Criminal Instruments

Shift-Share Designs and Collinearity

Enda Patrick Hargaden
María Padilla-Romo

July 2020

Abstract

This paper provides evidence that machine learning techniques can improve Bartik instrument research designs. We show that collinearity in a Bartik first-stage can render point estimates uninterpretable as a local average treatment effect (LATE). In addition to providing more precise estimates we show that addressing collinearity—for example, with Lasso—can also relax the assumptions needed for interpreting those estimates as a LATE. Our empirical focus is the effect of unemployment on crime. Compared to a baseline Bartik specification, a Lasso-style first-stage reduces the estimated effect of job losses on crime by forty-five percent.

Keywords: crime, unemployment, Bartik instruments, Ireland, Lasso, Rotemberg weights.
JEL classification: C36, J63, K42.

Hargaden: University of Tennessee, 702 Stokely Management Center, Knoxville, TN 37996-0570, USA. Email: enda@utk.edu. Padilla-Romo: University of Tennessee, 514 Stokely Management Center, Knoxville, TN 37996-0570, USA. Email: mpadill3@utk.edu.

1 Introduction

The shift-share instrument is a popular research design in applied econometrics. Attributable to Bartik (1991), the instrument uses pre-period sectoral shares to predict subsequent labor market conditions. Regions vary in their sectoral compositions, and using this differential exposure to common shocks makes the Bartik-style instrument widely applicable. Beyond the economics of crime (e.g. Gould et al., 2002; Aizer, 2010), the approach has been used to study international trade (Autor, Dorn and Hanson, 2013); behavioral economics (Clingingsmith, 2016); technological change (Hershbein and Kahn, 2018); location choice (Wozniak, 2010; Diamond, 2016); long-run population growth (Hanlon, 2017); political preferences (Brunner, Ross and Washington, 2011); and much more.

The first-stage of a Bartik instrument relies on the initial employment shares of many sectors, often correlated with one another. Collinearity occurs when the correlation ρ between independent variables is close to but less than one.[1] Standard estimators remain consistent in this scenario, but their variance can explode as $\rho \to 1$. Point estimates become unstable, and can vary wildly across specifications. We show that collinearity in a Bartik first-stage can render point estimates uninterpretable as a local average treatment effect (LATE), and propose a method to address this.

The methodology of this paper lies at the intersection of optimal instrumental variables and shift-share designs. We build on a result in Goldsmith-Pinkham, Sorkin and Swift (2020) that decomposes the Bartik instrument into a set of parameters ("Rotemberg weights").[2] If any of the weights are negative, which they often are, the parameter estimate does not in general permit a LATE interpretation. We show that Rotemberg weights are affected by collinearity in the first-stage. In particular, failure to tackle collinearity can result in negative Rotemberg weights. This itself is not a critique of Rotemberg weights any more than collinearity is a critique of OLS. Rather, we suggest that Rotemberg weights can be used to illuminate an underlying sensitivity of Bartik research designs to collinearity.

We propose machine-learning as a fix to the collinearity problem. Applying a Lasso-style estimator to a Bartik instrument can help ensure uniformly positive Rotemberg weights by addressing collinearity in the first-stage.[3]

[1] In the case of perfect collinearity ($\rho = 1$) the data matrix is singular (non-invertible) and standard statistical objects like $(X'X)^{-1}$ are not well-defined.

[2] Named in homage to Rotemberg (1983), these weights indicate which industries are driving the overall estimate (20% weight on furniture manufacturing, 15% weight for computer parts, and so on). Goldsmith-Pinkham et al. (2020) show that an overall Bartik instrument is numerically equivalent to identification from the industry shares interacted with national growth rates as a weighting matrix. Focusing on these shares as exogenous, they decompose the overall estimate into Rotemberg weights for each sector. Borusyak, Hull and Jaravel (2018) show the Bartik instrument can alternatively be consistently identified off the shocks themselves. Goldsmith-Pinkham et al. (2020) argue the former interpretation is more appropriate if researchers emphasize shocks to specific industries as central to the research design.

[3] We specify Lasso-*style* estimator to highlight that while we lean on the consistency results of Belloni, Chen,

The Lasso estimator is a simple machine-learning technique that is popular in economics (Tibshirani, 1996; Athey and Imbens, 2019). As the Lasso estimator tends to omit/drop statistically irrelevant variables, it can improve first-stage precision by selecting an optimal subset of potential instruments (Belloni et al., 2012). The Lasso can also be used to address collinearity (see e.g. Dormann et al., 2013). By mitigating collinearity the Lasso tends to alleviate the problems associated with Rotemberg weights.

We demonstrate this finding by applying a Lasso-optimized Bartik instrumental variable design to three canonical settings.[4] We show the approach substantially reduces the extent of negative Rotemberg weights. For example, Autor et al. (2013) harness data on 395 manufacturing sub-industries to generate their Bartik-style instrument, and forty percent of the sub-industries have negative Rotemberg weights.[5] We find that a Lasso-optimized Bartik shrinks the first-stage to nineteen sub-industries, all of which have positive Rotemberg weights. Uniformly positive Rotemberg weights ensure the IV results can be interpreted as a LATE in less restrictive conditions.[6] Thus applying a Lasso approach can lend itself to permitting a LATE interpretation.

The econometric assumption underpinning the Lasso-optimized design ("approximate sparsity") is most credible in settings where the first-stage is not implied by theory. Approximate sparsity requires that the conditional expectation of the endogenous variable can be well-approximated by a subset of the potential variables. In concrete terms, this means the researcher is primarily interested in predicting the endogenous variable (e.g. unemployment) and does not have a strong *a priori* theory/model on which specific instruments (e.g. sectors) best achieves that.

The effect of unemployment on crime is one such setting. We analyze this relationship with administrative data from Ireland. The labor market effects of the Great Recession were particularly severe in Ireland: unemployment rose by more than ten percentage points. A baseline Bartik instrument finds that one thousand additional unemployed people led to 106 additional crimes per quarter, with large standard errors. Rotemberg weights are negative and substantive for major 2-digit industries like Construction. This implies the point estimate can only be interpreted as a LATE under an additional assumption of homogeneous effects

Chernozhukov and Hansen (2012), alternative strategies to address collinearity could also be implemented.

[4]The three settings are estimating the inverse elasticity of labor supply in the US; the effect of trade on unemployment as per Autor, Dorn and Hanson (2013); and the inverse elasticity of substitution between immigrants and natives as per Card (2009). These examples follow Goldsmith-Pinkham et al. (2020). We thank Paul Goldsmith-Pinkham for providing replication files.

[5]We highlight the Autor et al. (2013) paper because of their data. Our empirical analysis of unemployment on crime has a baseline specification with fourteen sectors, so applying our research design to a dataset with 395 sub-industries is illuminating. There are fine justifications for including all sub-industries (e.g. if doing so is suggested by theory), and so our application should not be considered a refutation of Autor et al. (2013).

[6]Without uniformly positive Rotemberg weights, any treatment effect heterogeneity means the results are not a LATE in general (Goldsmith-Pinkham et al., 2020). Treatment effect homogeneity is a sufficient condition to interpret a Bartik estimate as a LATE.

across industries, which is questionable.

In contrast, the Lasso Bartik shrinks the first-stage to focus on four industries, with much of the weight on Information, Communication, Technology (ICT) and the Professional/Scientific sectors. Due to Ireland's low corporate tax rate, employment in these sectors is likely driven by international trends. This means the ICT and Professional/Scientific sectors are less sensitive to local conditions, and good candidate sectors for exogenous shocks to local unemployment.

Focusing on the sectors chosen by the Lasso substantially changes our estimates. The point estimate with this approach is 47% lower than the baseline, namely a quarterly increase of about 56 additional property crimes (theft, burglary, and robbery) per thousand people unemployed. The estimate is smaller, more precisely estimated, and fully consistent with a LATE interpretation. The policy advice stemming from the analysis thus depends on the chosen specification, and we believe there is good reason to prefer the Lasso specification.

While the focus of this paper is crime in Ireland, the empirical application can be considered illustrative. We believe the paper's contributions are of considerably broader interest. The Bartik shift-share instrument is a popular research design to extract causal evidence from observational data. While popular, recent work has queried the use of Bartik instruments.[7] Our finding that collinearity can make Rotemberg weights unstable builds on this new literature. This paper's result that the Lasso can alleviate the shift-share instrument's instability provides an avenue for researchers to improve future work using the shift-share design.

When faced with collinearity, textbooks encourage researchers to omit variables that generate excessive instability.[8] The standard Bartik instrument includes information from all available sectors. Inclusion of all sectors is a corner solution in a bias-variance tradeoff problem, and thus not necessarily optimal in general. The Lasso can scale down the number of selected variables in a pre-determined, hands-off way. To this insight, we add that Lasso is also used as a technique to address collinearity. In turn, optimizing a Bartik first-stage can ensure the Rotemberg weights are more robust and less sensitive to minor tweaks in the specification. This reduces the risk of collinearity-induced variance generating negative Rotemberg weights.

In addition to our headline results we also show cases where the Lasso Bartik approach fails to return uniformly positive Rotemberg weights. Thus we do not suggest that applying

[7] Goldsmith-Pinkham et al. (2020) is perhaps the most prominent paper in this vein. Other notable candidates for this title include Adao, Kolesár and Morales (2019) and Borusyak, Hull and Jaravel (2018).

[8] For example, Wooldridge (2006) suggests "we can try dropping other independent variables in an effort to reduce multicollinearity" but notes that OLS parameter estimates will be biased if we misspecify the model. More relevant to the case of Bartik instruments is Cameron and Trivedi (2005), who note that "the marginal value of an [additional] instrument may be very slight, because of increasing multicollinearity among the instruments, leading to a situation of weak instruments."

Lasso to the first-stage is a panacea to all identification and inference concerns about the Bartik instrument. Instead, we believe that the Lasso Bartik can improve the precision of estimates and lend itself to a credible interpretation. We show that the approach is an additional tool for the applied researcher.

2 Shift-share designs, Rotemberg weights, and collinearity

One contribution of this paper is applying a link between the optimal/sparse instrumental variable approach of Belloni et al. (2012) and the Rotemberg decomposition of Goldsmith-Pinkham et al. (2020). Specifically we note that collinearity-induced variance can result in negative Rotemberg weights, and that the approaches of Belloni et al. (2012) can be used to address this.

The starting-point of this analysis is the Bartik (1991)-style shift-share instrument. This is a popular research design for parameter identification using observational data. Consider the effect of local unemployment on crime. A researcher posits the following relationship:

$$Y_{rt} = \alpha_r + \gamma_t + \beta X_{rt} + C'_{rt}\delta + u_{rt} \tag{1}$$

where Y_{rt} represents crime in region r at time t; α_r are region fixed effects; γ_t are time fixed effects; X_{rt} is the number of unemployed people; C_{rt} is a set of control variables (e.g. region-specific time trends, controls for the share of young men, etc.); and u_{rt} is an error term. The parameter of interest is β and it measures the average effect of unemployment on crime.

As the error term is likely correlated with the unemployment rate, the estimate of β in this model will be inconsistent. The Bartik instrument predicts unemployment X_{rt} using pre-period regional industrial shares Z_{irt} and the growth rate of the sectors g_{irt} where subscript i stands for industry/sector i. Interacting the shares and the growth rates generates the Bartik instrument's predicted unemployment, B_{rt}. In practice we calculate a leave-one-out growth rate, where g_{irt} equals the average time-t growth rate of industry i in all regions excluding region r. Depending on the relevant geographic unit (e.g. nine U.S. Census divisions versus 709 commuting zones) the leave-one-out distinction can be quantitatively important. Identification can come from either the shocks (Borusyak et al., 2018) or from the shares (Goldsmith-Pinkham et al., 2020). In either case the intuition is that a shock to the automobile industry will affect Detroit more than Denver.

Belloni et al. (2012) provide theoretical foundations for using a Lasso-based IV. The key assumption for a Lasso-based procedure to provide an optimal instrument is *approximate sparsity*: that there exists a small set k of important instruments from the K potential instruments ($k < K$) that well-approximate the conditional expectation of the endogenous variable. In the Bartik setting, the relevant assumption is that the unemployment rate can be

well-approximated with a few sectors. Lasso penalizes the inclusion of variables which do not add enough predictive power. More formally, consider that OLS minimizes $\sum(Y_i - \beta' X_i)^2$ by choice of β. The Lasso alters this, placing a penalty on non-zero β, namely Lasso minimizes $\sum(Y_i - \beta' X_i)^2 + \lambda||\beta||_2^2$ by choice of β. The parameter λ is a subjective penalty parameter, a shadow cost of including additional variables. Consequently the Lasso tends to reduce, or shrinks, the number of included variables in the first-stage. Note that OLS, by including all possible variables, is a corner solution to this problem. Through cross-validation we can find an optimal λ that minimizes the out-of-sample prediction error (Athey and Imbens, 2019).

Belloni et al. (2012)'s contribution was to formalize the statistical properties of the optimal first-stage, showing that any optimal algorithm must also incorporate information from the second-stage, and including a data-driven method for optimal penalization. There have been extensions of this result such as Belloni, Chernozhukov, Hansen and Kozbur (2016) which extends the setup to clustered panels, as in our case. We rely on Ahrens, Hansen and Schaffer (2019)'s computational implementation, specifically their addition of the *ivlasso* command to Stata.

Goldsmith-Pinkham et al. (2020) is quite separate from the optimal IV literature. In their words, Goldsmith-Pinkham et al. (2020) "open the black box of the Bartik instrument." The authors derive what they term Rotemberg weights, unpacking the variation driving the Bartik instrument. Let ρ denote the Rotemberg weight for an industry with associated growth rate g. With X being unemployment (or similar) and Z being industry shares (or similar), and using single cross-section to avoid time subscripts, the Rotemberg weight for industry j equals

$$\rho_j = \frac{g_j Z_j' X^{\perp}}{\sum_i g_i Z_i' X^{\perp}}$$

where $\perp$ indicates an X Frisch-Waugh-Lovell (FWL)-residualized from controls and fixed effects. As the individual Rotemberg weights sum to one, removing an industry i from the analysis will mechanically change the remaining weights when renormalized. Our focus is not on this mechanical effect, but the post-FWL effect of removing industry i on weight j due to their collinearity.

To that end, more relevant to our consideration is the derivation in Appendix F of Goldsmith-Pinkham et al. (2020). The square of the Rotemberg weight for industry j is proportional to $\frac{\hat{\sum}_{\pi_j \pi_j}}{\hat{\sum}_{\pi\pi}}$ where $\hat{\sum}_{\pi_j \pi_j}$ is the estimated sampling variance on the first-stage coefficient on industry j and $\hat{\sum}_{\pi\pi}$ is the estimated sampling variance on the first-stage coefficient on the overall instrument. Similarly define F_j as the F-statistic of the jth instrument and F as the first-stage F-statistic for the overall Bartik, and let B denote the industry shares multiplied by the growth rates. Then

$$\frac{F_j}{F} = \rho_j^2 \left(\frac{\widehat{Var}(B^{\perp})}{g_j \widehat{Var}(Z_j^{\perp})} \right)^2 \frac{\hat{\sum}_{\pi\pi}}{\hat{\sum}_{\pi_j \pi_j}}$$

which defines the relationship between an industry's F statistic and its Rotemberg weight. The two will cohere (higher F leading to higher ρ, etc.) except where the precision of the first-stage coefficient for industry j is not proportional the variance of the overall instrument. In the presence of highly correlated instruments, minor changes in the first-stage specification (using a different base year, say) can have large effects on variances even if the change in coefficient is zero in expectation. Equally, removing a high-variance collinear industry that has little predictive power can affect the precision of the estimates for other industries, and thus change their Rotemberg weights.

Collinearity is a relatively mundane statistical feature and it lacks a rigorous statistical definition. It is an applied problem, and one which is more likely to occur as the number of independent variables increases. This is particularly relevant when we "open up" a Bartik instrument and consider which industries are driving the estimates. We argue that collinearity is an under-explored element of Goldsmith-Pinkham et al. (2020). While the authors derive the above relationship between the Rotemberg weight for industry i and its first-stage F statistic, and note a tendency for the two to move together, they are relatively silent on how collinearity could be responsible for departures from this tendency.

Data science researchers have found that the Lasso estimator can help address collinearity. Dormann et al. (2013) compare a suite of different methods to deal with collinearity. The authors find that Lasso consistently produces a low Root Mean Square Error (RMSE) in the presence of highly correlated explanatory variables. Using Lasso as a tool in data-driven variable selection has been used in applied problems, e.g. Munyo and Rossi (2015); de Bromhead, Fernihough and Hargaden (2020).

Taking *(i)* the sensitivity of Rotemberg weights to collinearity, with *(ii)* the finding that the Lasso estimator can address collinearity, and *(iii)* the benefits of using Lasso-based methods for optimal IVs as per Belloni et al. (2012), we have an algorithm for improving the credibility of shift-share instruments. The next section demonstrates this algorithm. We show how applying a Lasso estimator to the first-stage of Autor, Dorn and Hanson (2013) selects industries with uniformly positive Rotemberg weights.

2.1 Application to Autor, Dorn, Hanson (2013)

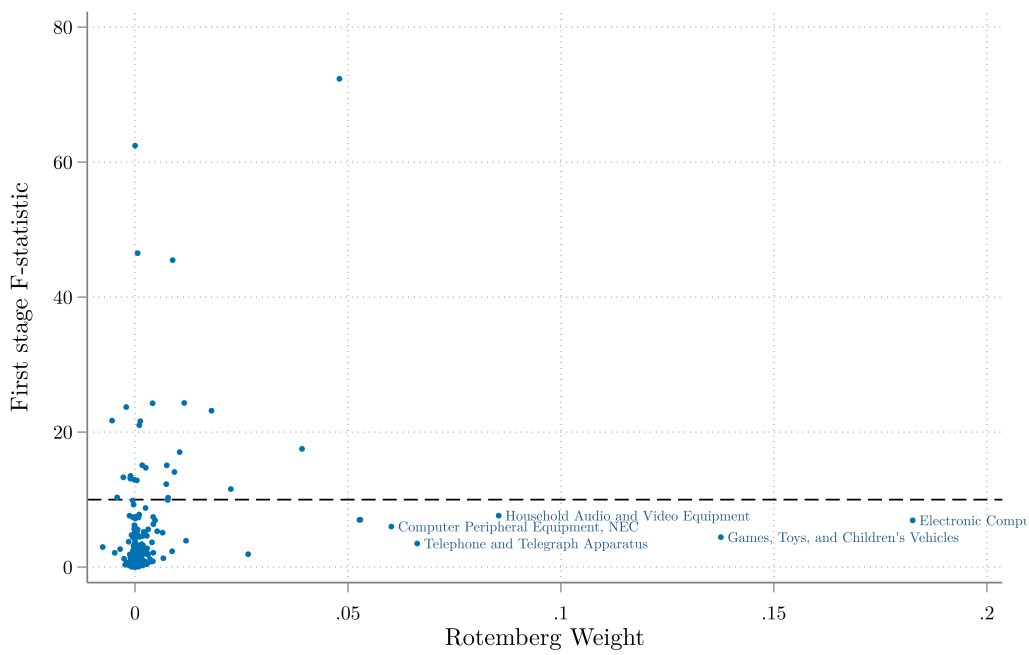
A seminal paper in the trade literature, Autor, Dorn and Hanson (2013) estimates the effect of Chinese trade shocks on US labor markets. Henceforth we refer this paper as "ADH". China's accession to the WTO introduced new competitors for American firms in the tradable goods sector. There is substantive theoretical motivation for the empirical specification ADH

employ. The underlying structure is a gravity model that maps imports to labor market outcomes like wages and unemployment.

While the incidence of the trade shock in affected industries is an empirical question, we note that a theoretically-motivated specification may not be an ideal setting for the approximate sparsity assumption. The Lasso Bartik is more appropriate for settings without theory-heavy underpinnings like ADH. Consequently this application should be viewed as an empirical illustration rather than a robust replication, and certainly neither a confirmation nor refutation.

We show the results of applying the optimal-first stage approach to the ADH data below. The baseline non-Lasso Bartik uses 395 sub-industries to generate predictions for exposure to the trade shock. Figure 1 depicts the distribution of Rotemberg weights and F-statistics for these 395 industries.

Figure 1: Rotemberg weights for the full Autor, Dorn, Hanson (2013) dataset



The Rotemberg weight analysis shows that five industries (electronic computers, games and toys, household audio and video, telephone apparatus, and computer equipment) comprise more than half of the absolute weight of the estimator. In this sense, most of the ADH results are driven by variation in these five industries. While these industries attract large Rotemberg weights, they have relatively modest first-stage power; zero satisfy the $F \geq 10$ rule of thumb.

A secondary finding is that 154 (39%) of the industries have negative associated Rotem-

berg weights. This finding comes with the caveat that although nearly forty percent of the individual industries have negative Rotemberg weights, most of these weights are quite small. While the cumulative weight attributable to these industries are relatively innocuous, negative Rotermberg weights mean point estimates only permit a LATE interpretation under a further assumption of homogenous effects across industries.

While there are 395 candidate industries for the ADH instrument, including them all may not be optimal. Applying the Lasso-optimized Bartik to the ADH instrument trims the number of included industries down to nineteen. Table 1 provides the details of these industries.

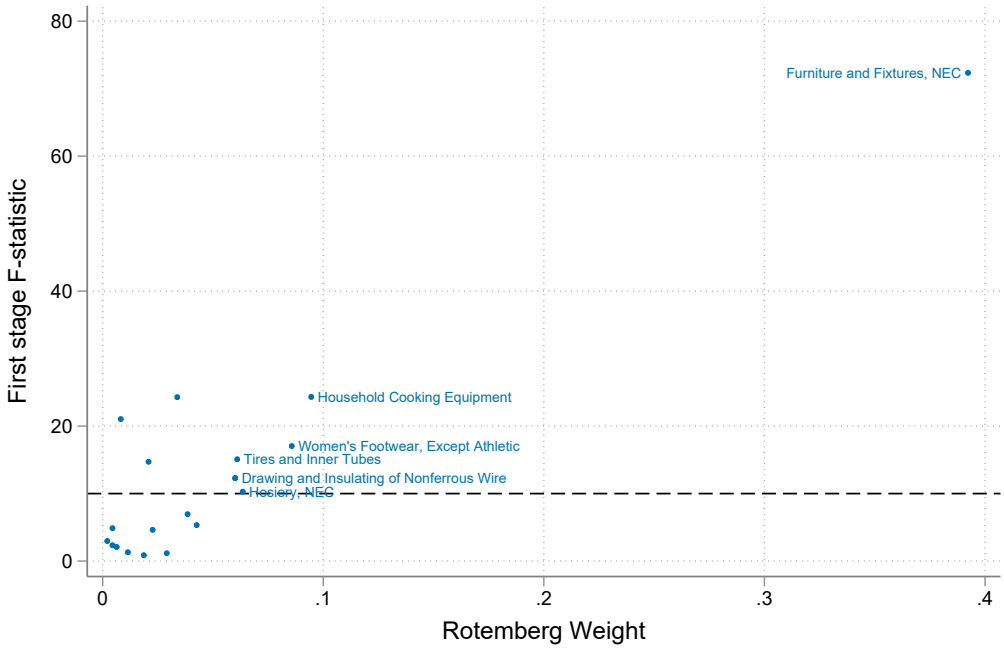
Table 1: Details of the nineteen sectors remaining from the post-Lasso specification

Industry Name	***F*-stat**	**Rotemberg Weight**	**Cum. Weight**
Furniture and Fixtures, NEC	72.34	0.39	0.39
Household Cooking Equipment	24.32	0.09	0.49
Women's Footwear, Except Athletic	17.04	0.09	0.57
Hosiery, NEC	10.29	0.06	0.64
Tires and Inner Tubes	15.08	0.06	0.70
Drawing and Insulating of Nonferrous Wire	12.29	0.06	0.76
Motors and Generators	5.32	0.04	0.80
Men's and Boys' Clothing, NEC	6.94	0.04	0.84
Costume Jewelry and Costume Novelties	24.28	0.03	0.87
Men's and Boys' Separate Trousers	1.15	0.03	0.90
Current-Carrying Wiring Devices	4.63	0.02	0.92
Glass Products, Made of Purchased Glass	14.72	0.02	0.94
Women's, Misses', and Juniors' Blouses	0.86	0.02	0.96
Plastics Products, NEC	1.30	0.01	0.97
Coated and Laminated Paper, NEC	21.03	0.01	0.98
Fabricated Rubber Products, NEC	2.09	0.01	0.99
Woodworking Machinery	2.33	<0.01	0.99
Vitreous China Plumbing Fixtures	4.86	<0.01	1.00
Metal Stampings, NEC	2.97	<0.01	1.00

Table shows compressed industry name, associated F-stat from the first stage, Rotemberg weight, and cumulative Rotemberg weights for the industries selected by the Lasso. NEC means 'Not elsewhere categorized.'

The industries are listed in order of their Rotemberg weight. All nineteen industries have positive Rotemberg weights. This ensures we can interpret the point estimate as a LATE without further assumptions on treatment heterogeneity. Note that higher Rotemberg weight tends to correlate with a higher F-stat (Stearman's rank correlation $\rho > 0.6$). The top five industries, which comprise seventy percent of the overall weight, each have an individual first-stage F-stat ≥ 10. Figure 2 depicts the relationship graphically.

Figure 2: Rotemberg weights are uniformly positive for the post-Lasso instrument



Six industries (with names listed on the figure) comprise three-quarters of the Rotemberg weights. The remaining thirteen industries are less important, sharing the residual twenty-five percent. Notably, the Lasso has deselected any industry with a negative associated Rotemberg weight. Eliminating the vast majority of the 395 industries means more of the renormalized weight is now shared between the remaining nineteen industries. It is unsurprising that the largest individual weight has grown to 0.39, though now it is attributable to the industry (Furniture and Fixtures) with the single largest first-stage F-stat.

The ADH example demonstrates the potential benefits of Lasso-optimized Bartik designs. It illustrates the interaction between Rotemberg weights and variables chosen from the first-stage. To the extent that negative Rotemberg weights are driven by collinearity, applying the Lasso to the first-stage can improve the properties of the estimator. The focus of the section has been the effect of the Lasso on the ADH Rotemberg weights, and we include further details of the application (e.g. the second-stage results) in the Appendix.

While the Lasso-optimized Bartik eliminated negative Rotemberg weights in the ADH case, the approximate sparsity assumption is more appropriate in cases where researchers have no prior beliefs about which sectors are chosen to approximate the endogenous variable. We believe the effect of unemployment on crime is a particularly policy-relevant case. The next section includes one such application.

3 Unemployment and Crime Application

3.1 Institutional Setting and Data

This section contributes with analysis of crime data from Ireland during the early 2000s Celtic Tiger economic boom and the Great Recession beginning in 2008. We believe there are two reasons why this might provide a compelling setting for analyzing the effect of the labor market on crime. The first advantage is the centralized nature of policy-making in Ireland. This mitigates concerns that policy heterogeneity is an omitted variable. Ireland is a relatively small country, and so central government is more important than in larger jurisdictions like the USA or the United Kingdom. Ireland has a single, national police force and the crime data are collated by the Census agency directly from the police's internal reporting database. While regional councils have control over policies like road maintenance and waste collection, they have minimal influence on education, social insurance policy, or criminal enforcement. With the central government controlling practically all of the relevant policy levers, it is less likely that regions have unobserved heterogeneity on these dimensions.

The second advantage is the severity of the business cycle. The literature has typically relied on relatively modest variation as a basis for parameter identification. For example, estimates from France are based on a 3.6% increase in the unemployment rate,[9] and estimates from Italy are based on 3.2% increase.[10] In contrast Ireland's unemployment rate rose more than ten percentage points between 2007–2011.[11] Unemployment peaked at 16% before dropping to 12.8% at the end of 2013.[12] Such labor market volatility is essentially unprecedented for a developed country. The enormous variation in unemployment rates improves the precision of estimates of the effect of unemployment on crime.

The consensus in the crime literature is that deteriorating labor markets increases crime. Freedman and Owens (2016) identify the effect of unemployment on crime in San Antonio using the effect of a federal construction project. This program created new construction jobs in the area but excluded previously convicted felons, creating differential treatment. Neighborhoods with more convicts, who were ineligible for the federal jobs, saw relatively smaller decreases in crime. The focus on previously incarcerated people is continued in Schnepel (2018), which uses a Bartik research design to estimate the effect of unemployment on criminal recidivism.

The crime data used in this paper are the reported crime statistics provided by An

[9] Fougére, Kramarz and Pouget (2009), from 8.9% to 12.5%.

[10] de Blasio, Maggio and Menon (2016), from 6% to 9.2%.

[11] From 4.7% to 14.9%. Of course this is not a critique of the cited papers — authors do not control unemployment rates — but rather an explanation of why this setting is fruitful.

[12] The census agency phased out the Quarterly National Household Survey in 2016, replacing it with Labor Force Survey in 2017. Problems with missing values mean we stop our analysis at the end of 2013.

Garda Síochána[13] to the Central Statistics Office (CSO). The CSO is the independent agency charged with publishing official statistics and conducting the census. Raw administrative data from the Garda PULSE computing system forms the basis of these reports, and the CSO verifies and classifies these records using the Irish Crime Classification System (ICCS).

We focus on three types of property crime. Thefts and burglaries are standard measures of property crime.[14] We also classify robbery as a property crime.[15] The numbers for any sub-component can be small, and so we sum all three to generate an omnibus 'All property crime' variable. The crime data run from the start of 2003 through to the end of 2013.[16]

The geographic unit of analysis is a NUTS (Nomenclature of Territorial Units for Statistics) Region, which is the standard practice across the EU. There are eight NUTS-3 regions in Ireland, averaging about half a million people, and they provide the basic unit of analysis for this paper. With over one million people, the Dublin Region is a considerable outlier and so all regressions are weighted by population. The population weights come from the censuses, with linear interpolation between year-quarters.

All data are quarterly. The labor market data primarily come from the CSO's Quarterly National Household Survey (QNHS). The QNHS is a statutorily-based administrative survey, and details the number of people employed in each of the fourteen NACE-2 economic sectors by region.[17] We use the number of people on the unemployment insurance rolls ("Live Register") as our measure of unemployment, also provided by the CSO.

3.2 Empirical Specification

To estimate the effects of unemployment on property crime, we use variation in unemployment levels within regions and across year-quarters. Our regression model is given by:

$$Crime_{rt} = \alpha_r + \gamma_t + \lambda_r t + \beta \widehat{Unemployment}_{rt} + C'_{rt}\delta + u_{rt} \tag{2}$$

where $Crime_{rt}$ represents different measures of property crime in region r at year-quarter t; α_r are region fixed effects; γ_t are time fixed effects; $\lambda_r t$ are region-specific linear time trends; $\widehat{Unemployment}_{rt}$ is the predicted number unemployed people (measured in thousands); C_{rt} is the share of the population that are male; and u_{rt} is an error term, which we allow to be

[13] The police force in Ireland is called An Garda Síochána, which translates as "The Guardians of the Peace".

[14] The distinction between theft and burglary is an unlawful entry.

[15] Robbery is defined as theft achieved with physical force, so it could alternatively be considered a violent crime.

[16] The code to calculate Rotemberg weights requires a balanced panel with no missing values. There are missing values in the data beginning in 2014.

[17] Agriculture, forestry and fishing; Construction; Wholesale and retail; Transportation and storage; Accommodation and food service; Information and communication; Professional, scientific and technical; Administrative and support services; Public administration and defence; Education; Human health and social work; Industry; Financial, insurance and real estate; and Other.

correlated within regions. Our parameter of interest is β and it measures the average effect of unemployment on property crime.

One potential threat to the validity of our estimates is the misreporting of crime.[18] Region and time fixed effects will capture many of these differences in reporting. Region-specific variation (e.g. high or low reporting in Dublin) is partialled out from the estimates, and so are not a concern. Similarly, time trends in reporting will be captured by the year$\times$quarter fixed effects. Further, region-specific linear time trends will capture first-order trends within each region. Differential trends of misreporting across regions will, however, bias the estimates.[19]

Measurement error and concerns about reverse causality mean OLS is unlikely to identify treatment effects. Consequently we employ a Bartik shift-share instrument as discussed in Section 2. To predict unemployment, our Bartik specification uses pre-period shares of employment (four-period lag), with a leave-one-out growth rate for each industry. The prediction for unemployment will depend on whether we use a traditional Bartik or the Lasso-optimal specification.

The implicit economic theory underlying the analysis is a standard Becker (1968)-type model where crime can be represented as an alternative to traditional employment. A negative shock to the economy transfers people from the labor market to the 'informal alternative.' Consequently we expect a strong relationship between unemployment and property crime, primarily through the mechanism of increased marginal utility of consumption from lower income levels.

3.3 Results

Table 2 presents the estimated effects of unemployment on crime and is the main empirical table of the paper. We report the results in three panels. Panel A depicts standard OLS estimates, while Panels B and C show two variants of the Bartik instrument. We denote the baseline Bartik using all industries as the 'Traditional Bartik.' This is in contrast to the Belloni et al. (2012) Lasso-based instrument depicted in Panel C which we call the 'Optimal-first stage Bartik.' Columns 1 through 4 present the point estimates for the effect of unemployment (measured in thousands of people on the unemployment rolls, per region, per quarter) for each of Theft, Burglary, Robbery, and All property (the sum of all three) crimes respectively.

The first result is that OLS finds that 1000 extra unemployed people is associated with about 31 extra property crimes per quarter. This effect is concentrated in thefts, which

[18] Reporting behavior is correlated with a number of variables that are correlated with unemployment such as the socioeconomic profile of the population, the number of police officers, variation in local policies, etc.

[19] The relationship between the police force and the CSO is often contentious. The CSO have publicly questioned the classifications of the data provided by the police e.g. the police's designation of homicides as manslaughter versus murder. To our knowledge there is no disagreement (on e.g. time-varying trends) that would bias the estimates in this paper.

Table 2: Estimated Effects of Unemployment on Quarterly Crime

	Theft (1)	Burglary (2)	Robbery (3)	All property (4)
Panel A: OLS				
Unemployment level (1000s)	17.08	9.58	4.34	30.99
p-value	(0.028)	(0.026)	(0.022)	(0.047)
95% CI	[8.45, 24.00]	[1.99, 11.45]	[1.09, 4.83]	[16.16, 37.94]
N	352	352	352	352
Panel B: Traditional Bartik				
Unemployment level (1000s)	71.99	22.89	11.08	105.96
p-value	(0.103)	(0.046)	(0.025)	(0.039)
95% CI	[-6.39, 134.17]	[0.61, 49.25]	[-510.03, 19.89]	[.2921, 193.7]
Cluster-robust weak ID F-stat	78.05	76.92	83.31	117.91
N	352	352	352	352
Panel C: Optimal-first stage Bartik				
Unemployment level (1000s)	26.31	22.08	7.71	56.10
p-value	(0.040)	(0.001)	(0.003)	(0.021)
95% CI	[1.57, 57.41]	[8.12, 30.77]	[2.38, 10.77]	[12.58, 99.98]
Cluster-robust weak ID F-stat	41.91	72.61	59.98	66.15
N	352	352	352	352

Notes: Each column within a panel represents a different regression. All specifications include the share of the population that are male, region fixed effects, year-by-quarter fixed effects and region-specific linear time trends, and are weighted by regions' population. We report wild-cluster bootstrap p-values and 95% confidence intervals constructed with 999 replications and Webb weights, using the Stata command *boottest* developed by Roodman, Nielsen, MacKinnon and Webb (2019).

comprise about half of the increase. The point estimates are all positive, and collectively are significant using the wild-cluster bootstrap SEs. Wild-bootstrap SEs are not symmetric in general.

The traditional Bartik IV finds substantially larger point estimates than OLS. This is a result commonly found in the literature, e.g. Raphael and Winter-Ebmer (2001). In general the results are three or four times as large, e.g. an increase in all property from 31 extra crimes to 106 extra crimes per quarter. These results are not precisely estimated. The 95% CI for the effect of unemployment on All property crime spans from just above zero to nearly 200, though the p-value on the null of zero effect is significant at the 5% level.

As shown in Panel C, estimates are smaller than those in Panel B using the optimal first-stage approach. All results are positive, and the CI on the estimate in Column 4 is halved. The coefficients in Panel C fall in between the OLS and traditional Bartik estimates. While the traditional Bartik found point estimates about four times as large as OLS, the estimates

using the the optimal first-stage approach are about twice as large the OLS estimates.

Given that we have few clusters, we report cautious measures for both the first- and second-stages. Firstly our measure of first-stage strength is based on the weak IV test of Olea and Pflueger (2013). Secondly, owing to the relatively small sample size, the confidence intervals (CIs) and p-values and bootstrapped as per Roodman, Nielsen, MacKinnon and Webb (2019). These CIs penalize the small number of clusters in our data, and thus represent conservative values.

When faced with substantially different results (the point estimates using optimal-first stage Bartik specification are 47% smaller than those found by the traditional Bartik), the researcher faces the question of which results are more credible. We propose that unpacking the variation from the instrument can aid the researcher in this task. The Rotemberg weights underlying the results in Panels B and C are depicted in Figure 3.

The top panel in Figure 3 shows the Rotemberg weights for the traditional Bartik. We see considerable dispersion in the weights and the first-stage F statistics. These results are troubling on at least two counts. Firstly, three of the fourteen sectors have negative Rotemberg weights, and in two cases the associated weights are quite large ($|\rho| > 0.3$). This complicates the interpretation of the traditional Bartik treatment effects. Secondly, at $\rho > 0.5$, the weight attributed to Health is very large. This means the overall treatment effect is heavily dependent on this sector. In contrast to the thought experiment of increased import competition, Health is a questionable candidate sector. The full name of this sector is 'Human health and social work'. The number of social workers in an area is possibly endogenous with local crime.

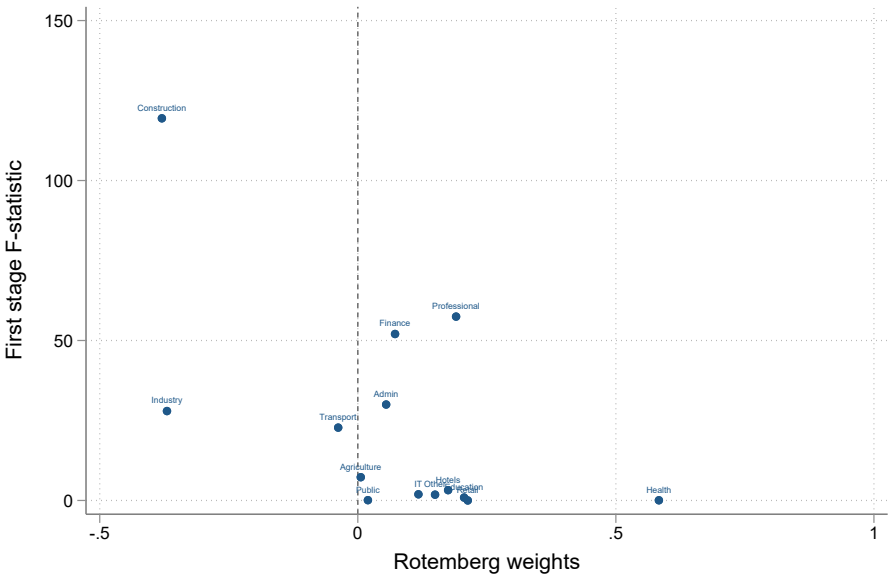
The bottom panel in Figure 3 depicts the Rotemberg weights for the 'optimal' Lasso Bartik. Several features should be noted. Firstly, the number of selected industries is substantially reduced, from fourteen to four. Secondly, all weights are positive. Thirdly, three of the selected industries (IT, Professional/Scientific, and Hotels) are largely driven by international considerations.

The Rotemberg weights can be interpreted as a sensitivity-to-misspecification elasticity as in Andrews, Gentzkow and Shapiro (2017). The Bartik instrument is more sensitive to misspecifation/endogeneity in higher weight industries, so it is instructive to consider the IT and professional/scientific industries in Ireland. Ireland has a low corporate tax rate and has been called a tax haven.[20] Multinational corporations employ over 300,000 people, about fifteen percent of the labor force. Examples of IT firms include Google in Dublin, Apple in Co. Cork, and Intel in Co. Kildare. When related to crime, the exclusion restriction is that employment by these companies (and expansions to the workforce) only affect crime through the employment channel. Concerns about reverse causality (e.g. firms moving location based

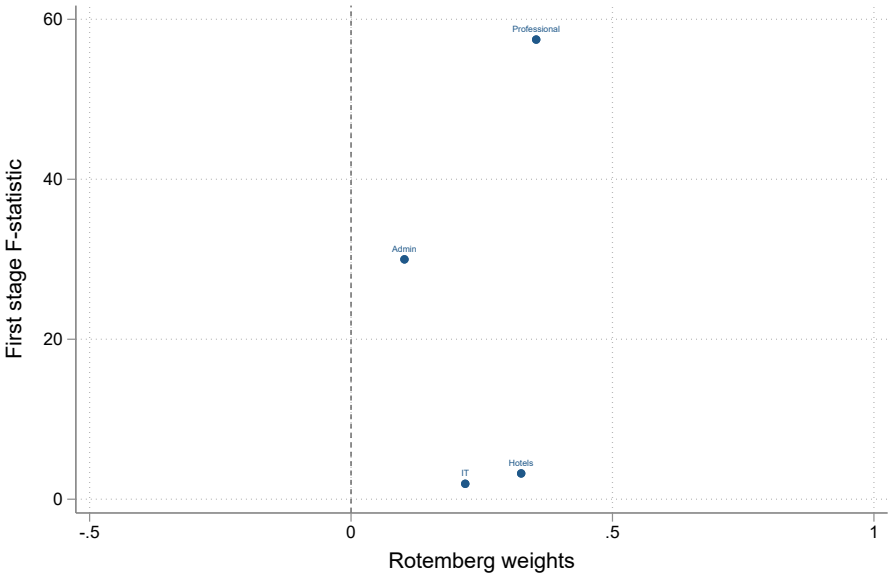
[20]For the tax haven claim, see e.g. Hines Jr and Rice (1994). This is a contested claim. Ireland is not considered a tax haven by the OECD or the European Commission (Tobin and Walsh, 2013).

Figure 3: Rotemberg Weights by Industrial Sector: Unemployment

(a) Traditional Bartik



(b) Optimal-first stage Bartik



on crime levels) seem less relevant in the context of multinational companies.

In terms of Professional/Scientific, Ireland is the seventh largest exporter of medicinal and pharmaceutical products in the world.[21] Twenty-four of the largest twenty-five bio-pharmaceutical companies have offices in Ireland.[22] This sector, too, seems like a good candidate for a Bartik instrument where the pre-period shares interacted with internationally driven shifts are a good candidate for exogenous shifts to employment. Both of these sectors are given negative weights in the baseline specification. A similar argument could be made for the Hotels sector, though domestic demand makes that case less clear-cut.

We conclude that analysis of the Rotemberg weights leaves the researcher less satisfied with the traditional Bartik. The optimal-first stage Bartik has compelling first-stage properties and point estimates more amenable to a LATE interpretation. As the point estimates are over forty percent lower than a traditional Bartik, we conclude that the effects of unemployment on crime could otherwise have been overstated. This is not to say the effects are unimportant, particularly as the preferred specification still indicates point estimates approximately twice as large as OLS suggests, but the optimal-first stage Bartik provides nuance that the existing approaches would miss.

4 Conclusion

This paper provides one potential mechanism to revive the Bartik shift-share research design. The econometrics of optimal sparse IVs has shown the benefits of bringing machine learning into causal inference (Belloni et al., 2012). Our application shows how Rotemberg weights can be sensitive to collinearity. We show that Lasso estimators tend to reduce collinearity in the first stage. By incorporating elements on the optimal sparse IV literature into the framework of Bartik instruments, we combine elements of both literatures to mitigate concerns about negative Rotemberg weights. With our point estimates of the effect of unemployment on crime reduced by more than forty percent, we show that the methodological distinction substantively affects estimates.

Correlational evidence likely suffers from a selection problem, and the Bartik (1991)-style shift-share instrument is a popular research design to recover causal parameters of interest. The strategy (or variants) has been extensively employed to show negative labor market shocks increasing crime, e.g. Gould et al. (2002); Fougére et al. (2009); Schnepel (2018). Bartik-style instruments can also be found in other literatures, e.g. Autor et al. (2013); Brunner et al. (2011); Hershbein and Kahn (2018); Wozniak (2010); Deming and Walters (2018).

[21] See https://www.idaireland.com/doing-business-here/industry-sectors/bio-pharmaceuticals. Accessed in June 2020.

[22] See https://www.siliconrepublic.com/careers/biotech-pharma-companies-ireland. Accessed in June 2020.

The statistical properties of traditional shift-share approaches have been queried by recent research (Adao et al., 2019; Goldsmith-Pinkham et al., 2020; Borusyak et al., 2018). There is a consensus that the validity of the Bartik instrument requires stronger assumptions than previously thought, and subsequently that the design may previously have been mis-applied. In particular, Goldsmith-Pinkham et al. (2020) provide sufficient conditions for second-stage coefficients to be interpreted as a LATE. These conditions, named uniformly positive Rotemberg weights, are rarely met. As the conditions are rarely met, previous results in the literature may require some reinterpretation.

In addition to being an under-explored data source, the institutional features of Ireland make the setting a useful test-case. Owing to the relatively modest population, government policy in Ireland is heavily centralized. Consequently there is no policy heterogeneity across regions, which helps alleviate concerns about confoundedness. Drawing on data from the Great Recession where the unemployment rate quickly rose by 10 percentage points, we find clear evidence that unemployment increases property crime. The most important sectors for the Bartik instrument in the Irish case are dominated by foreign direct investment. As many employers in this sector use Ireland as a low tax international base, expansions by firms in these sectors are heavily driven by factors unrelated to local crime conditions. Our results suggest that job creation generates the positive externality of lower crime. A cohesive crime reduction policy could thus include labor market activation measures.

We do not suggest that applying the Lasso to the first-stage is a panacea to all identification and inference concerns of a Bartik instrument. Rather, we note that our approach allows researchers to estimate a more precise first-stage in a way that addresses collinearity. Our work is best considered in the context of the new literature. For example the concerns about inference/inflated significance of Adao, Kolesár and Morales (2019) persist with a different first-stage, and we encourage researchers to remain mindful of potential problems with shift-share instruments.

References

Adao, Rodrigo, Michal Kolesár, and Eduardo Morales, "Shift-share designs: Theory and inference," *The Quarterly Journal of Economics*, 2019, *134* (4), 1949–2010.

Ahrens, Achim, Christian B Hansen, and Mark E Schaffer, "lassopack: Model selection and prediction with regularized regression in Stata," *arXiv preprint arXiv:1901.05397*, 2019.

Aizer, Anna, "The gender wage gap and domestic violence," *American Economic Review*, 2010, *100* (4), 1847–1859.

Andrews, Isaiah, Matthew Gentzkow, and Jesse M Shapiro, "Measuring the sensitivity of parameter estimates to estimation moments," *The Quarterly Journal of Economics*, 2017, *132* (4), 1553–1592.

Athey, Susan and Guido W Imbens, "Machine learning methods that economists should know about," *Annual Review of Economics*, 2019, *11*, 685–725.

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

Bartik, Timothy J, *Who Benefits from State and Local Economic Development Policies?*, WE Upjohn Institute for Employment Research, 1991.

Becker, Gary S, "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, 1968, *76* (2), 169–217.

Belloni, Alexandre, Daniel Chen, Victor Chernozhukov, and Christian Hansen, "Sparse models and methods for optimal instruments with an application to eminent domain," *Econometrica*, 2012, *80* (6), 2369–2429.

_ , Victor Chernozhukov, Christian Hansen, and Damian Kozbur, "Inference in high-dimensional panel models with an application to gun control," *Journal of Business & Economic Statistics*, 2016, *34* (4), 590–605.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel, "Quasi-experimental shift-share research designs," Technical Report, National Bureau of Economic Research 2018.

Brunner, Eric, Stephen L Ross, and Ebonya Washington, "Economics and policy preferences: Causal evidence of the impact of economic conditions on support for redistribution and other ballot proposals," *Review of Economics and Statistics*, 2011, *93* (3), 888–906.

Cameron, A Colin and Pravin K Trivedi, *Microeconometrics: methods and applications*, Cambridge University Press, 2005.

Card, David, "Immigration and inequality," *American Economic Review: Papers and Proceedings*, 2009, *99* (2), 1–21.

Central Statistics Office, "CJQ03: Recorded Crime Offences by Garda Division, Type of Offence and Quarter," 2016. Accessed on cso.ie.

_ , "LRM07: Persons on Live Register by Age Group, Sex, Social Welfare Office and Month," 2016. Accessed on cso.ie.

_ , "QNQ40: Persons aged 15 years and over in Employment by Sex, NACE Rev 2 Economic Sector, Region and Quarter," 2016. Accessed on cso.ie.

Clingingsmith, David, "Negative emotions, income, and welfare: Causal estimates from the PSID," *Journal of Economic Behavior & Organization*, 2016, *130*, 1–19.

de Blasio, Guido, Giuseppe Maggio, and Carlo Menon, "Down and out in Italian towns: Measuring the impact of economic downturns on crime," *Economics Letters*, September 2016, *146*, 99–102.

de Bromhead, Alan, Alan Fernihough, and Enda Hargaden, "Representation of the people: Franchise extension and the "Sinn Féin election" in Ireland, 1918," *Journal of Economic History*, 2020, *Forthcoming.*

Deming, David J and Christopher R Walters, "The impact of state budget cuts on US postsecondary attainment," *Working Paper*, February 2018.

Diamond, Rebecca, "The determinants and welfare implications of US workers' diverging location choices by skill: 1980–2000," *American Economic Review*, 2016, *106* (3), 479–524.

Dormann, Carsten F, Jane Elith, Sven Bacher, Carsten Buchmann, Gudrun Carl, Gabriel Carré, Jaime R García Marquéz, Bernd Gruber, Bruno Lafourcade, Pedro J Leitão et al., "Collinearity: a review of methods to deal with it and a simulation study evaluating their performance," *Ecography*, 2013, *36* (1), 27–46.

Fougére, Denis, Francis Kramarz, and Julien Pouget, "Youth unemployment and crime in France," *Journal of the European Economic Association*, 2009, *7* (5), 909–938.

Freedman, Matthew and Emily G Owens, "Your friends and neighbors: Localized economic development and criminal activity," *Review of Economics and Statistics*, 2016, *98* (2), 233–253.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift, "Bartik instruments: What, when, why, and how," *American Economic Review*, 2020, *Forthcoming.*

Gould, Eric D., Bruce A. Weinberg, and David B. Mustard, "Crime Rates And Local Labor Market Opportunities In The United States: 1979-1997," *Review of Economics and Statistics*, February 2002, *84* (1), 45–61.

Hanlon, W Walker, "Temporary Shocks and Persistent Effects in Urban Economies: Evidence from British Cities after the US Civil War," *Review of Economics and Statistics*, March 2017, *99*, 67–79.

Hershbein, Brad and Lisa B Kahn, "Do Recessions Accelerate Routine-Biased Technological Change? Evidence from Vacancy Postings," *American Economic Review*, 2018, *108* (7), 1737–1772.

Hines Jr, James R and Eric M Rice, "Fiscal paradise: Foreign tax havens and American business," *The Quarterly Journal of Economics*, 1994, *109* (1), 149–182.

Munyo, Ignacio and Martín A Rossi, "First-day criminal recidivism," *Journal of Public Economics*, 2015, *124*, 81–90.

Olea, José Luis Montiel and Carolin Pflueger, "A robust test for weak instruments," *Journal of Business & Economic Statistics*, 2013, *31* (3), 358–369.

Raphael, Steven and Rudolf Winter-Ebmer, "Identifying the Effect of Unemployment on Crime," *Journal of Law and Economics*, 2001, *44* (1), 259–283.

Roodman, David, Morten Ørregaard Nielsen, James G MacKinnon, and Matthew D Webb, "Fast and wild: Bootstrap inference in Stata using boottest," *The Stata Journal*, 2019, *19* (1), 4–60.

Rotemberg, Julio, "Instrument variable estimation of misspecified models," *Working Paper*, 1983.

Schnepel, Kevin T, "Good jobs and recidivism," *The Economic Journal*, 2018, *128* (608), 447–469.

Tibshirani, Robert, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, 1996, *58* (1), 267–288.

Tobin, Gary and Keith Walsh, "What makes a country a tax haven? An assessment of international standards shows why Ireland is not a tax haven," *The Economic and Social Review*, 2013, *44* (3, Autumn), 401–424.

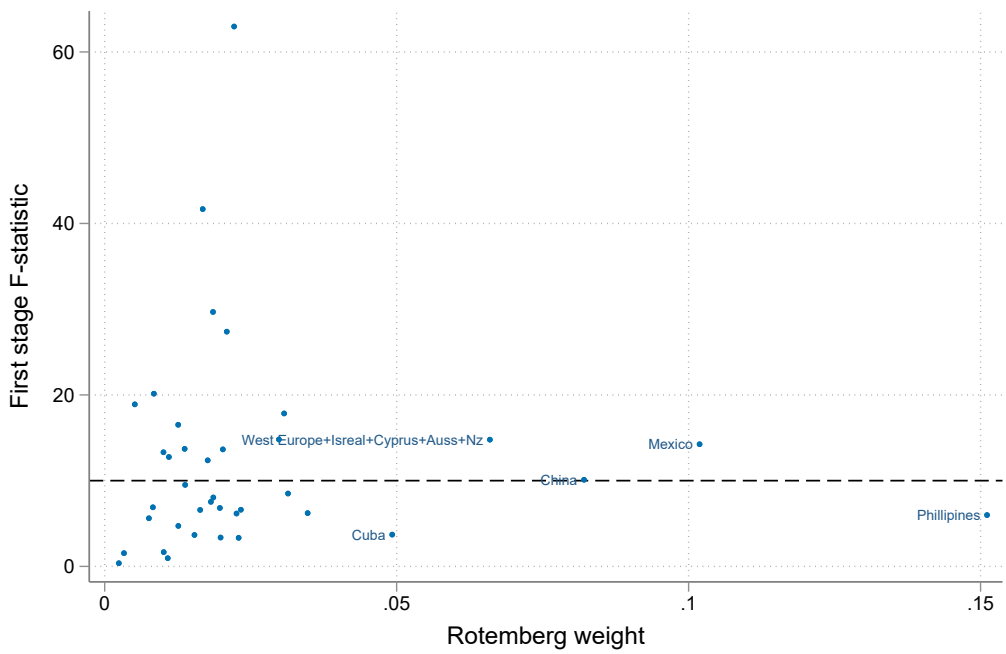
Wooldridge, Jeffrey M, *Introductory Econometrics: A modern approach*, New York: Thomson, 2006.

Wozniak, Abigail, "Are college graduates more responsive to distant labor market opportunities?," *Journal of Human Resources*, 2010, *45* (4), 944–970.

A Application to Card (2009)

Following Goldsmith-Pinkham et al. (2020) we investigate Rotemberg weights David Card's 2009 Ely Lecture (Card, 2009). The principal parameter of interest is the elasticity of substitution between high-skill (college educated) immigrants and natives in 2000. The 'industrial' shares are immigrant community country of origin (e.g. Mexico, China, Philippines) across 124 U.S. cities. There are 38 candidate countries of origin, noting Card aggregated some countries (e.g. Western Europe) together into single units. Figure 4 depicts the distribution of the Rotemberg weights for these countries.

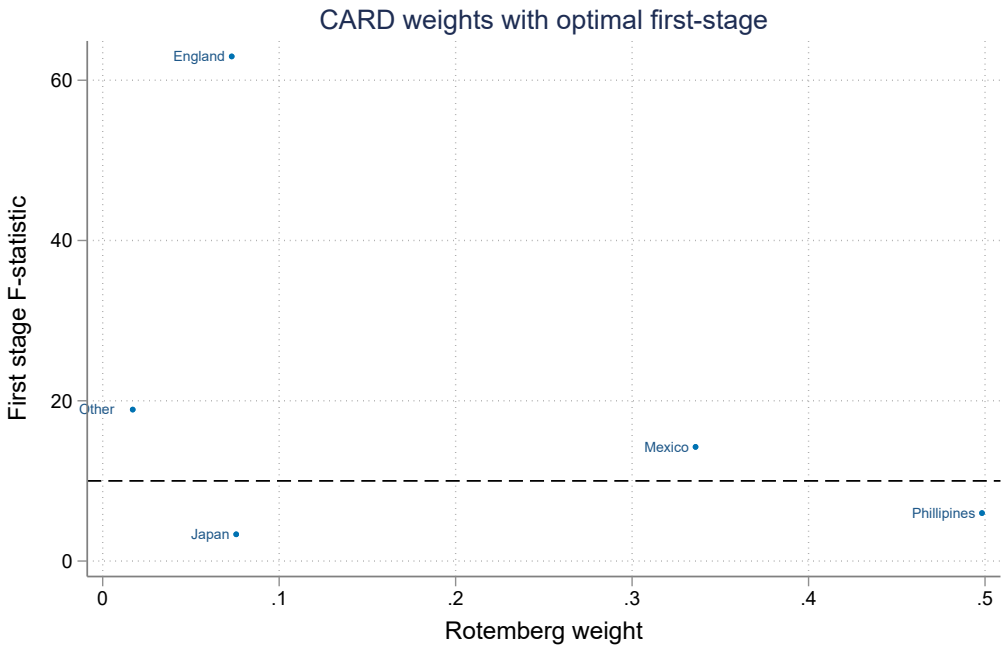
Figure 4: Rotemberg weights for the full Card (2009) dataset



The point estimate for the inverse elasticity of substitution in this setting is $-.078$. In this specific case, all countries receive positive weight in this application. We can thus interpret the $-.078$ point estimate as a LATE even if we believe there is heterogeneity in treatments across countries of origin. The uniformly positive Rotemberg weights present a potential placebo test for the Lasso-optimized Bartik: any large deviations in weights or point estimates may be difficult to reconcile with our preceding motivation. Figure 5 presents the post-Lasso Bartik weights.

We see that from the 38 candidate countries, the Lasso shrinks the first stage to just five. With the removal of 33 countries, the weight attributed to the Philippines has substantially increased. Variation from that country is responsible for half of the overall Bartik estimate,

Figure 5: Rotemberg weights for the post-Lasso Bartik instrument



comparable to the weight Card attributed to Mexico for lower-skilled workers. The point estimate ($\beta = -0.064$) is slightly smaller in magnitude than for the full Bartik, though still quite comparable. All Rotemberg weights remain positive, which is not surprising.

Table 3: Estimated Effects of Immigration on Native Wage Differentials (Card, 2009)

	Traditional Bartik	**Optimal-First Stage Bartik**
Log relative supply of immigrants/natives	-0.078***	-0.064***
	(0.010)	(0.013)
Observations	124	124
First-stage F	33.22	27.73

Notes: Each column represents a different regression. All specifications include college shares in 1980 and 1990, wages for natives and immigrants in 1980 and 1990, the log of city's size in 1980 and 1990, the share of employment in manufacturing in 1980 and 1990, and are weighted by city's population. Standard errors in parentheses are clustered at the city level.

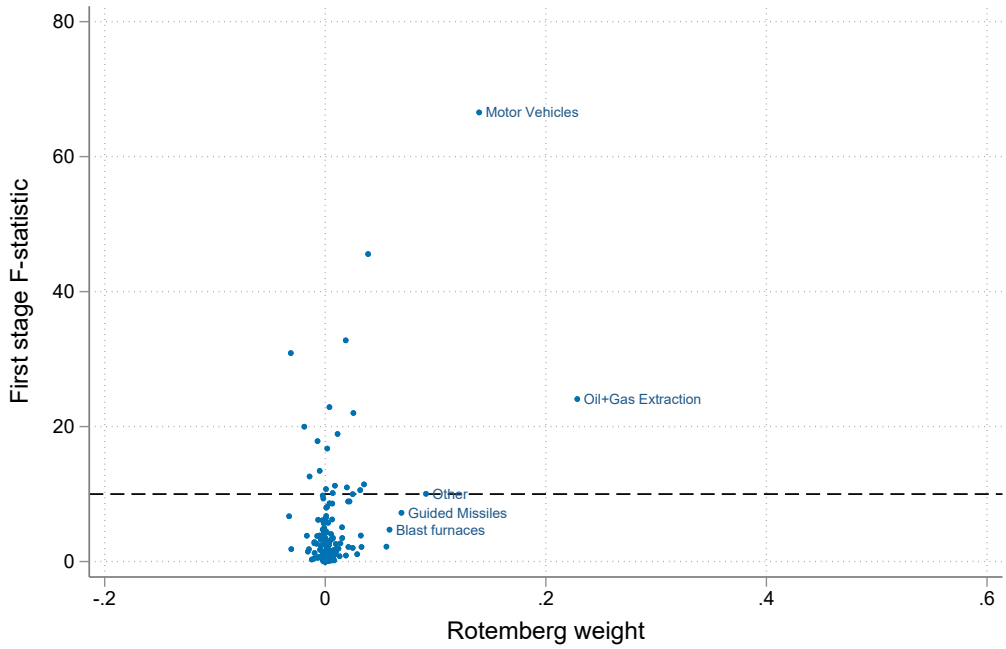
B Application to estimating the elasticity of labor supply

In addition to Autor, Dorn and Hanson (2013) and Card (2009), we provide a third replication.[23] The question is the (inverse) elasticity of labor supply.

The data come from the United States, and include employment shares for 228 industries in 722 commuting zones. The estimates are based on a 5% sample of U.S. decennial census, 1980–2000, supplemented with pooled American Community Survey from 2009–2011 to generate a quasi-2010 observation.

Figure 6 depicts the distribution of Rotemberg weights in the baseline setting. The principal feature of this figure is that two industries, Motor Vehicles and Oil & Gas Extraction, are the dominant drivers of the overall Bartik estimate. We also see that a large number of the 228 candidate industries cluster around zero, including a large number that are negative. While admittedly small on an individual level (mean of -0.004), the fact there are nearly one-hundred negative weights means their sum accumulates to a quite substantial -0.368.

Figure 6: Rotemberg weights for the labor supply elasticity application



The distribution of Rotemberg weights in the Lasso Bartik is shown in Figure 7. We note three major features. Firstly, the Lasso shrinks the first-stage down from 228 candidate industries to just fifteen. This is consistent with what we found in the ADH replication. Secondly, the same two industries that were outliers in Figure 6 are retained, albeit this time

[23] This is the final example presented by Goldsmith-Pinkham et al. (2020).

with increased prominence. Thirdly, we continue to have negative Rotemberg weights. This confirms that the Lasso-optimized first-stage is not a panacea to problems associated with Bartik research designs. While the Lasso reduces the number of negative-weight industries by ninety percent, and the cumulative weight of the negative industries has more than halved (from -0.368 to -0.178), the fact that thirteen percent ($\frac{0.178}{1.178+0.178} = 0.131$) of the overall weight is negative means further assumptions on treatment homogeneity is necessary to ensure interpreting the point estimate as a LATE.

Figure 7: Rotemberg weights for the for the post-Lasso Bartik instrument

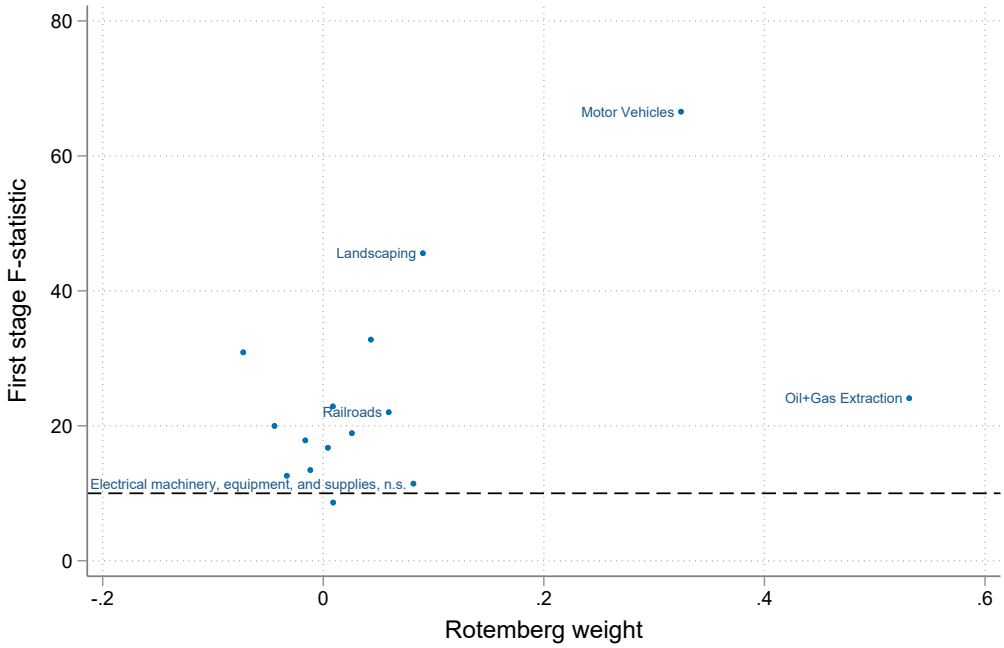


Table 4 shows the decline in negative Rotemberg weights using the Lasso Bartik. The sum of negative weights falls by about one-half, which implies a reduction in the share of negative weights by about one-third. Table 5 shows the second-stage results, where the Lasso-optimized Bartik has a significantly smaller point estimate than the baseline case. This estimate is very close to the OLS estimate (0.63). For the reasons stated above, we are cautious in interpreting this estimate.

Table 4: Summary of Rotemberg weights by sign: inverse elasticity of labor supply

Panel A: Baseline Bartik			
	Sum	Mean	Share
Negative	-0.368	-0.004	0.212
Positive	1.368	0.010	0.788
Panel B: Optimal-First stage Bartik			
	Sum	Mean	Share
Negative	-0.178	-0.036	0.131
Positive	1.178	0.118	0.869

Table 5: Inverse Labor Supply Estimate: Change in Earnings

	Traditional Bartik	**Optimal-First Stage Bartik**
Change in Employment	1.20***	0.64***
	(0.14)	(0.13)
Observations	2,166	2,166
First-stage F	50.44	22.86

Notes: Each column represents a different regression. Both specifications include year fixed effects, control variables, and CZ fixed effects, and are weighted by start of the period CZ share of the population. Standard errors in parentheses are clustered at the CZ level.

C Further table and figures

C.1 First-stage for Crime and Unemployment application

Table 6: *Effect of Unemployment on Crime*
First-Stage Estimates

	Coefficient	p-value	95% CI
Panel A: Traditional Bartik			
Omnibus Bartik sector shares	-0.21	0.005	[-0.31, -0.10]
Panel B: Optimal-first stage Bartik			
Accommodation and food service	-0.44	0.292	[-0.87, 0.20]
Administrative and support services	-0.73	0.004	[-1.12, -0.53]
Information and communication	-0.81	0.145	[-1.43, 0.16]
Professional, scientific and technical	-0.72	0.000	[-1.16, -0.48]

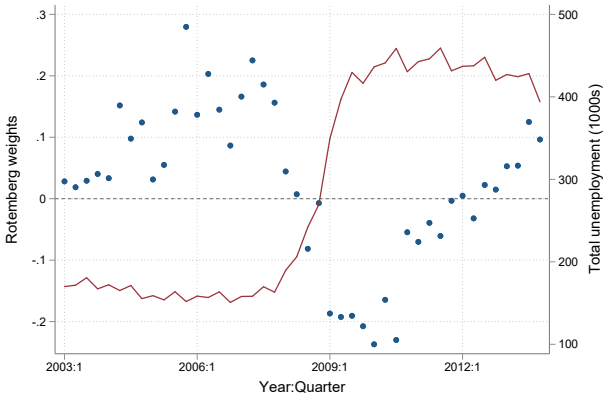
Notes: Table shows the first-stage estimates for the Traditional Bartik and Lasso-Bartik approaches in Table 2. The regression specification includes the share of the population that are male, region fixed effects, year-by-quarter fixed effects and region-specific linear time trends, and are weighted by regions' population. We report wild-cluster bootstrap p-values and 95% confidence intervals constructed with 999 replications and Webb weights, using the Stata command *boottest* developed by Roodman, Nielsen, MacKinnon and Webb (2019).

C.2 Intertemporal Rotemberg weights

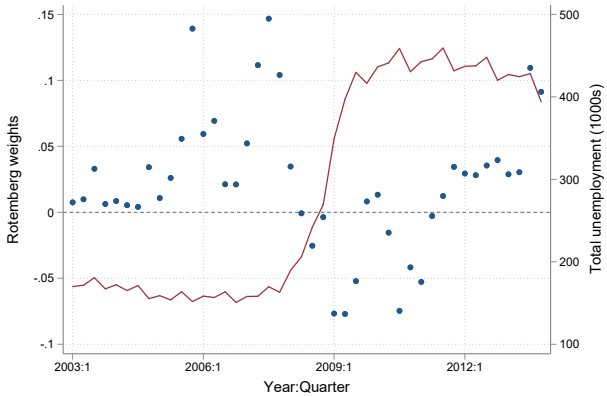
Figure 8 shows the pattern of Rotemberg weights (dots, measured on left y-axis) compared to unemployment (solid line, measured on right y-axis) over time. Panel (a) depicts the distribution in the Traditional Bartik case. There is a cluster of negative weights (dots) around 2009–2010 when unemployment rates rose. There is a large dispersion of the weights, with many of the negative weights substantially large in magnitude ($|\rho| \simeq 0.2$). Cumulatively this cluster of weights adds to more than one. These large and negative weights are compensated against a considerable number of moderately large ($0.15 < \rho < 0.3$) positive weights.

Figure 8: Rotemberg Weights by Year-by-Quarter: Unemployment

(a) Traditional Bartik



(b) Optimal-first stage Bartik



Panel (b) of the figure depicts the distribution when the first-stage is Lasso-based. As it depicts the number of people unemployed, the solid line is identical to Panel (a). The distribution of Rotemberg weights, in particular the left y-axis, is different. Although the overall pattern of weights is similar in Panel (a) and Panel (b), there are two differences. Firstly, the distribution is condensed: while many weights are quite far from zero in Panel (a), all weights are now within $[-0.1, -.15]$. The 'modal weight' has a relatively modest magnitude $-0.05 < \rho < 0.05$. Secondly, the number of negative weights has fallen by about a third. Compared to the baseline Bartik, the number of negative weights has fallen from fifteen to eleven.

C.3 Second-stage for Autor, Dorn, Hanson (2013) application

Table 7 shows the effect of implementing the Lasso Bartik on the point estimate in the Autor et al. (2013) example. The baseline point estimate of -0.746 (with a standard error of 0.07) suggests that a $1,000 rise in import exposure reduces a commuting zone's manufacturing employment per working-age population by three-quarters of a percentage point. The Lasso point estimate is larger, closer to one percentage point. Autor et al. (2013) note that the -0.746 point estimate is dampened by the inclusion of control variables, and we suspect the same result would hold in the Lasso case.

Table 7: Effect of Imports from China on Manufacturing Employment (Autor et al., 2013)

	Traditional Bartik	**Optimal-First Stage Bartik**
Δ Imports	-0.746***	-0.978***
	(0.068)	(0.164)
Observations	1,444	1,444
First-stage F	97.54	71.27

Notes: Each column represents a different regression. All specifications include year fixed effects and CZ fixed effects, and are weighted by start of the period CZ share of the population. Standard errors in parentheses are clustered at the state level.