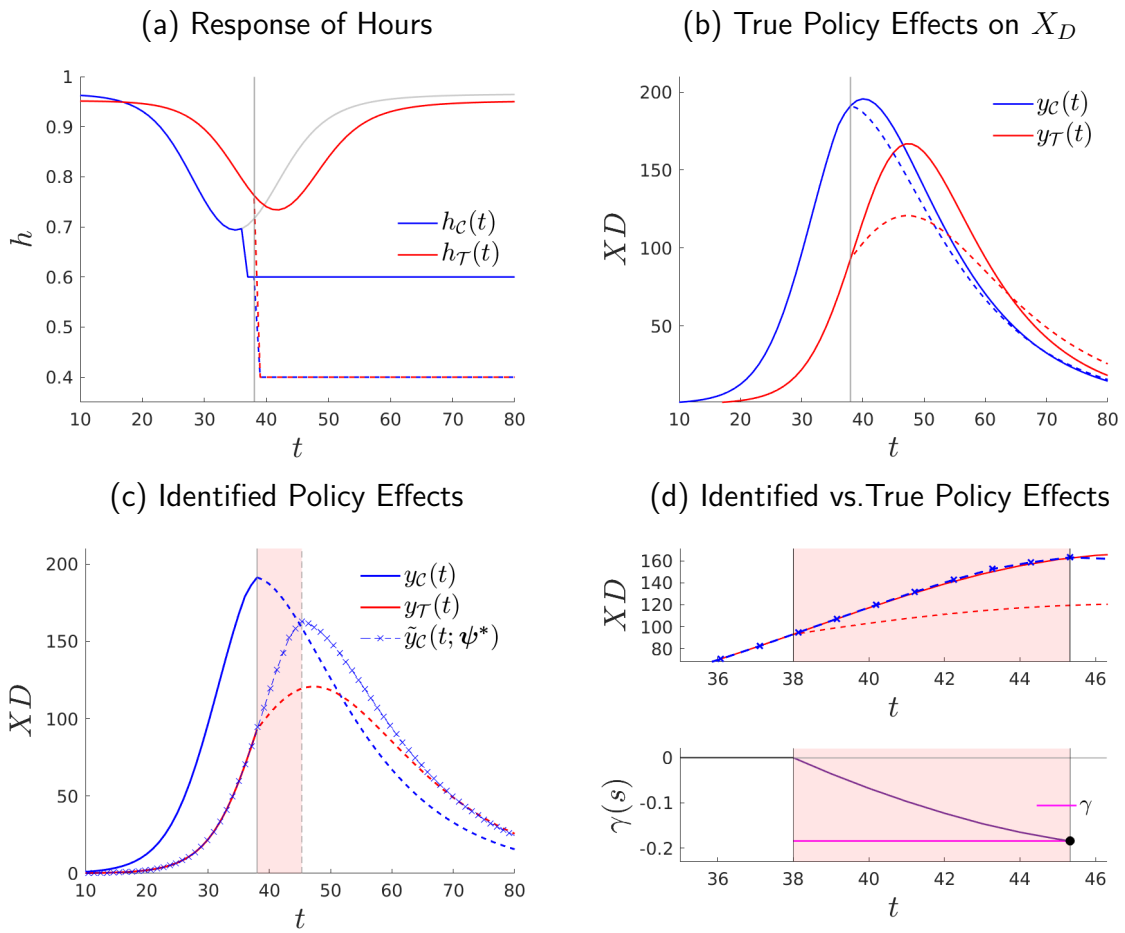
Figure 26: Time-Varying Latent Heterogeneity: Assumed Path of Beliefs $\xi_{\mathcal{P}}$ for region $\mathcal{T}$



D.2 More on Confounding Policies

Here, we show how our method performs where the confounding policy that happens before the stay-home policy is implemented occurs in region $\mathcal{C}$ and not in region $\mathcal{C}$, i.e. the opposite case studied in our Section **??**. In this case, we find that the policy effect is recovered with an error of 2.34%. Again, we can make this error larger if the confounding policy drives the outcome path of region $\mathcal{C}$ further away from the outcome path of regin $\mathcal{T}$.

Figure 27: Stage-Based Identification of Policy Effects: With Confounding Policy in $\mathcal{C}$

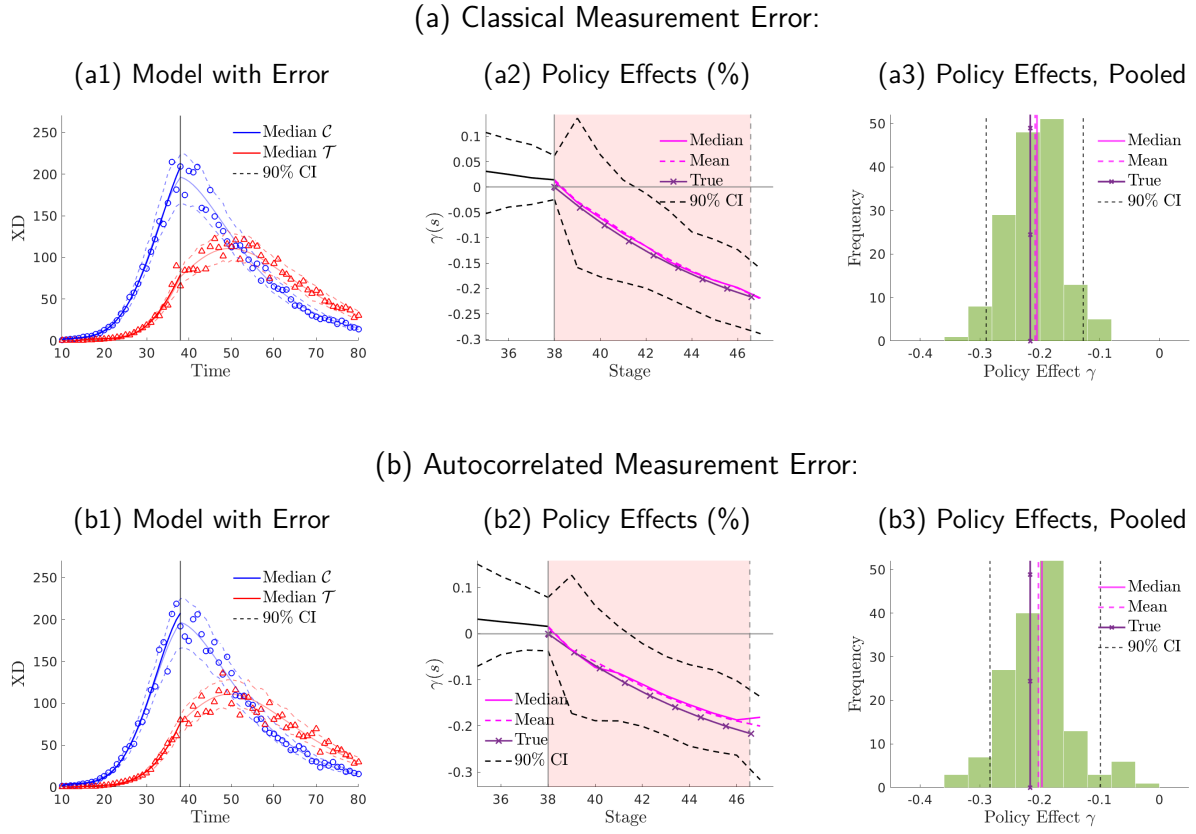


Notes: Where $\bar{h}_1 = 0.6$ at $t = 37$ in $\mathcal{C}$, his policy is unobserved, $\bar{h}_2 = 0.4$ (lockdown) at $t_p = 38, t_f = 250$, $\gamma = -18.34\%, \epsilon(\gamma) = 2.34\%$.

E Further Details on Inference

An altogether alternative way to conduct inference with the recovered estimates for the error terms $u_r(t)$ is to estimate the sample variance of the errors, i.e. $\widehat{\sigma}_r$. Then, under a normality assumption on the error term in (27), we simulate $Q = 1,000$ paths of errors and, hence, the same number of pre-policy outcome paths onto which we apply the smoothing step in order to recover a simulation-specific estimand $\widehat{\widehat{y}}_{r,q}(t)$. Since the estimand $\widehat{\widehat{y}}_{r,q}(t)$ differs by simulation $q \in Q$, each simulation delivers an stage-based identified policy effect, γ_q. We show the results of this different inference in Figure 28. Overall, we find similar insights with an identified mean policy effect of 21.12% [14.41,28.23] and 19.61% [12.32,29.35] with classical ME and with auto-correlated ME, respectively. The recovered policy effect is not significantly different from the true (model-generated) policy effect. In Figure 29, we show the policy effects from directly using the $Q = 1,000$ simulations of data $\widehat{y}_{r,q}(t)$, that is, without applying the smoothing step. The identified mean policy effect obtained without the smoothing step is 20.92% [7.88,30.32] and 21.55% [7.51,30.81] with classical ME and with auto-correlated ME respectively.

Figure 28: Stage-Based Identification of Model-Generated Policy Effects: Inference

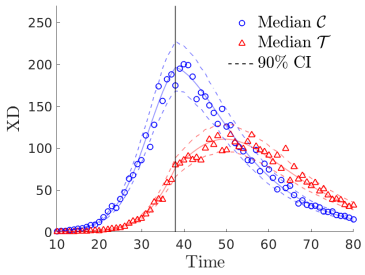


Notes: We use the benchmark calibration in Section 3.1.1. the top panels (a), we introduce classical measurement error in our model with $\{\sigma^2_{\mathcal{C}}, \sigma^2_{\mathcal{T}}\} = \{0.01, 0.01\}$. In the bottom panels (b), we introduce non-classical measurement error with $\{\rho_{\mathcal{C}}, \rho_{\mathcal{T}}\} = \{0.13, 0.13\}$ and $\{\sigma^2_{\mathcal{C}}, \sigma^2_{\mathcal{T}}\} = \{0.01, 0.01\}$.

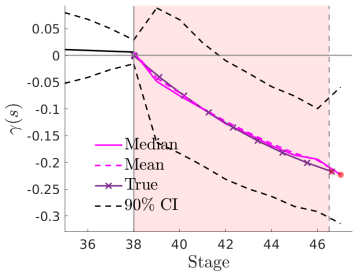
Figure 29: Stage-Based Identification of Model-Generated Policy Effects: Inference, No Smoother

(a) Classical Measurement Error:

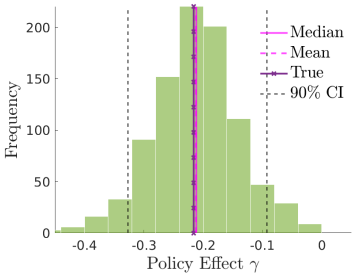
(a1) Model with Error



(a2) Policy Effects (%)

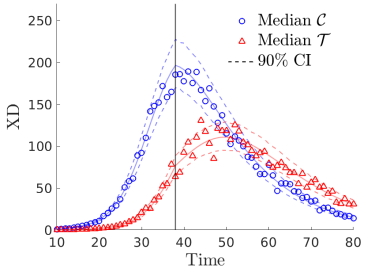


(a3) Policy Effects, Pooled

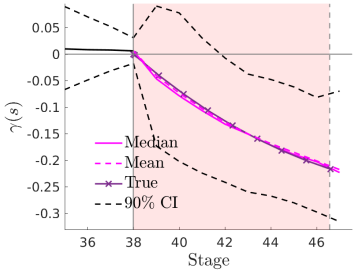


(b) Autocorrelated Measurement Error:

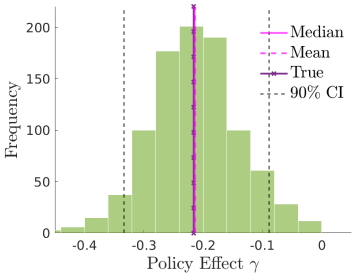
(b1) Model with Error



(b2) Policy Effects (%)



(b3) Policy Effects, Pooled



Notes: We use the benchmark calibration in Section 3.1.1. the top panels (a), we introduce classical measurement error in our model with $\{\sigma^2_{\mathcal{C}}, \sigma^2_{\mathcal{T}}\} = \{0.008, 0.008\}$. In the bottom panels (b), we introduce non-classical measurement error with $\{\rho_{\mathcal{C}}, \rho_{\mathcal{T}}\} = \{0.13, 0.13\}$ and $\{\sigma^2_{\mathcal{C}}, \sigma^2_{\mathcal{T}}\} = \{0.008, 0.008\}$.