

ORIGINAL ARTICLE

A Bayesian structural time series analysis of the effect of basic income on crime: Evidence from the Alaska Permanent Fund*

Richard Dorsett 

University of Westminster, London, UK

Correspondence

Richard Dorsett, University of Westminster,
35 Marylebone Road, London NW1 5LS,
UK.

Email: r.dorsett@westminster.ac.uk

Abstract

This paper examines the impact of Alaska's Permanent Fund Dividend on crime. The Dividend has been payable annually to state residents since 1982 and is the world's longest-running example of a basic income. Initially universal, from 1989 onwards eligibility was withdrawn from an increasing proportion of those in prison. A Bayesian structural time series estimator is used to simulate Alaskan crime rates in the counterfactual no-Dividend case. A comparison with actual rates provides an estimate of the Dividend's impact. This does not provide strong evidence of an effect. However, incorporating information on Dividend amounts suggests the size of payments is important, with larger amounts reducing crime. There is no evidence that this effect is reinforced by the 1989 change to eligibility. The results demonstrate the potential for a basic income to encourage positive outcomes and lend support to payment being universal rather than conditional.

KEYWORDS

Alaska Permanent Fund, basic income, Bayesian structural time series, case study, crime

*I am grateful to Mouhcine Guettabi, Ioana Marinescu, the Joint Editor, the Associate Editor and an anonymous reviewer for helpful comments. The usual disclaimer applies.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. Journal of the Royal Statistical Society: Series A (Statistics in Society) published by John Wiley & Sons Ltd on behalf of Royal Statistical Society

1 | INTRODUCTION

The idea of a universal basic income is certainly not new (More 1516 [1963]) but recent years have seen increased interest. Sloman (2018) identifies the current decade as representing the latest distinct wave of enthusiasm for the concept. In the United Kingdom, the Labour Party included a commitment to a pilot of basic income in its 2019 general election manifesto (Labour Party, 2019). More recently, the United Nations has issued guidance identifying basic income as one means by which countries can help safeguard against the acute effects of the COVID-19 pandemic (Office of the United Nations High Commissioner for Human Rights, 2020).

Advocates highlight its simplicity; in principle, a universal payment could replace a complex system of welfare payments (Murray, 2016). Furthermore, they view the payment as providing a means of addressing problems of poverty and inequality, extending financial security to all, while freeing individuals from need to comply with job search requirements when out of work. Critics worry about a basic income acting as a disincentive to work. Furthermore they argue that any ambition to replace welfare systems with a basic income would cost so much as to be non-viable (Hoynes & Rothstein, 2019).

It is useful to distinguish between a full basic income, designed to free individuals from poverty and allow full social inclusion, and a partial basic income which falls short of that and is more of an income supplement. A full basic income could replace existing state benefits. Under a partial basic income, the interaction with the tax and welfare system becomes particularly relevant and provides a link to the negative income tax experiments carried out between 1968 and 1982 in the United States and Canada (Hum & Simpson, 1993; Robins, 1985). Martinelli (2019) assesses the trade-off between adequacy (meeting individuals' needs) and affordability of partial schemes, with a particular focus on how the purported benefits of a basic income vary with its generosity.

A number of randomised trials have assessed how basic income may function in practice. Trials in Namibia (Haarmann, 2009) and India (Bharat & UNICEF, 2014) have demonstrated the poverty reduction potential in a developing country context. Among developed countries, Finland experimented with a basic income for recipients of unemployment benefits aged 25–58; evaluation evidence suggests no employment effect but a positive benefit with respect to well-being and attitudes such as trust in institutions (Kangas et al., 2019). In Canada, a negative income tax trial for low-income people was launched but cut short in March 2019 (<https://www.ontario.ca/page/ontario-basic-income-pilot>). The US examples include a small trial currently underway in Stockton, California (<https://www.stocktondemonstration.org/>) and a larger experiment planned by the start-up accelerator *Y Combinator* (<https://basicincome.ycr.org/our-plan>).

The Alaska Permanent Fund Dividend is the world's longest-running example of a basic income and, as such, provides unique opportunities for learning about long-term impacts. This paper explores the effect of the Dividend on crime. When introduced in 1982, the Dividend provided a payment to all citizens of Alaska and one might expect the increase in individuals' income to reduce incentives to engage in criminal activity. From 1989, individuals who had served prison sentences in the previous year became ineligible. One might further expect that this would act as a deterrent to offending. The analysis in the paper distinguishes between these two phases and therefore offers evidence on the broader question of how income and income penalties affect crime. The primary outcome considered is property crime since this has a clear economic motivation.

A small number of papers have looked at various effects of the Alaskan Dividend. Hsieh (2003) examined individuals' consumption responses, Evans and Moore (2011) looked at mortality rates around the time of receipt and Goldsmith (2012) considered economic and social effects more broadly, while noting the absence of robust evidence. Jones and Marinescu (2018) examined the impact on employment, finding no effect. Only one paper has looked at the effect of the Dividend on crime.

Watson et al. (2019) used daily police reporting data to examine how crime changes around the time of Dividend payment. They showed an increase in substance abuse incidents on the day following payment which, since the precise timing of payment is quasi-random, can be interpreted as causal. No such ‘day-after’ effect was found for property crime but a significant reduction over the 2 weeks following payment day was detected.

The aim of this paper is to estimate how the *existence* of the Dividend and the change to eligibility affects crime. It does this by comparing observed crime to an estimate of crime in the counterfactual scenario of an Alaska without the Dividend. This is the first time such an analysis has been attempted.

The analysis uses a Bayesian structural time series approach (Brodersen et al., 2015) to simulate crime rates in a counterfactual ‘no-Dividend’ Alaska. The results did not find strong evidence of an effect. However, examining the relationship between year-on-year variations in estimated impacts and changes in the size of the Dividend revealed an association which, based on an autoregressive distributed lag model, has a causal interpretation. Before the change to eligibility for prisoners, a one-off increase in the Dividend would reduce property crime for at least 2 years. After the change to the eligibility rule, the impact was smaller and possibly shorter-lived.

The findings of this paper offer novel evidence on the impact of a basic income on crime. In so doing, they highlight a potential benefit of basic income that has not so far been fully explored. Furthermore, the impacts appear, if anything, to be stronger without the conditionality that was introduced by the 1989 change to eligibility for prisoners. The results do not support the intuition that such conditionality should deter crime but rather provide an argument for retaining the principle of universality.

The remainder of the paper has the following structure. In Section 2, the Alaska Permanent Fund is described in more detail. Trends in Alaskan crime are presented and compared to the other states and the United States as a whole. Section 3 describes how impacts were estimated. These are presented and further analysed in Section 4. Section 5 concludes.

2 | BACKGROUND—THE ALASKA PERMANENT FUND AND CRIME IN ALASKA

2.1 | The Alaska Permanent Fund

In 1976, the Alaska Permanent Fund was established through an amendment to the Alaska state constitution. Each year, a proportion of Alaskan oil revenues is diverted into the Fund. Since 1982, the Fund has paid a Dividend to every resident Alaskan.

The size of the Dividend varies each year. The total available for distribution is calculated as the Fund balance at the time of calculation, plus Fund earnings averaged over the previous 5 years, less appropriations and reductions. This is then divided by the estimated number of eligible applicants to give the size of the Dividend. The formula is public knowledge but relies on some elements that are unrealised until the time of calculation and one, the number of eligible applicants, that must be estimated. The Fund Commissioner must announce the value of the Dividend by October 1 of each year. Citizens have a rough sense of the likely size of the payment before the announcement since the Fund Corporation provides estimates some months in advance. The majority of Alaskans receive their annual Dividend in October. The minority whose eligibility is not yet established may receive their payment later in the year.

Figure 1 shows how the nominal value of the Dividend has varied since its introduction in 1982. It was initially set at \$1000 but was considerably smaller in 1983, after which it followed a mostly upward

trend until 2000 (\$1963). Its value has fluctuated since then, but reached its record level in 2008, when the Dividend of \$2069 was boosted by a one-time ‘Resource Rebate’ supplement of \$1,200, bringing the total amount that year to \$3269. In considering the size of the Dividend, it should be noted that parents can claim on behalf of their children. Consequently, families with children will receive multiple Dividends each year. Watson et al. (2019) report that over the period 2000–2016, average household size was 2.83 and the Dividend represented, on average, 6.28% of total household income.

While initially a universal payment, eligibility restrictions for those in prison were introduced in 1989 and strengthened thereafter. Since 1989, those incarcerated on felony charges (i.e. for a crime punishable by imprisonment for more than one year) during the year to which the Dividend relates (the reference year) have been ineligible. In 1996, misdemeanants (those convicted of a crime for which a prison term of more than one year is not authorised) in the reference year who had two prior crimes became ineligible. In 2002, misdemeanants in the reference year who had one prior felony or two prior misdemeanours also became ineligible. It should be noted that the children of ineligible adults themselves become ineligible. Consequently, the financial penalty deriving from the change to eligibility may be greater where the sentenced or incarcerated individual has dependent children.

Our expectations of how the Dividend may affect crime is informed chiefly by Becker (1968) who conceptualised individuals’ offending behaviour as a function of the probability of conviction, the penalty associated with conviction and a range of other background factors such as education, civic values and income. In this framework, individuals make a rational decision to engage in criminal activity based on whether the expected benefits exceed the expected costs. A forward-looking individual will compare expected utility flows from criminal and non-criminal activity over an extended period. Lee and McCrary (2009) develop a model that incorporates such dynamics. Its implications are that the

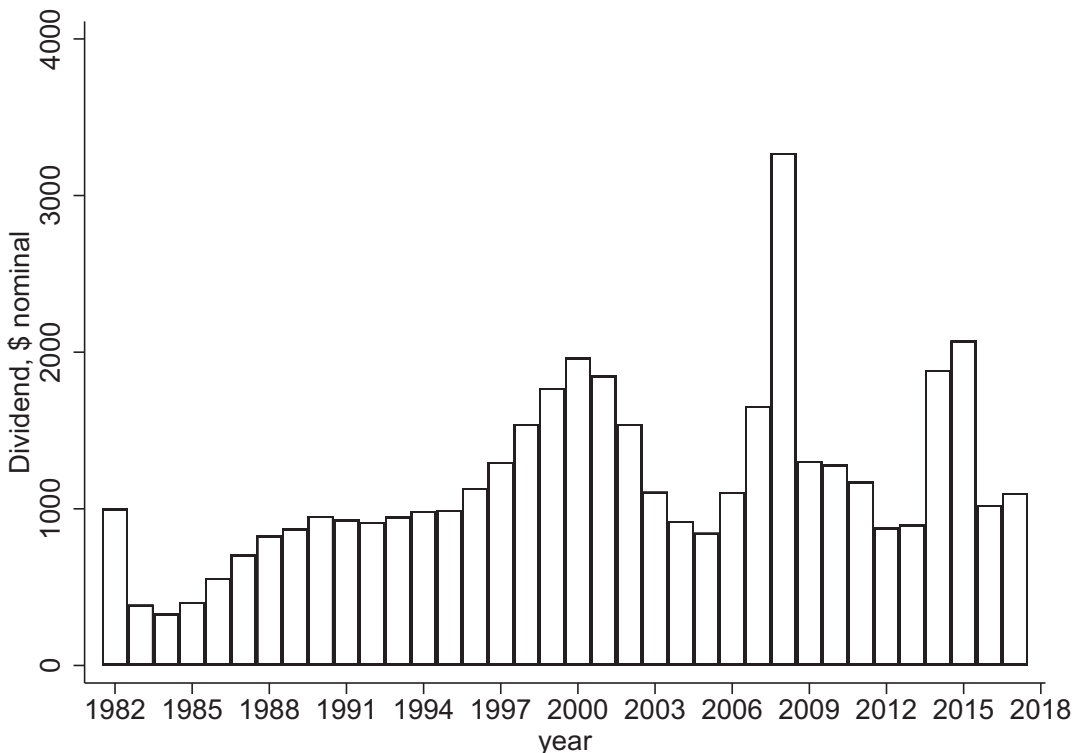


FIGURE 1 Dividend amount, 1982–2017.

attraction of crime will be lower among individuals for whom the per-period utility cost is higher and individuals with higher discount factors (for those who value the future more, lengthy punishment imposes a higher cost). It is also negatively related to sentence length and the probability of being caught.

An advantage of this model is that it provides a framework for interpreting and, to some extent, quantifying the results. Furthermore, it is particularly suited to the case of property crime, where one might expect a strong economic motive. However, it should be noted that other theories of crime exist, within sociology for example, and these might predict different effects of income. Indeed, some have argued that criminology as a discipline is theoretically over-determined, with too many competing theories that have not been rigorously tested, compared, refuted or integrated (Wikström, 2007).

In the Alaskan case, the introduction of the Dividend served to increase individuals' income and reduce inequality. It would therefore be expected, if anything, to reduce property crime (see, for instance, Blakeslee and Fishman (2018) and Choe (2008)). The 1989 change to rules governing Dividend eligibility increased the financial penalty associated with conviction. Intuitively, this should act as a disincentive to commit a crime. A simple decision model is developed in the web-based supporting materials to show how this disincentive is greater when the Dividend is higher. Our expectation then is that the Dividend will, if anything, reduce crime and this reduction should be reinforced by the change to eligibility rules in 1989 and later years.

2.2 | Crime and imprisonment in Alaska

State-level crime data are available from the Federal Bureau of Investigation Uniform Crime Reports, covering the period from 1960 to 2017. Data are missing for New York State between 1960 and 1964, so these years are excluded from the impact analysis. In all cases, crimes are expressed relative to the state population; the number of reported crimes per 100,000 residents. Fuller details on data sources are given in the web-based supporting materials.

The primary focus in this paper is on property crime rather than violent crime since this is more likely to be economically driven and therefore responsive to the threat of reduced income. Property crime itself covers burglary, larceny theft and motor vehicle theft. Results for violent crime—murder, rape, aggravated assault and robbery—are also shown. There are two reasons for this. First, while not so readily accommodated within a model of incentives as property crime, violent crime is nevertheless liable to the influence of socio-economic conditions and circumstances. Second, robbery falls within the category of violent crime despite having an economic motivation.

Figure 2 shows trends over the period 1960 to 2017, displaying both the Alaskan crime rates, the rates for the United States as a whole and the rates for individual states. For reference, vertical lines are included at 1982 (the year of the first Dividend) and 1989 (the year conditionality was introduced for criminals). Property crime as a whole grew in Alaska until the early 1980s and then steadily declined. This was broadly the pattern seen in the United States as a whole, although the rates in Alaska were relatively elevated during the peak years around 1980. Alaskan trends were within the range seen across other states. Looking at types of property crime allows a more detailed insight and highlights differences between Alaska and the United States as a whole. Alaska has tended to have lower rates of burglary and higher rates of larceny. Vehicle theft was initially higher in Alaska but this difference mostly disappeared after the mid-1980s.

With regard to violent crime, this grew in both Alaska and the United States until the early 1990s. There was then a reversal in both cases but this was temporary in the case of Alaska where, after 2000, violent crime reverted to its upward trend, thereby diverging from the United States as a whole. For much of this period, Alaska did not look notably different from other states. Since about 2000,

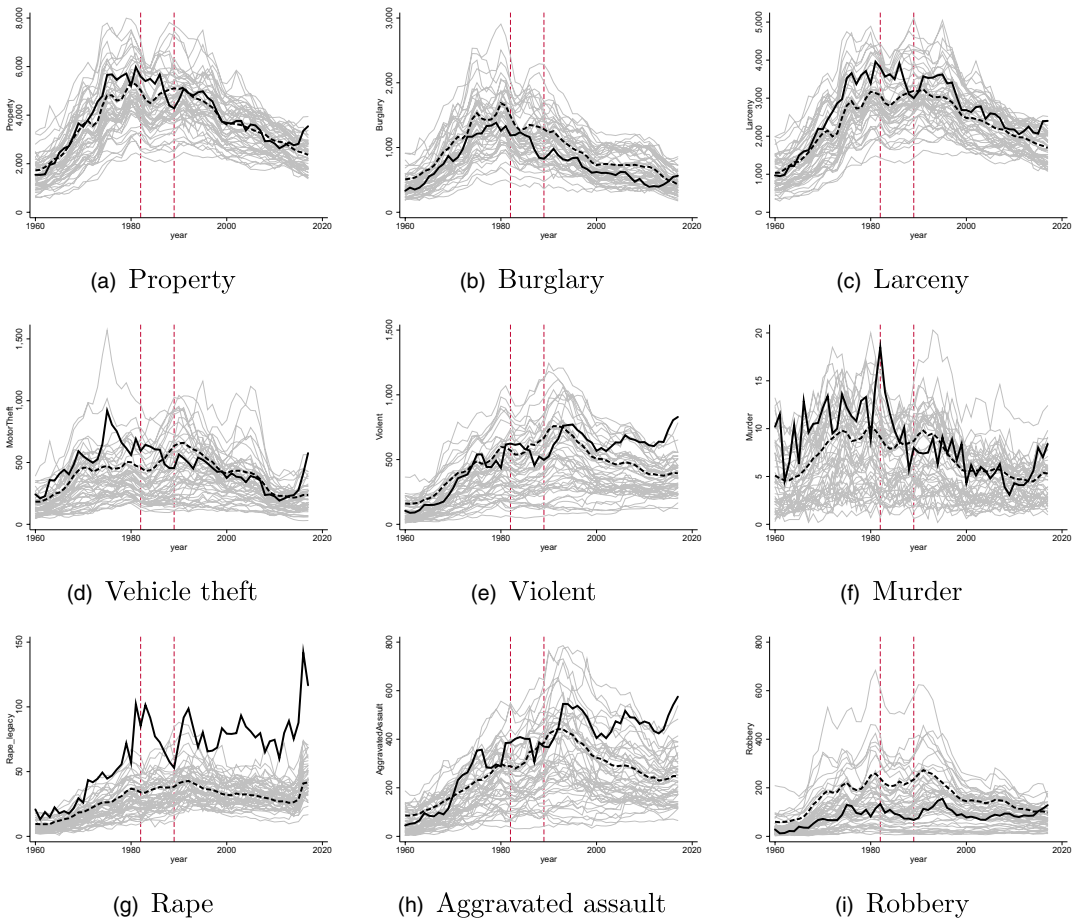


FIGURE 2 Crime trends, 1960–2017 (number of crimes per 100,000 population) Alaska is shown by the solid black line, the United States by the dashed black line and states other than Alaska by grey lines.

though, its ranking has grown to the point where it had the highest rate of violent crime of any US state in 2017. In more detail, murder was initially relatively high in Alaska but converged to roughly the US rate from the mid-1980s. Rape, on the other hand, has consistently been higher in Alaska than the United States, and this difference has increased over time (the definition of rape used in this paper changed in 2016; see the online supplementary material for details). Aggravated assault grew roughly in line with the national trend but the sustained decline that began in the early 1990s in the United States was not matched in Alaska. It is this category that dominates the trends in violent crime in the United States as a whole. Lastly, rates of robbery have mostly been considerably lower in Alaska than in the United States as a whole, although again that has changed in more recent years.

3 | ESTIMATION APPROACH

To estimate the impact of the Dividend on crime requires an estimate of how criminal activity would have evolved in Alaska had the Dividend not been introduced. A comparison of actual outcomes in Alaska with these estimated counterfactual outcomes can then provide an estimate of the impact of the Dividend.

In the next sub-section, some possible evaluation approaches are discussed and the relative strengths of the BSTS estimator are considered. Following this, the BSTS approach to estimating model parameters is then outlined, including the choice of priors. The final subsection describes how counterfactual time series are simulated for the post-intervention period and used to estimate impacts.

3.1 | Possible evaluation approaches

With aggregate data, one option is to conduct a pure time series analysis. However, this risks conflating the impact of interest with the impact of other factors, particularly if the environment has changed over the period analysed. This can be controlled for through difference-in-differences, incorporating data from other areas into the estimation. The identifying assumption in this case is that mean outcomes in these other areas follow the same trend that would have also been seen in the treatment area had the intervention not occurred.

A common strategy for conducting a case study is to adopt the synthetic control approach of Abadie and Gardeazabal (2003) and Abadie et al. (2010). This involves estimating counterfactual (no-treatment) outcomes for a treated unit as a weighted average of the outcomes of non-treated units. Weighting the non-treated units in this way results in a synthetic control, where weights are typically chosen such that it looks similar to the treated unit in respect of those characteristics thought to influence outcomes and also the trajectory of outcomes in the pre-intervention period. It generalises the difference-in-differences estimator by allowing for the effect of unobserved influences to vary over time, thereby relaxing the common trends assumption.

Two features need to be taken into account when assessing its suitability in a particular application. First, the restriction that weights be non-negative limits its usefulness when the treated unit is an outlier (lies outside the ‘convex hull’). An attraction of the approach is that such cases can be readily identified through basic descriptive statistics. Specifically, it is reassuring when it can be shown that the synthetic control resembles the treated unit with regard to observed characteristics and pre-treatment trends. Second, identification relies on the number of pre-treatment periods being sufficiently large relative to the conditional volatility of the outcome. Inspecting pre-intervention trends may help inform the judgment of whether this is achieved in practice but in doing so it should be noted that, as shown in Appendix B of Abadie et al. (2010), the potential bias derives from the volatility of outcomes in the non-treated units rather than the treated unit itself.

In this application, an alternative approach is used; the Bayesian structural time series (BSTS) estimator. Its key features are outlined later in this section (for the full detail, see Brodersen et al. (2015)). Analogous to the synthetic control approach, counterfactual outcomes for Alaska are estimated as a weighted average of the outcomes for other US states. However, unlike the synthetic control approach, these weights derive from the coefficients of an assumed structural model estimated in the pre-Dividend period.

There are two main reasons why this approach is well suited to the research question at hand. First, it avoids the synthetic control requirement of non-negative weights. The coefficients of the structural model are not constrained to be non-negative so other states can contribute negatively to the predicted counterfactual. The flexibility allowed by this property is useful when the treated unit has characteristics that are outside the range seen in other units. Since some characteristics of Alaska differ markedly from other states, the limitation of non-negative weights may well have empirical relevance to the extent that it does not permit a sufficiently similar synthetic Alaska to be formed.

Second, the BSTS estimates take fuller account of model uncertainty. The synthetic control approach provides a single estimate of counterfactual outcomes and no account is taken of the fact

that the weights used for this purpose are subject to uncertainty. BSTS adopts a Bayesian approach, constructing multiple synthetic controls based on draws from the posterior distribution of simulated coefficients over the post-intervention period. Each such synthetic control provides counterfactual outcomes which can be subtracted from the observed outcomes in the treated unit to give a distribution of simulated impacts, each with draw-specific time series properties. This feature is particularly useful in this study where it is exploited to examine how impacts vary year-on-year with changes in the size of the Dividend.

Set against this, the synthetic control approach has its own relative strengths. First, it avoids the need to specify priors. In the analysis that follows, sensitivity of the BSTS results to the choice of priors is reported in order to provide some reassurance that results are not unduly dependent on this. However, it remains the case that the basis for some priors required for BSTS may be difficult to evidence. Second, the synthetic control approach allows the inclusion of characteristics that are not observed in every time period, and need not be observed at all in the post-intervention period. Incorporating covariates can capture some variation in outcomes, and is useful if these covariates—and indeed the influence of these covariates—remain stable across the pre- and post-intervention periods. By contrast, the BSTS estimates in this application are based purely on time series of outcomes in Alaska and other states.

3.2 | Bayesian structural time series-estimating parameters

The first element of the BSTS approach is to simulate draws of the model parameters and the state vector given the observed data in the pre-Dividend period.

Assume we have data over T units of time (in this application, years), of which T_0 units precede the intervention, $1 < T_0 < T$. Write the (observed) crime rate in Alaska as y_t and contemporaneous outcomes in the J other states as a $(J \times 1)$ vector \mathbf{x}_t .

The observation equation is

$$y_t = \mu_t + \beta' \mathbf{x}_t + \varepsilon_t \quad (1)$$

where $\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. The state equation shows how the intercept evolves over time as a random walk

$$\mu_{t+1} = \mu_t + \eta_t \quad (2)$$

where $\eta_t \sim \mathcal{N}(0, \tau^2)$.

States other than Alaska are assumed not to be affected by the intervention. This allows the counterfactual outcomes, y_t^0 , to be estimated as the predicted values from the model, $y_t^0 = \hat{y}_t$, where \hat{y}_t is predicted from coefficients estimated using data from the pre-intervention period.

Denoting by θ the set of all model parameters and by $\alpha = (\alpha_1, \dots, \alpha_T)$ the full state sequence, Bayesian estimation requires specification of a prior distribution $p(\theta)$ and a distribution $p(\alpha_0 | \theta)$ on the initial state. MCMC can then be used to sample from $p(\alpha, \theta | \mathbf{y})$.

In Equation (1), Bayesian variable selection and model averaging is handled through a ‘spike-and-slab’ prior (Scott & Varian, 2014). This data-driven approach sets to zero some elements of β such that the associated states do not contribute to the counterfactual. Regularising in this way offers protection against the risk of over-fitting. The set of states used to form the counterfactual will vary with each draw.

The details of the approach are set out in Brodersen et al. (2015) and summarised here for convenience. Write $\rho = (\rho_1, \dots, \rho_J)$, where $\rho_j = 1$ if $\beta_j \neq 0$ (0 otherwise) and let β_ρ be the non-zero elements of β . The spike-and-slab prior is then $p(\rho, \beta, 1/\sigma_\varepsilon^2) = p(\rho)p(\sigma_\varepsilon^2|\rho)p(\beta_\rho|\rho, \sigma_\varepsilon^2)$.

The ‘spike’ component is specified as the product of independent Bernoulli distributions, $p(\rho) = \prod_{j=1}^J \pi_j^{\rho_j} (1 - \pi_j)^{1-\rho_j}$. In this application, the prior probability of inclusion, π_j , is set to $J_0/J, \forall j$, where J_0 is the number of states used to construct the synthetic Alaska using the synthetic control approach (see results in online supplementary material).

For the ‘slab’ component, the conditionally conjugate pair $\beta_\rho|\sigma_\varepsilon^2 \sim \mathcal{N}(\mathbf{b}_\rho, \sigma_\varepsilon^2(\Sigma_\rho^{-1})^{-1})$ and $\frac{1}{\sigma_\varepsilon^2} \sim \mathcal{G}(\frac{v_\varepsilon}{2}, \frac{ss_\varepsilon}{2})$ is used, where $\mathcal{G}(a, b)$ is the Gamma distribution with expectation a/b and \mathbf{b} , the expected values of each element of β , is initialised to zero. Writing the full-model prior information matrix as Σ^{-1} , Σ_ρ^{-1} is the submatrix for the rows and columns corresponding to $\rho = 1$. The values ss_ε and v_ε are the prior sum of squares and prior sample size respectively. With v_ε the number of pre-intervention periods, ss_ε can be set using an expected R^2 for Equation (1), R_0^2 , using $ss_\varepsilon = v_\varepsilon(1 - R_0^2)s_y^2$, where s_y^2 is the sample variance of y in the pre-Dividend period. In this application, R_0^2 is set to the R^2 from the regression of y_t on $\hat{Y}_{1t}^N, t \in \{1, \dots, T_0\}$, where \hat{Y}_{1t}^N is the counterfactual series estimated using the synthetic control approach (again, see online supplementary materials). Setting $\Sigma^{-1} = \frac{g}{n} \{wX'X + (1-w)\text{diag}(X'X)\}$ adapts Zellner’s g-prior $\Sigma^{-1} = \frac{g}{n} X'X$ to avoid it becoming improper when $X'X$ is not positive definite. Estimation proceeds with default values of $g = 1$ and $w = 1/2$.

In Equation (2), the variance of the random walk is governed by the prior distribution $\frac{1}{\tau^2} \sim \mathcal{G}(\frac{v_\tau}{2}, \frac{ss_\tau}{2})$ where ss_τ is the prior sum of squares, and v_τ is the prior sample size, so ss_τ/v_τ is a prior estimate of τ^2 . In this application, the specification of the prior is completed (i.e. the parameters of the inverse-Gamma prior can be derived) using a prior estimate of $ss_\tau/v_\tau = 0.01s_y$. This reflects a weak expectation that the variance of the state equation will be small relative to the observation equation. Nevertheless, since it is arbitrary, sensitivity to this choice is examined by also considering the results under an alternative prior of $ss_\tau/v_\tau = 0.1s_y$.

The initial state, μ_1 , has a normally distributed prior with mean y_1 (the first observed value of the outcome) and standard deviation s_y .

Posterior simulation uses a Gibbs sampler to simulate a sequence of parameter-state combinations from a Markov chain with stationary distribution $p(\theta, \alpha|\mathbf{y}_{1:T_0})$. The sampler alternates between a data-augmentation step that simulates from $p(\alpha|\mathbf{y}_{1:T_0}, \theta)$ and a parameter-simulation step that simulates from $p(\theta|\mathbf{y}_{1:T_0}, \alpha)$. The results are based on 10,000 simulations, of which the first 500 are discarded (‘burn-in’).

3.3 | Simulating impacts

The second element of the approach is to use the posterior simulations to simulate counterfactual outcome time series. Writing the density $p(\mathbf{y}_{1:T}^0|\mathbf{y}_{1:T_0}, \mathbf{x}_{1:T})$ highlights that these simulations are conditional on pre-Dividend outcomes in Alaska but pre- and post-Dividend outcomes in other states. Simulated impacts are constructed as the difference between observed and counterfactual outcomes. Across all draws, this gives the posterior distribution of impacts.

It is important to note that these posterior predictive simulations are coherent in the sense that they are defined as a joint distribution over all time periods. That is, for each draw, the counterfactual for any period is related to that of any other period through the shared time series properties unique to that draw. This allows inference for summary statistics that relate to multiple points in time, such as cumulative effects.

The BSTS analysis was implemented using the R program *CausallImpact* (Brodersen et al., 2015), available from <http://google.github.io/CausallImpact>.

4 | ESTIMATION RESULTS

4.1 | The estimated impacts of the Dividend

In the results that follow, the years 1965–1981 constitute the pre-treatment period. This implies that impacts may be seen from 1982 onwards. A distinction is drawn between the 1982–1988 period and the years from 1989 onwards, when conditionality relating to criminals was introduced. Consequently, during the earlier (1982–1988) period, the results capture the impact on crime of the Dividend whereas from 1989 onwards, the results capture the combined impact of the Dividend and the conditionality whereby certain prisoners are ineligible.

The BSTS results are shown graphically. In Figure 3 the thick black lines trace out the estimated impacts over time (shown with 95% credible intervals). Prior to 1982, the lines should ideally be close to the x-axis since this pre-dates the Dividend and so a zero impact is expected. The extent to which this is the case provides a visual check of model fit. From 1982 onwards, the lines show the impact of the Dividend. Looking across all categories, the results provide little evidence of impact. This is true for property crime as a whole and for its subcategories. It is also true for violent crime with the occasional exception of murder, for which there is some evidence of a negative effect.

Table 1 summarises and quantifies these impacts. Before considering these, note that the penultimate column reports the root mean squared error (RMSE) over the pre-intervention period, expressed as a percentage of the estimated counterfactual. This suggests that, overall, the fit is better for property crime than violent crime. Consequently, the results for property crime may be felt to be more reliable than those for violent crime. Within this overall category, burglary and larceny show a better fit than motor theft. The final column shows the Bayes factor for model H1 ($\Delta < 0$) vs. model H2 ($\Delta \geq 0$), where Δ is the impact; a higher value constitutes evidence in favour of H1. None of the factors provides strong evidence of H1, that is, a negative effect, except in the case of murder. While intriguing, the high RMSE for murder suggests caution in reading too much into this.

For each category of crime, Table 1 provides the mean impact over the post-intervention period (1982 onwards) as well as impacts for 1982–1988 and 1989 onwards. The median across all draws of the impact (expressed as a percentage of the counterfactual outcome) is also provided. For each outcome, R_0^2 and J_0 are shown in the left hand column.

Over the full post-intervention period (1982 onwards), the number of property crimes is estimated to have been reduced by 238 per 100,000 residents per year, or 3.6%. The impact in the years 1982–1988 was positive while the impact for 1989 onwards was negative. With burglary, larceny and vehicle theft the overall impacts were -16.4% , -3.7% and -1.8% respectively. However, all were estimated with wide intervals (that spanned zero). This was also the case for violent crimes as a whole. Only in the case of murder was there a strong suggestion of a (negative) effect, although the high RMSE raises a concern about the reliability of that result.

Given its primacy among case study methods, impacts were also estimated using the synthetic control approach. The choice of predictors to incorporate in this approach was informed by the theoretical framework discussed and included pre-1982 measures of income, inequality, unemployment, education, conviction rates and crime (see web-based supporting materials for precise definitions).

Estimated impacts are provided in full detail in the web-based supplementary materials and summarised here for the purpose of comparison. In Figure 4, the thick black lines again show the impact of the Dividend from 1982. To give a sense of the prediction error, the charts include the results of separate placebo tests for each of the other 49 US states, shown as thin grey lines. These are estimated

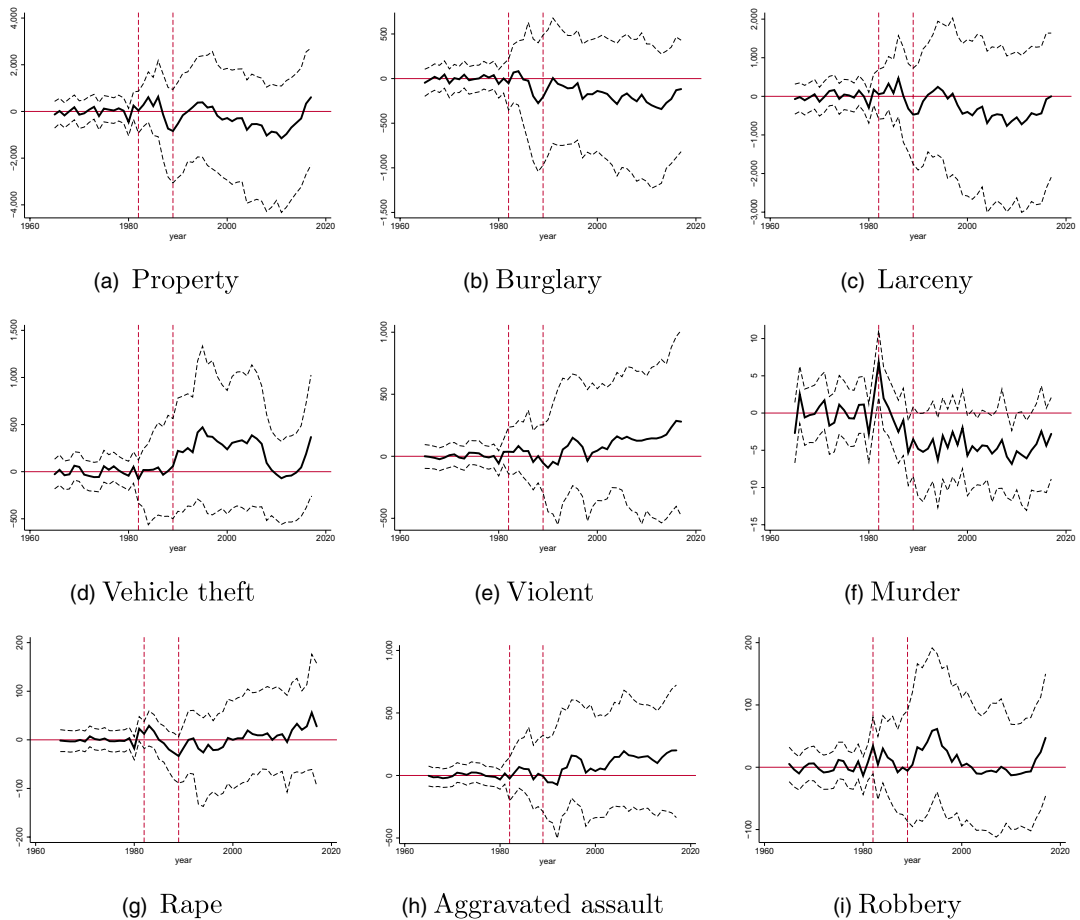


FIGURE 3 Impacts of the Dividend on the number of crimes per 100,000 population in Alaska using Bayesian structural time series to construct counterfactual outcomes on the basis of crime trends in other states. In each case, the solid black line shows the impact on the number of crimes per 100,000 population, while the dashed lines are the 95% prediction intervals. Using priors given in Table

in the same way as for Alaska, the only difference being that Alaska itself is not included among the pool of states that could possibly make up the synthetic control. They provide an indication of how likely it is that differences of the size estimated for Alaska could arise by chance. Where the actual-counterfactual difference in the Alaskan case is outside the top or bottom 2.5% of the distribution of placebo actual-counterfactual differences, it is marked with a diamond.

From Figure 4(a), it appears that the rate of property crime in Alaska prior to 1982 was mostly not markedly different from that in the synthetic Alaska, although this was not uniformly true with, for example, a difference in 1980 that was outside the range seen in other states. Motor vehicle theft shows numerous such differences in the pre-1982 period. Violent crime shows a broadly similar pattern to property crime in this period (Figure 4(e)), Murder, rape and aggravated assault all are revealed as problematic; robbery is less so.

The impression from this graphical evidence of a better model ‘fit’ for property crime than violent crime is reinforced by additional diagnostics described in the online supplementary material. This also shows that Alaska looks quite different from its synthetic counterpart in respect of some of the predictor variables, with higher unemployment and income, for example.

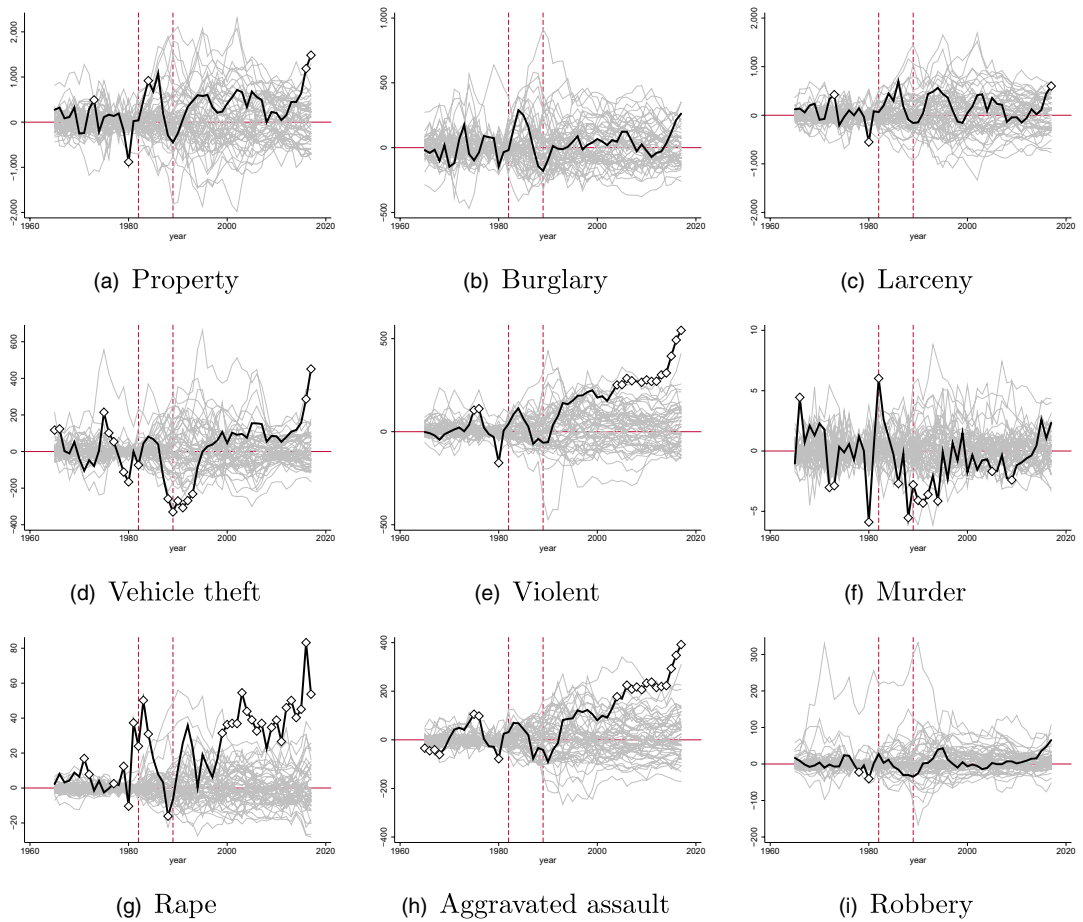


FIGURE 4 Impacts of the Dividend on the number of crimes per 100,000 population in Alaska using synthetic control approach to construct counterfactual outcomes on the basis of crime trends in other states. In each case, the solid black line shows the difference between Alaska and the synthetic Alaska in the number of crimes per 100,000 population. The thin grey lines are placebo tests for other states. Markers indicate where the Alaskan difference is outside the top or bottom 2.5% of the placebo distribution. Estimates control for the mean pre-1982 rate of the crime in question, rates 1, 3 and 5 years pre-1982 and for state characteristics over those pre-1982 years for which data are available (these characteristics include ratio of convictions to crimes, unemployment rate, educational attainment and income inequality).

With these caveats in mind, Figure 4 suggests a lack of impact on property crime except in 1984 and the last 2 years, which saw positive impacts. Burglary and larceny are the most reliable sub-categories and these mostly suggest no effect. Given the performance diagnostics discussed above, the results for violent crimes should be viewed with caution.

Comparing the BSTS impact estimates with the synthetic control impact estimates, two points are worth highlighting. First, counterfactual outcomes under BSTS resemble true outcomes more closely in the pre-Dividend years than is the case for the synthetic control approach; for property crime, the RMSEs are 3.8% and 7.8% respectively. Second, neither approach finds strong evidence of an impact on property crime.

TABLE 1 Mean and total impacts, estimated using BSTS controlling for outcome trends in other states, with crime-specific priors. The ‘Median impact, %’ column gives the median impact expressed as a percentage of the estimated counterfactual outcome. The ‘RMSE’ column expresses the root mean squared error as a percentage of the estimated counterfactual outcome. The final column shows the Bayes factor for model H1 ($\Delta < 0$) vs. model H2 ($\Delta \geq 0$). The prior odds of these hypotheses are assumed to be 1; $p(H_2)/p(H_1) = 1$. A higher value of the Bayes Factor constitutes evidence in favour of H1

		Mean impact, Δ	95% interval	Median impact, %	95% interval	RMSE	Bayes factor
Property:	1982+	-237.9	[-2512.4, 1380.9]	-3.6	[-38.2, 72.9]	3.8	1.4
$R_0^2 = 0.97$	1982–1988	137.1	[-1093.9, 1254.6]	3.3	[-16.3, 33.2]		0.6
$J_0 = 5$	1989+	-328.5	[-2934.9, 1479.9]	-5.4	[-44.4, 85.0]		1.6
Burglary:	1982+	-148.0	[-785.9, 385.9]	-16.4	[-55.9, 123.8]	3.5	2.5
$R_0^2 = 0.95$	1982–1988	-51.7	[-516.5, 366.5]	-2.6	[-30.2, 57.6]		1.2
$J_0 = 8$	1989+	-171.2	[-875.5, 409.8]	-19.9	[-63.4, 143.1]		2.8
Larceny:	1982+	-233.8	[-1901.6, 1103.0]	-3.7	[-38.7, 83.4]	4.0	1.3
$R_0^2 = 0.97$	1982–1988	105.4	[-622.3, 869.1]	3.4	[-14.4, 35.1]		0.6
$J_0 = 6$	1989+	-315.7	[-2267.2, 1211.5]	-5.9	[-45.6, 99.4]		1.5
Motor:	1982+	173.4	[-305.8, 610.6]	-1.8	[-1096.7, 1033.5]	7.4	0.3
$R_0^2 = 0.78$	1982–1988	-0.6	[-405.1, 318.5]	1.5	[-113.5, 192.0]		0.7
$J_0 = 5$	1989+	215.4	[-332.6, 707.1]	-2.7	[-1337.0, 1252.4]		0.2
Violent:	1982+	85.7	[-236.1, 479.4]	11.0	[-319.8, 464.6]	5.7	0.5
$R_0^2 = 0.93$	1982–1988	27.3	[-135.3, 231.2]	4.1	[-18.9, 78.2]		0.7
$J_0 = 5$	1989+	99.8	[-280.7, 561.0]	12.5	[-404.9, 567.4]		0.5
Murder:	1982+	-3.6	[-7.3, -1.1]	-32.4	[-49.1, -11.3]	14.0	231.3
$R_0^2 = 0.36$	1982–1988	0.0	[-2.1, 1.9]	-0.2	[-15.3, 20.8]		1.0
$J_0 = 5$	1989+	-4.5	[-8.7, -1.6]	-40.4	[-58.2, -16.2]		286.9
Rape:	1982+	3.6	[-60.9, 66.6]	9.3	[-224.9, 302.5]	12.5	0.7
$R_0^2 = 0.79$	1982–1988	0.9	[-27.6, 23.5]	3.7	[-22.3, 46.2]		0.9
$J_0 = 2$	1989+	4.2	[-70.0, 82.5]	8.6	[-276.0, 373.2]		0.7
Assault:	1982+	84.0	[-186.4, 421.8]	16.6	[-601.3, 728.4]	9.0	0.5
$R_0^2 = 0.90$	1982–1988	21.3	[-135.9, 238.9]	3.4	[-27.9, 216.8]		0.9
$J_0 = 4$	1989+	99.2	[-215.6, 475.3]	19.7	[-743.7, 870.8]		0.4
Robbery:	1982+	11.0	[-58.2, 88.8]	11.6	[-385.4, 391.7]	10.4	0.5
$R_0^2 = 0.84$	1982–1988	12.2	[-38.9, 60.6]	14.0	[-37.4, 232.9]		0.3
$J_0 = 9$	1989+	10.7	[-69.9, 102.6]	10.1	[-460.6, 438.5]		0.6

4.2 | Modelling the relationship between estimated impact and Dividend Size

The BSTS impact estimates have wide credible intervals, as apparent from Figure 3. These intervals summarise uncertainty in the estimates and correspond to the top and bottom 2.5% of simulated impacts at any point in time. What is not visible from Figure 3 is the relationship between impacts in

different years. The BSTS simulation approach provides multiple time series of impact estimates. If time series vary in their mean level, this will lead to wide intervals for point-in-time impact estimates. However, it is possible that there is more consistency across time series in how year-on-year changes in estimated impacts vary with changes in the size of the Dividend. Put differently, the imprecision of the impact estimates visible in Figure 3 may be due to draw-specific effects; controlling for these makes it possible to focus directly on the relationship between crime and Dividend amount.

In this section, a separate autoregressive distributed lag (ARDL) model is estimated for every simulated impact time series generated through the BSTS simulation. The results across all such time series are then summarised and used to illustrate the impact on crime of a one-off increase in the Dividend. Pesaran and Shin (1998) show that ARDL models can be appropriate for the examination of long-run relationships regardless of the time series properties of the individual regressors.

A generic ARDL(P,Q) model can be written

$$\Delta_t = \phi + \sum_{p=1}^P \gamma_p \Delta_{t-p} + \sum_{q=0}^Q \kappa_q D_{t-q} + v_t \quad (3)$$

where, in this application, Δ_t is the estimated impact of the Dividend on crime at time t , D_t is the (real) Dividend amount at time t , and ϕ , γ_t and κ_t are parameters to be estimated. The disturbances, v_t , are assumed to be independently and identically distributed, with zero mean. They are also assumed to be distributed independently of the regressors.

In considering this assumption, it is relevant to highlight that D_t is not influenced by contemporaneous or lagged values of Δ_t . Such a scenario could arise if the Dividend amount were calculated by simply dividing Fund earnings by the size of the eligible population. In that case, a negative impact on crime at time t would be likely to reduce the number of people sentenced or incarcerated, thereby increasing the denominator in the calculation and leading to a subsequent reduction in the Dividend. Instead, the Dividend is calculated by dividing Fund earnings *less appropriations* and reductions by the size of the eligible population. Rather than being shared out among eligible applicants, the Dividends that would otherwise have been paid to those ineligible on criminal grounds are retained by the state and used to contribute to some of the costs associated with incarceration and probation, to provide victim support and to fund grants for domestic violence and sexual assault programmes. Consequently, the numerator and denominator in the calculation of the Dividend amount are both affected, with the consequence that there is no such feedback.

Equation (4) generalises Equation (3) by allowing coefficients to differ across simulations. Introducing the superscript d to identify a unique draw, the draw-specific ARDL model is then

$$\Delta_t^d = \phi^d + \sum_{p=1}^P \gamma_p^d \Delta_{t-p}^d + \sum_{q=0}^Q \kappa_q^d D_{t-q} + v_t^d \quad (4)$$

Across all D draws, Equation (4) involves $(1 + P + (1 + Q)) \times D$ parameters. With $D = 9500$, this amounts to a large number of parameters. An ARDL model of the form of Equation (4) is estimated for every draw. The estimated coefficients will be different for every draw but looking across all draws allows the distribution of estimated γ_p^d and κ_q^d coefficients to be summarised. From Equation (4) it is apparent how a relationship between crime and Dividend amount might exist despite the earlier finding of no overall impact of the Dividend. Intuitively, the wide intervals around the impacts presented already may correspond to wide variation in the estimates of ϕ^d despite possibly precise estimates of

TABLE 2 The distribution of P and Q values that minimise the BIC

	Lags of D_t				
	1	2	3	4	5
Lags of Δ_t					
1	0.38	0.00	0.02	0.01	0.00
2	0.46	0.01	0.05	0.01	0.00
3	0.01	0.00	0.01	0.00	0.00
4	0.02	0.00	0.00	0.00	0.00
5	0.00	0.00	0.01	0.00	0.00

Cells show the proportion of times each combination of i and j minimises the BIC across 9500 simulated impact series where, for each series, Equation (3) is estimated for property crime, with i and j allowed to vary between 1 and 5.

TABLE 3 Results of estimating the ARDL(2, 1) specification of Equation (4) for property crime, varying priors

	Using priors given in Table 1	Using alternative priors
γ_1	0.96 (0.70, 1.23)	0.98 (0.62, 1.29)
γ_2	-0.18 (-0.44, 0.10)	-0.18 (-0.50, 0.17)
κ_0	-0.45 (-0.77, -0.16)	-0.47 (-0.82, -0.14)
κ_1	0.23 (-0.16, 0.51)	0.23 (-0.11, 0.55)
ϕ	140.31 (-327.88, 540.69)	44.29 (-465.67, 490.24)

The results in the first column are based on the preferred impact estimates, using priors described in Table 1. The results in the second column use default *CausalImpact* priors apart from using $ss_\tau/v_\tau = 0.1s_y$ to specify the inverse-Gamma prior on τ^2 rather than $ss_\tau/v_\tau = 0.01s_y$. Each reported coefficient is the median across 9,500 estimates; the 95% interval shown in parentheses corresponds to the smallest and largest 2.5% of estimates.

γ_p^d and κ_q^d . Equation (4) offers the potential to control for draw-specific fixed effects and thereby make the relationships of interest more visible.

Estimation used simulated time series over the period 1982–2017. While this amounts to only 36 years of data in each case, Monte Carlo results provided by Pesaran and Shin (1998) suggest good small-sample performance of ARDL estimators.

A preliminary analysis was carried out to examine the required order of the ARDL model. ARDL(i, j) models were estimated with $i = 1, \dots, 5, j = 1, \dots, 5$ for each draw from the posterior distribution for property crime impacts. The preferred ARDL(P, Q) specification for each draw was chosen as having P, Q equal to the i, j combination that achieved the smallest Bayesian Information Criterion (BIC) for that draw. The joint distribution of P and Q is shown in Table 2. This suggests an ARDL(2, 1) specification is sufficiently flexible in 84% of cases. Furthermore, the estimated coefficients are shown to be quite robust to the choice of ARDL specification (see the online supplementary material).

Using this preferred specification, the results of estimating Equation (4) for property crime for each draw from the posterior are summarised in Table 3. Each reported coefficient is the median across

TABLE 4 Results of estimating Equation (4), specified as ARDL(2, 1)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Prop	Burg	Larc	Motr	Viol	Murd	Rape	Asst	Robb
γ_1	0.82 (0.55, 1.08)	0.64 (0.33, 0.97)	1.08 (0.67, 1.42)	0.80 (0.44, 1.16)	0.09 (-0.25, 0.47)	0.64 (0.16, 0.85)	0.79 (0.46, 1.16)	0.77 (0.42, 1.12)
γ_2	0.02 (-0.28, 0.29)	0.11 (-0.17, 0.39)	-0.25 (-0.56, 0.18)	-0.00 (-0.33, 0.39)	0.47 (-0.09, 0.68)	0.03 (-0.24, 0.26)	-0.05 (-0.40, 0.33)	0.03 (-0.30, 0.36)
κ_0	-1.93 (-3.53, -0.68)	-1.14 (-2.26, 0.04)	-0.05 (-0.45, 0.66)	0.01 (-0.37, 0.26)	0.00 (-0.02, 0.01)	-0.01 (-0.11, 0.03)	-0.01 (-0.43, 0.25)	-0.03 (-0.12, 0.03)
κ_1	-0.26 (-1.05, 0.73)	-0.28 (-0.86, 0.44)	0.02 (-0.24, 0.34)	-0.07 (-0.26, 0.08)	-0.00 (-0.01, 0.00)	-0.01 (-0.06, 0.02)	-0.06 (-0.20, 0.10)	-0.01 (-0.07, 0.05)
ϕ^c	-703.24 (-1578.34, 90.10)	-498.49 (-1272.89, 166.93)	129.09 (-155.84, 428.56)	26.35 (-192.16, 163.42)	0.79 (-14.69, 5.91)	5.10 (-84.12, 42.77)	9.17 (-147.40, 141.94)	-11.53 (-60.20, 32.48)
κ_0^c	1.50 (0.35, 2.92)	0.82 (-0.23, 1.86)	-0.05 (-0.86, 0.36)	-0.06 (-0.29, 0.34)	-0.00 (-0.01, 0.02)	-0.01 (-0.05, 0.10)	-0.01 (-0.26, 0.40)	0.03 (-0.03, 0.10)
κ_1^c	0.38 (-0.66, 1.14)	0.35 (-0.33, 0.92)	0.00 (-0.30, 0.31)	0.09 (-0.06, 0.24)	0.00 (-0.01, 0.00)	0.02 (-0.00, 0.06)	0.05 (-0.11, 0.17)	0.01 (-0.05, 0.06)
ϕ	956.99 (184.49, 1809.68)	169.74 (-181.73, 439.89)	-35.69 (-286.00, 266.76)	19.87 (-127.27, 246.37)	-1.61 (-4.90, 5.13)	0.77 (-31.46, 55.13)	27.75 (-99.47, 278.11)	20.12 (-20.62, 74.34)

Columns relate to different types of crime as follows: (1) property crime (2) burglary (3) larceny (4) motor vehicle theft (5) violent crime (6) murder (7) rape (8) aggravated assault (9) robbery. Using priors given in Table 1. Each reported coefficient is the median across 9,500 estimates; the 95% interval shown in parentheses corresponds to the smallest and largest 2.5% of estimates.

9500 ARDL estimates, and the 95% interval corresponds to the lowest and highest 2.5% of these estimates. For each draw from the posterior predictive distribution, the estimated ARDL coefficients represent a transformation of the time series that constitutes that draw. The coefficients thereby define new parameters that summarise the relationship between impact and Dividend amount. Looking across all draws gives the intervals that appear in Table 3 and later tables. These intervals ignore the fact that the draw-specific coefficients are themselves estimates. Instead, they capture uncertainty arising from posterior variation in the time series and the fact that the transformation applied to each time series is itself draw-specific.

Two columns of results are shown in order to provide some reassurance that the results are not unduly sensitive to the choice of priors assumed by the BSTS approach. The results in the first column are based on the preferred impact estimates, using the priors described in Table 1. The results in the second column are based on $J_0 = 3$ and $R_0^2 = 0.8$ (the default in the *CausalImpact* program) but use $ss_\tau/\nu_\tau = 0.1s$, to specify the inverse-Gamma prior on τ^2 (which implies more variability in the trend term relative to the preferred specification). The column 2 estimates are less precise than the preferred results (the RMSE is 5.4 rather than 3.8), yet the estimated relationship with the Dividend amount is much the same as before; an estimated κ_0 of roughly -0.45 and an estimated κ_1 of 0.23 . Hence, introducing more imprecision into the impact estimates does not appear to alter the overall finding. Specifically, increasing the Dividend amount reduces property crime. However, the direct effect does not last beyond the year of increase, as indicated by the imprecise κ_1 and κ_2 coefficients. Instead, the contemporaneous effect transmits to later years through γ_1 and γ_2 .

Since the preliminary analysis suggests the ARDL(2, 1) specification is preferred on the basis of the BIC in the majority of cases and given that the results seem robust to the choice of specification, an ARDL(2, 1) specification is used in the remainder of this paper.

To allow for the relationship between impacts and Dividend size to change following the introduction of conditionality, a generalised version of Equation (4) is estimated

$$\Delta_t^d = \phi^d + \sum_{p=1}^P \gamma_p^d \Delta_{t-p} + \sum_{q=0}^Q \kappa_q^d D_{t-q} + \sum_{q=0}^Q \kappa_q^{dC} D_{t-q} C_{t-q} + v_t^d \quad (5)$$

where $C_t = 1$ if $t \geq 1989$, 0 otherwise such that κ_q^{dC} captures the extent to which the impact of the Dividend changes under conditionality.

The results are presented in Table 4. For property crime (column 1), there is a negative contemporaneous effect of Dividend size (the κ_0 coefficient) but only a weak lagged effect. With the introduction of conditionality, the contemporaneous effect changes by κ_0^C . This is smaller in absolute size than κ_0 but is positive, suggesting that under conditionality the impact of Dividend size is reduced. Again, the lagged effect is weak. Looking within sub-categories of property crime, no strong effect of the Dividend is seen. The same is true for violent crime and all its sub-categories.

The results in Table 4 provide the econometric underpinning of the estimated relationships between crime and Dividend amount. However, it is helpful to present this information in a way that shows the evolution over time of the estimated effects. Figures 4 and 5 do this using impulse response functions (IRFs) which show the effect of a one-off \$100 Dividend increase, contemporaneously and over the subsequent 10 years. As with the results above, these figures are based on 9500 estimated ARDL models. For each draw from the posterior distribution of impacts, the estimated ARDL coefficients are used to calculate the dynamic effect of the \$100 impulse. To illustrate, for draw d , the contemporaneous effect in the absence of conditionality is calculated as $\Delta_0^d = 100\kappa_0^d$, the impact after 1 year is $\Delta_1^d = \gamma_1^d \Delta_0^d + 100\kappa_1^d$ and the impact after $t > 1$ years is $\Delta_t^d = \gamma_1^d \Delta_{t-1}^d + \gamma_2^d \Delta_{t-2}^d$

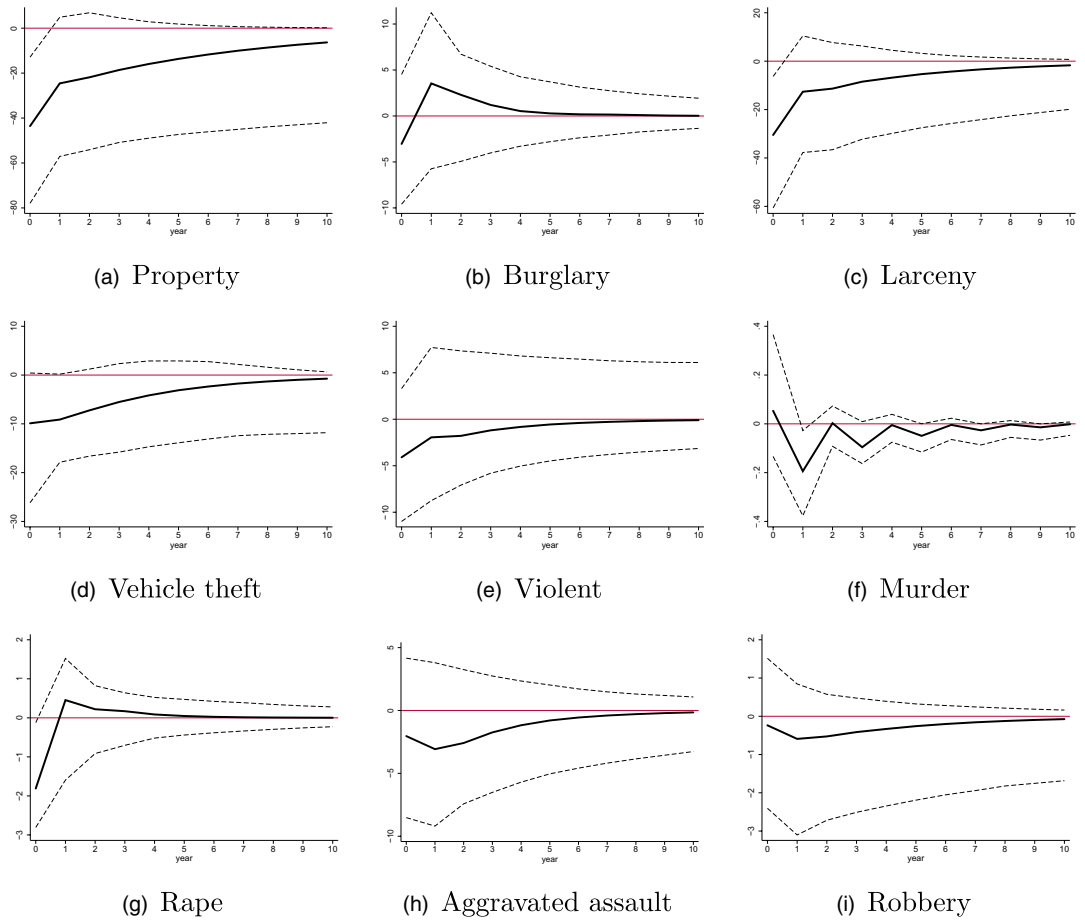
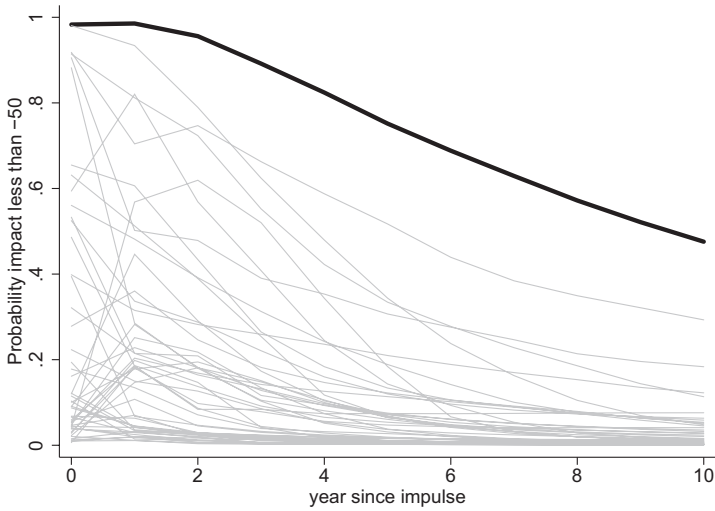


FIGURE 5 Evolution of effect of one-off \$100 Dividend increase on the number of crimes per 100,000 population, without conditionality. Estimated using ARDL(2, 1) model. Dashed lines show 95% prediction intervals.

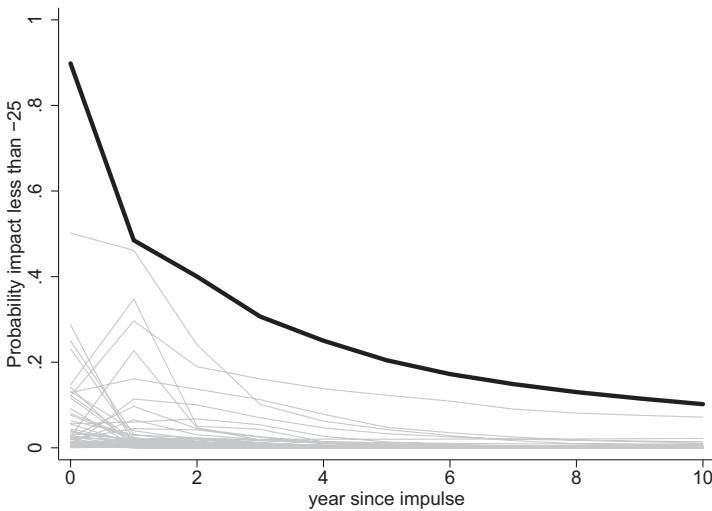
. IRFs under conditionality are derived as $\Delta_0^d = 100(\kappa_0^d + \kappa_0^{dC})$, $\Delta_1^d = \gamma_1^d \Delta_0^d + 100(\kappa_1^d + \kappa_1^{dC})$ and $\Delta_t^d = \gamma_1^d \Delta_{t-1}^d + \gamma_2^d \Delta_{t-2}^d$, as before. Figures 4 and 5 summarise the calculated IRFs across all 9500 draws, plotting the median, along with 95% intervals.

Figure 4 relates to the case where there is no conditionality (i.e. it sets $C_{t-q} = 0, \forall q$ in Equation (5)). It is readily apparent that the strongest effects are for property crime. Consistent with the regression results in Table 4, the immediate effect of the \$100 increase is to reduce by 193 the number of property crimes per 100,000 population. One year later, the reduction is 184 and the following year it is 157. Beyond that point, the effect slowly decays such that 10 years later the estimated impact is a reduction of 46. The impact appears to be driven mainly by the reduction in larcenies. For all other categories, the results suggest impacts that are weaker and close to zero by the end of the 10-year period.

Figure 5 shows the IRFs under conditionality. Again the strongest results are found for property crime. However, the reduction in crime resulting from the Dividend increase is smaller when conditionality is in place than when it is not in place (Figure 4). Substantively, there is no evidence from this analysis that denying eligibility to offenders reduces crime.



(a) Without conditionality - probability of reduction of 50 crimes per 100,000 residents



(b) With conditionality - probability of reduction of 25 crimes per 100,000 residents

FIGURE 6 Probability of a one-off \$100 Dividend increase reducing property crime by at least 50 per 100,000 residents without conditionality (top panel) or 25 per 100,000 residents with conditionality (bottom panel).

4.3 | Assessing robustness of the BSTS results

To provide reassurance that the result for Alaska is not spurious, the analysis was repeated for all other states, with property crime as the outcome. For each state, the BSTS impact was estimated using all other states (excluding Alaska) as the basis for estimating counterfactual outcomes, and the ARDL(2, 1) model of Equation (5) was estimated for each draw of the posterior distribution. Since the Dividend is

not paid in any states other than Alaska, the expectation is that any suggestion of an impact from these placebo IRFs should only ever arise chance. The results are given in the web-based supporting materials and confirm that in the majority of states the credible intervals mostly span the x-axis throughout the period considered.

To take account of the size of effect, the upper panel of Figure 6 plots the probability over time (in the no-conditionality case) that a one-off \$100 Dividend increase causes a reduction of more than 50 property crimes per 100,000 residents. Alaska is shown by the thick black line and thin grey lines show placebo results for other states. In the spirit of permutation tests, the placebo results provide the null distribution, and the results for Alaska can be seen to be quite distinct, with a reduction of 50 or more crimes consistently more likely than elsewhere. The bottom panel shows the results with conditionality, this time focusing on a reduction of 25 property crimes per 100,000 residents. Again, the probability of such a reduction is higher in Alaska.

Overall, these results show that the estimated impacts of a Dividend increase are stronger in Alaska than in other states. Clearly, since the Dividend only exists in Alaska, this is in line with expectations as the impacts estimated in other states are of a placebo. The fact that the results for Alaska stand out so prominently provides additional evidence that the main results are capturing a true relationship.

5 | CONCLUSION

This paper has examined the impact of Alaska's Permanent Fund Dividend on crime. Despite no strong evidence of a negative effect overall, the results suggest the rate of crime responds to the size of the Dividend, with higher payments reducing property crime.

The overall impact estimates have low statistical power due to high variance. The BSTS estimation approach simulates the posterior distribution of counterfactual outcome time series, from which a posterior distribution of estimated impact time series can be readily constructed. By estimating the relationship between impact and Dividend amount for each draw from this distribution, it has been possible to control for draw fixed effects and thereby achieve more precise estimation of the relationship of interest.

Substantively, the results imply that increasing the Dividend reduces property crime for an extended period. This may appear to contrast with Watson et al. (2019) who find the (negative) marginal effect on property crime of an increase in the Dividend to be non-significant. However, as noted already, their analysis addresses a different question—the effect of Dividend *receipt* rather than the Dividend *per se*—and sensitivity to the size of the Dividend is considered only in the week of receipt, rather than over a longer period.

While the evidence supports the finding of a negative impact on crime, there is a further question regarding whether the effect is of a size that is meaningful. Further manipulation of the posterior distribution suggests that, over 10 years, a hypothetical one-off \$100 hike in the Dividend reduces the number of property crimes by 1200 per 100,000 residents in the no-conditionality case and 166 per 100,000 residents in the conditionality case. Hence, for the cost of the Dividend to be recouped through averted property crimes alone, these would have to average over \$8000 or \$60,000 in the respective cases. Since this is unlikely to hold, it is difficult to argue that increasing the Dividend is warranted on crime-prevention grounds alone. Rather, the reduction in property crime is best viewed as a beneficial side effect of the Dividend.

Two final points about the estimated relationship should be emphasised. First, the posterior distribution of impact time series was constructed without using information on the Dividend amount. This is required if the relationship estimated by the ARDL models is to be meaningful in the sense of not

merely reproducing assumed features of the underlying model. Instead, the results are more likely to be detecting a true effect.

Second, aside from the fact that ARDL models can be given a causal interpretation, the basis for viewing the relationship as causal is further strengthened by the fact that Dividend amount is plausibly exogenous in its relationship to crime. Nevertheless, it should be borne in mind that the credibility of the ARDL estimates depends on the ability to estimate impacts of the Dividend itself. BSTS uses estimated coefficients based on pre-Dividend years to form counterfactual outcomes from the observed outcomes of selected other states in post-Dividend years. Should the relationship between outcomes in Alaska and other states be disrupted by factors other than the introduction of the Dividend, the credibility of these impact estimates would reduce. As an identifying assumption, this cannot be directly verified. However, the fact that the ARDL results for Alaska are stronger than for any other state provides some reassurance that a true effect has been captured.

With these points in mind, the findings demonstrate the potential of the Dividend to influence an important social outcome. As such, it provides evidence that may be considered alongside the other potential benefits of a basic income. In this case, a higher payment level results in a lower rate of property crime. Removing eligibility from criminals does not appear to reinforce this effect as one might expect if criminals are forward-looking. In view of this, the change to eligibility does not seem to act as a deterrent to crime.

ORCID

Richard Dorsett  <https://orcid.org/0000-0002-4180-8685>

REFERENCES

- Abadie, A., Diamond, A. & 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. & Gardeazabal, J. (2003) The economic costs of conflict: A case study of the Basque Country. *American Economic Review*, 93, 113–132.
- Becker, G.S. (1968) Crime and punishment: An economic approach. *Journal of Political Economy*, 76(2), 169–217.
- Bharat, S. & UNICEF. (2014) A little more, how much it is...Piloting basic income transfers in Madhya Pradesh, India, New Delhi.
- Blakeslee, D.S. & Fishman, R. (2018) Weather shocks, agriculture, and crime evidence from India. *Journal of Human Resources*, 53(3), 750–782.
- Brodersen, K.H., Gallusser, F., Koehler, J., Remy, N., Scott, S.L., et al. (2015) Inferring causal impact using bayesian structural time-series models. *The Annals of Applied Statistics*, 9(1), 247–274.
- Choe, J. (2008) Income inequality and crime in the United States. *Economics Letters*, 101(1), 31–33.
- Evans, W.N. & Moore, T.J. (2011) The short-term mortality consequences of income receipt. *Journal of Public Economics*, 95(11), 1410–1424.
- Goldsmith, S. (2012) The economic and social impacts of the permanent fund dividend on Alaska, In Widerquist, K. & Howard, M.W.(Eds.) *Alaska's Permanent Fund Dividend: Examining its Suitability as a Model*. London: Palgrave.
- Haarmann, C. (2009) Making the difference!: The BIG in Namibia: Basic Income Grant pilot project assessment report, April 2009, NANGOF.
- Hoynes, H. & Rothstein, J. (2019) Universal basic income in the United States and advanced countries. *Annual Review of Economics*, 11, 929–958.
- Hsieh, C.-T. (2003) Do consumers react to anticipated income changes? Evidence from the Alaska Permanent Fund. *American Economic Review*, 93, 397–405.
- Hum, D. & Simpson, W. (1993) Economic response to a guaranteed annual income: Experience from Canada and the United States. *Journal of Labor Economics*, 11(1, Part 2), S263–S296.
- Jones, D. & Marinescu, I. (2018) The labor market impacts of universal and permanent cash transfers: Evidence from the Alaska Permanent Fund, Working Paper w24312, National Bureau of Economic Research.

- Kangas, O., Jauhiainen, S., Simanainen, M., Ylikännö, M.. (2019) The basic income experiment 2017–2018 in Finland. preliminary results.
- Labour Party. (2019) It's time for real change: The labour party manifesto 2019.
- Lee, D. & McCrary, J. (2017) The deterrence effect of prison: Dynamic theory and evidence. In Cattaneo, M. & Escanciano, J.C. (Eds.) *Regression Discontinuity Designs: Theory and Applications*, Advances in Econometrics, Vol. 38, pp. 73–146.
- Martinelli, L. (2019) A basic income trilemma: Affordability, adequacy, and the advantages of radically simplified welfare. *Journal of Social Policy*, 49, 461–482.
- More, T. (1516 [1963]) *Utopia*. Harmondsworth: Penguin Classics. English translation by Paul Turner.
- Murray, C. (2016) *In our hands: A plan to replace the welfare state*. Washington DC: Rowman & Little-field.
- Office of the United Nations High Commissioner for Human Rights. (2020) COVID-19 Guidance, Technical report.
- Pesaran, M.H. & Shin, Y. (1998) An autoregressive distributed-lag modelling approach to cointegration analysis. *Econometric Society Monographs*, 31, 371–413.
- Robins, P.K. (1985) A comparison of the labor supply findings from the four negative income tax experiments. *Journal of Human Resources*, 20, 567–582.
- Scott, S.L. & Varian, H.R. (2014) Predicting the present with Bayesian structural time series. *International Journal of Mathematical Modelling and Numerical Optimisation*, 5(1), 4–23.
- Sloman, P. (2018) Universal basic income in british politics, 1918–2018: From a ‘vagabond's wage’ to a global debate. *Journal of Social Policy*, 47(3), 625–642.
- Watson, B., Guettafi, M. & Reimer, M. (2019) Universal cash and crime. *Review of Economics and Statistics*, 102, 1–45.
- Wikström, P.-O.H. (2007) *In search of causes and explanations of crime. In Doing research on crime and justice*. Oxford: Oxford University Press, pp. 117–139.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.
Supplementary Material

How to cite this article: Dorsett R. A Bayesian structural time series analysis of the effect of basic income on crime: Evidence from the Alaska Permanent Fund*. *J R Stat Soc Series A*. 2020;00:1–22. <https://doi.org/10.1111/rssa.12619>