The author(s) shown below used Federal funding provided by the U.S. Department of Justice to prepare the following resource:

Document Title:	The Interaction and Impacts of State DNA Database Laws, Final Summary Overview
Author(s):	Jennifer L. Doleac
Document Number:	251035
Date Received:	July 2017
Award Number:	2013-IJ-CX-0001

Final Summary Overview:

The Interaction and Impacts of State DNA Database Laws

Jennifer L. Doleac *

Final Draft date: February 2017

This project was supported by Award Number 2013-IJ-CX-0001 awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. The points of view in this document are those of the author and do not necessarily represent the official position or policies of the U.S. Department of Justice.

* Frank Batten School of Leadership and Public Policy, University of Virginia, Charlottesville, VA 22904. Email: jdoleac@virginia.edu.

1 Purpose

Despite the widespread use of criminal offender DNA databases by the law enforcement community, relatively little is known about their real-world effects on criminal behavior. The purpose of DNA databases is to quickly and accurately match crime scene evidence with known offenders. They aim to reduce crime by increasing the probability of punishment, conditional on offending. They also exhibit large returns to scale. For these reasons, DNA databases are likely to be far more cost-effective than traditional law enforcement tools. Indeed, previous research has shown that DNA databases have large crime-reducing effects at the state level (Doleac, 2015).

However, because DNA database expansions are – like most crime policies – legislated at the state level, the resulting policies may not be efficient from a national perspective. Each state will weigh its own costs and benefits when making policy decisions, without regard for their effects on other states. But state DNA databases do not exist in a vacuum: they could have effects on criminal behavior elsewhere. The federal government can improve the national efficiency of state-level policies by inducing states to consider the costs and benefits to their neighbors – i.e., to internalize the externalities.

The federal government has invested a large amount of money in helping states expand their databases and clear sample backlogs, presumably because it perceives positive externalities to state database expansions. Positive externalities would exist if offenders are highly mobile, and are deterred or incapacitated upon being added to a state database. If they stop committing crime, this would reduce crime in other states. Positive externalities could also exist if profiles from one state help other states catch and incapacitate offenders. One important aspect of DNA database policy in the United States is that states can search for profile matches across state lines, using

CODIS. This should increase the positive externalities of state-level expansions. For instance, if Georgia expands its database, neighboring states such as Florida could benefit from being able to search those additional profiles. Depending on how often crimes in Florida are committed by out-of-state offenders, the expansion of Georgia's database could help Florida identify and convict more offenders, reducing its own crime rate. When Georgia is weighing the costs and benefits of expanding its database, it will probably not consider these external benefits, and so its database will be inefficiently small; hence the justification for federal funding to encourage more expansions.

However, there could be negative externalities, particularly if a state's database expansion induces probable offenders to move to another state. For instance, suppose Georgia adds felony arrestees to its database, but Florida does not include that group. Offenders in Georgia who think they could be arrested for a felony will have an incentive to move to Florida, to avoid DNA profiling and therefore detection for past or future crimes. (Such a behavioral effect is not unreasonable to expect: we typically assume that human capital is mobile, and that workers move to places with better jobs.) In this case, Georgia's database expansion would displace some of its own crime to Florida. As before, Georgia would not include this external cost in its cost-benefit analysis, but in this case a state database that seems optimal from the state's perspective will be inefficiently large from the national perspective.

Which of these externalities dominates in the DNA database context is an empirical question. The goal of this project is to answer that question. This report summarizes the main findings. Specifically, I address the following research questions:

1. Does the addition of an offender or arrestee profile to CODIS by one state help or hurt other states, in terms of crime outcomes?

2. Do these effects vary with distance from the state that added the profile?

3. Do these effects vary with the type of offender who is added?

4. Do these effects depend on whether a state allows partial-match searches of CODIS?

2 Project Design and Methods

The primary goal of this project is to test the effect of other-state DNA profiles on own-state crime rates. I follow Doleac (2015) and use an instrumental variable strategy, to exploit exogenous variation in state database size. That study shows that the timing of state database expansions is random with respect to underlying crime trends in those states. However, the implementation of those laws (the rate at which qualifying offenders' profiles are analyzed an uploaded to CODIS) might not be, which could introduce omitted variable bias to simple OLS estimates of crime rates on database size. In addition, state database size and crime rates are simultaneously determined; this biases OLS estimates upwards. I use the timing of state database expansions, along with pre-period convict, arrest, and incarceration rates, to estimate the number of qualifying offenders in each state in each year. I then use that number as an instrument for actual database size in each state. This isolates the "good"/random variation in database size from the "bad" variation stemming from differences in implementation, removing the omitted variable bias. Using pre-period rates also removes the simultaneity bias.

2.1 Data

Part of this project involved expanding the dataset collected in Doleac (2015). Historical data on state database sizes come from state crime labs, FBI statistics, and media reports. I include every state-year that information on database size was available, focusing on the years 2000 to 2014.

To construct instrumental variables, I estimate the number of qualifying offenders, using pre-period (1999) state-level data on arrests, convictions, and prison populations, as well as information on the timing of state database expansions. Arrest estimates are based on the 1999 Uniform Crime Reports (UCR) multiplied by 1999 clearance rates, by crime type. Conviction estimates are based on those arrest estimates, multiplied by the share of arrestees who are ultimately convicted, by crime type, in the 2000 State Court Processing Statistics. State prison populations are from the 1999 National Jail Census.

Information on the timing and details of state database expansions comes directly from state legislative histories. Information on which states conduct partial-match searches (which are typically not legislated) comes from Ram (2009), the FBI website, and media reports. For states listed as allowing partial-match searches in Ram's review, I used 2009 (the year of Ram's data collection) as the effective date unless a specific date was available elsewhere.

I construct three measures of other-state database sizes as the treatment: (1) total other-state profiles, (2) other-state profiles in nearby states only (those within 500 miles), and (3) total other-state profiles divided by the distance in miles to each state (that is, weighted by the inverse of the distance between states, so that nearby states get more weight). For the distance between states, I used Google maps to determine the shortest driving distance, in miles, from each state to each other state.

For outcome measures, I use Uniform Crime Report (UCR) data on reported crime and National Incident-Based Reporting System (NIBRS) data on arrests in reported crimes. In the latter, I create an indicator variable for whether an arrest was made in a particular case, then (because the effect sizes are very small) multiply that indicator by 100 so that the estimated effects are on a 0-100 scale instead of 0-1.

All crime, database size, and qualifying offender variables are converted to rates, so that the numbers used are per 100,000 own-state residents.

Summary statistics are in Table 1.

2.2 Empirical Strategy

I first consider the effect of the number of other-state profiles in CODIS on own-state crime outcomes. I use each of the three treatment variables described above, in an instrumental variable (IV) framework.

The baseline 2SLS IV specification is:

$$\text{Own-state outcome}_{i,s,t} = \beta_1 + \beta_2 * \text{Other-state profiles}_{s,t} + \beta_3 * \text{Own-state qualifying offenders}_{s,t} + \gamma_{State} + \delta_{Year*Region} + \mathrm{e}_{i,s,t}, \quad (1)$$

where *Other-state profiles* is estimated in the first-stage regression:

$$\text{Other-state crime}_{i,s,t} = \alpha_1 + \alpha_2 * \text{Other-state qualifying offenders}_{s,t} + \alpha_3 * \text{Own-state qualifying offenders}_{s,t} + \gamma_{State} + \delta_{Year*Region} + \mathrm{e}_{i,s,t}, \quad (2)$$

In both specifications, i indexes crime types, s indexes states, and t indexes years. State fixed effects, γ_{State}, control for average differences in state crime rates. Year-by-region fixed effects, $\delta_{Year*Region}$, control for annual variation in average crime rates, at the Census region level. The identification therefore comes form within-region-year variation in other-state profiles. In all specifications, I control for own-state qualifying offenders (the instrument, not actual database size), to account for own-state policy effects on own-state crime. *Own-state outcome* is either a crime rate or probability of making an arrest in new cases. Standard errors are clustered at the state level in the crime analysis, and at the police department level in the arrest analysis. The coefficient of interest is β_2, which measures the marginal effect of an other-state DNA profile on own-state crime outcomes.

First stage results are presented in Table 2. Each instrument (qualifying offenders) is highly correlated with the actual number of profiles from other states. F-statistics range from 1,199 to 426,897, implying that the instruments are extremely strong.

After measuring baseline effects, I consider the differential effects by type of other-state profile. Of particular interest in the current policy environment is whether arrestee profiles are more or less valuable than convict profiles. To test this, I interact the treatment (profiles) and the IV (qualifying offenders) with an indicator of whether arrestees were included in the relevant other-state database in that year.

I also consider whether the effect of other-state profiles depends on whether states allow searching for partial matches in CODIS. (Partial-match searches could make each profile more valuable by, for instance, leading investigators to a family member of someone in the database.) To do this, I interact the treatment and IV with an indicator of whether the focal state allowed partial-match searches.

Note that CODIS did not allow arrestee profiles or partial-match searches until 2006, which provides useful variation in addition to state-level policy changes.

Finally, I test for effects of other-state profiles on the probability of making arrests in own-state cases. If CODIS is helping states solve more crimes, we would expect other-state profiles to have a positive effect here.

3 Summary of Results

The effects of other-state profiles on own-state crime are show in Table 3, using the instrumental variable strategy described above. The first panel shows effects on violent crime (murder, forcible rape, aggravated assault, and robbery). The second panel shows effects on property crime (burglary, larceny, and motor vehicle theft).

In all specifications I control for the number of own-state qualifying offenders (the instrument for own-state database size). Note that the effect of own-state qualifying offenders is both negative and statistically significant in all cases. While the magnitude of the coefficient cannot be directly compared with the magnitudes of the other coefficients in the table (because it is the instrument, not actual database size), this implies that adding additional offenders to a state database reduces crime in that state. This is consistent with the findings of Doleac (2015).

Column 1 shows the effect of a marginal other-state profile, without regard to which other state it was from. The effect is positive and statistically significant. It implies that an other-state profile *increases* own-state violent crime by 0.0001 offenses, and own-state property crime by 0.0003 offenses.

Column 2 restricts attention to the effects of nearby states – that is, those less than 500 miles away. If the mechanism is offenders traveling between states, we would expect the effect of nearby states' databases to be larger than that of more distant states' databases. Indeed, that is what we find. An additional profile in nearby states' databases increases own-state violent crime by 0.0011 offenses, and own-state property crime by 0.0063 offenses – larger effects than we saw in Column 1.

Column 3 considers the same issue slightly differently, using a distance-weighted measure of other states' database sizes. Each other state's database is multiplied by the inverse of the distance to the state of interest, so that states farther away therefore get less weight than states nearby. For instance, an additional profile in a state 5000 miles away will count as 0.0002 profiles, while an additional profile 100 miles away counts as 0.01 profiles. We again see that other-state profiles have a positive and statistically-significant effect on both violent crime and property crime. The coefficients suggest that an additional profile 100 miles away increases own-state violent crime by 0.0012 offenses and property crime by 0.0042 offenses, while an additional profile 5000 miles away increases own-state violent crime by 0.00003 offenses and property crime by 0.00008 offenses.

Tables 4 and 5 show these effects separately by specific crime type. We see that other-state profiles consistently have a positive effect on own-state crime in all cases, though the proximity of the state matters more for the crimes of burglary and larceny. For those crimes, only profiles from nearby states have a statistically-significant effect on own-state crime.

Table 6 considers the differential effects of other-state profiles when arrestees are included in other-state databases, or when own-state policies permit CODIS searches for partial DNA matches. Columns 1–3 replicate the main results with an interaction term for whether arrestees are included in each state's database. Overall, those interaction terms are not significant. This implies that arrestee profiles have the same effect on other-state crime that convict profiles do. However, it is important to note here that I do not observe how many arrestee profiles are actually in each state's database. I only observe whether arrestees were part of the overall total. Having more detailed information on the composition of state databases would allow more precise estimates.

Column 4 tests for differential effects of other-state profiles when a state is allowed to search for partial matches in CODIS. This should make each other-state profile more valuable in reducing crime. Indeed, I find that a marginal other-state has a less-positive effect on both violent and property crime when states can use it to look for partial matches. The coefficient on the interaction term is negative and statistically significant, largely cancelling out the positive coefficient on other-state profiles. However, given the apparent rarity of partial match searches, and the more controversial nature of such searches, it seems likely that this significant effect is due to those states' more aggressive use of forensic evidence and CODIS, more broadly. Without more information about the relative frequency of partial match searches at the state level, I urge caution in interpreting this result as due to those searches themselves. However, this does suggest that states have some power to mitigate the crime-increasing effects of other-state profiles.

Finally, Table 7 shows the effect of other-state profiles on the probability of making arrests in own-state crimes. Recall that the outcome is measured on a 0-100 scale (instead of 0-1) to make the table more readable. There are no statistically significant effects in any specification.

4 Implications for Criminal Justice Policy and Practice

The cross-state effects of DNA database policies are consistent with the hypothesis that offenders are mobile and respond rationally to states' crime-reduction strategies. In particular, it appears that a DNA expansion in one state incentivizes probable offenders in that state to move elsewhere, which increases crime rates in other states. In other words, DNA databases' negative effects on own-state crime appear partly due to shifting crime to other states.

Doleac (2015) estimated that a marginal own-state profile reduces own-state violent crime by 0.047 offenses and own-state property crime by 0.328 offenses. Here I estimate that a marginal other-state profile increases own-state crime by 0.0001 violent offenses, and 0.0003 property offenses. To get the net effect of a marginal own-state profile on national crime, we can multiply the other-state effect by 49 and add the own-state effect. This suggests that a marginal DNA profile reduces violent crime by 0.042 offenses, and property crime by 0.313 offenses. That is still a large net benefit, though it is smaller than the benefit perceived by the state uploading the profile.

This has implications for federal DNA database policy. As long as state database laws vary, adding additional offenders in one state will have negative externalities on other states. Individual states weighing the costs and benefits of database expansions will therefore settle on databases that are inefficiently large. Federal programs that decrease the costs by funding state expansions therefore exacerbate this problem. To reduce these negative externalities, the federal government should stop subsidizing state-level expansions, or standardize DNA database policies nationwide so that offenders have no incentive to move.

These results have broader implications for U.S. criminal justice policy and the evaluation of policy effects. Most crime policy is determined at the state level, and the social efficiency of those policies is partly dependent on their effects on other states. If criminal offenders were not mobile, effects would be negligible, and states should reach efficient outcomes independently. But this study suggests effects are not zero. Research on policy effects should be more careful to consider external effects of state and local policies; in particular, to what degree are local reductions in crime due to shifting crime elsewhere? (Addressing this issue can be difficult since detailed data are often unavailable at the national level, so this study argues for improved efforts to make such data available.) The federal government should then use its funding and policy options to help states internalize any externalities.

References

Doleac, J. L. (2015, January). The effects of DNA databases on crime. Working paper. Available at http://jenniferdoleac.com/wp-content/uploads/2015/03/Doleac_DNA_Databases.pdf.

Ram, N. (2009). DNA confidential: State law enforcement policies for genetic databases lack transparency. *Science Progress*. Available at http://scienceprogress.org/2009/11/dna-confidential/.

A Appendix: Tables

Table 1: Summary Statistics

	N	Mean	SD
Year	660	2007	4.3238
Violent crime rate	660	390.53	161.37
Property crime rate	660	3161.2	799.06
Other states' profiles	660	172498	278713
Nearby states' profiles	660	13420	19562
Distance-weighted profiles	660	170.73	242.60
Other states' qualifying offenders	660	372609	518556
Nearby states' qualifying offenders	660	29225	35873
Distance-weighted qualifying offenders	660	371.87	446.66
Own-state qualifying offenders	660	3760.6	3111.4

Table 2: First Stage for Instrumental Variable Regressions

	Other states' profiles	Nearby states' profiles	Distance-weighted profiles
Other states' qualifying offenders	0.6321***		
	(0.0010)		
Nearby states' qualifying offenders		0.6087***	
		(0.0176)	
Distance-weighted qualifying offenders			0.6218***
			(0.0067)
F-statistic	426897	1199	8656

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are shown in parentheses.

Table 3: Effect of Other-State Profiles on Own-State Crime

	(1)	(2)	(3)
Violent Crime			
Other states' profiles	0.0001***		
	(0.0000)		
Nearby states' profiles		0.0011***	
		(0.0003)	
Distance-weighted profiles			0.1273***
			(0.0417)
Own qualifying offenders	-0.0155***	-0.0172***	-0.0148***
	(0.0033)	(0.0042)	(0.0032)
2000 violent crime rate	426.83		
Observations	660	660	660
Property Crime			
Other states' profiles	0.0003**		
	(0.0002)		
Nearby states' profiles		0.0063***	
		(0.0021)	
Distance-weighted profiles			0.4241**
			(0.1998)
Own qualifying offenders	-0.0743**	-0.0711**	-0.0706*
	(0.0377)	(0.0349)	(0.0387)
2000 property crime rate	3558.8		
Observations	660	660	660

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are shown in parentheses.

Table 4: Effect of Other-State Profiles on Own-State Violent Crime

	(1)	(2)	(3)
Murder			
Other states' profiles	0.0000^{***}		
	(0.0000)		
Nearby states' profiles		0.0000^{**}	
		(0.0000)	
Distance-weighted profiles			0.0015^{***}
			(0.0003)
Own profiles	-0.0001^{***}	-0.0001^{***}	-0.0001^{***}
	(0.0000)	(0.0000)	(0.0000)
Rape			
Other states' profiles	0.0000^{***}		
	(0.0000)		
Nearby states' profiles		0.0001^{***}	
		(0.0000)	
Distance-weighted profiles			0.0117^{***}
			(0.0037)
Own profiles	-0.0004^{*}	-0.0006^{**}	-0.0003^{*}
	(0.0002)	(0.0002)	(0.0002)
Assault			
Other states' profiles	0.0001^{**}		
	(0.0000)		
Nearby states' profiles		0.0006^{**}	
		(0.0003)	
Distance-weighted profiles			0.0840^{**}
			(0.0341)
Own profiles	-0.0108^{***}	-0.0123^{***}	-0.0104^{***}
	(0.0029)	(0.0036)	(0.0029)
Robbery			
Other states' profiles	0.0000^{***}		
	(0.0000)		
Nearby states' profiles		0.0004^{***}	
		(0.0001)	
Distance-weighted profiles			0.0301^{***}
			(0.0066)
Own profiles	-0.0042^{***}	-0.0043^{***}	-0.0040^{***}
	(0.0008)	(0.0008)	(0.0007)
Observations	660	660	660

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are shown in parentheses.

Table 5: Effect of Other-State Profiles on Own-State Property Crime

	(1)	(2)	(3)
Burglary			
Other states' profiles	0.0000		
	(0.0000)		
Nearby states' profiles		0.0012^{**}	
		(0.0005)	
Distance-weighted profiles			0.0632^{*}
			(0.0383)
Own profiles	-0.0129^{**}	-0.0116^{*}	-0.0122^{*}
	(0.0065)	(0.0061)	(0.0066)
Larceny			
Other states' profiles	0.0001		
	(0.0001)		
Nearby states' profiles		0.0035^{***}	
		(0.0013)	
Distance-weighted profiles			0.1706
			(0.1414)
Own profiles	-0.0435^{*}	-0.0387^{*}	-0.0411^{*}
	(0.0243)	(0.0229)	(0.0250)
Vehicle Theft			
Other states' profiles	0.0002^{***}		
	(0.0000)		
Nearby states' profiles		0.0017^{**}	
		(0.0007)	
Distance-weighted profiles			0.1903^{***}
			(0.0455)
Own profiles	-0.0179^{**}	-0.0207^{**}	-0.0173^{*}
	(0.0087)	(0.0081)	(0.0088)
Observations	660	660	660

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are shown in parentheses.

Table 6: Differential Effects of Other-State Profiles on Own-State Crime

	(1)	(2)	(3)	(4)
Violent crime				
Other states' profiles	0.0001**			0.0001***
	(0.0000)			(0.0000)
Other states' profiles * Arrestees included	0.0000			
	(0.0001)			
Nearby states' profiles		0.0010**		
		(0.0004)		
Nearby states' profiles * Arrestees included		0.0001		
		(0.0006)		
Distance-weighted profiles			0.0765	
			(0.0468)	
Distance-weighted profiles * Arrestees included			0.0733	
			(0.0839)	
Other states' profiles * Partial match allowed				-0.0001***
				(0.0000)
Own profiles	-0.0155***	-0.0173***	-0.0149***	-0.0144***
	(0.0033)	(0.0043)	(0.0032)	(0.0033)
Observations	660	660	660	660
Property crime				
Other states' profiles	0.0002			0.0005***
	(0.0004)			(0.0001)
Other states' profiles * Arrestees included	0.0001			
	(0.0006)			
Nearby states' profiles		0.0061***		
		(0.0021)		
Nearby states' profiles * Arrestees included		0.0004		
		(0.0037)		
Distance-weighted profiles			0.6055*	
			(0.3449)	
Distance-weighted profiles * Arrestees included			-0.2621	
			(0.5739)	
Other states' profiles * Partial match allowed				-0.0004***
				(0.0001)
Own profiles	-0.0742**	-0.0713**	-0.0701*	-0.0685*
	(0.0377)	(0.0347)	(0.0388)	(0.0361)
Observations	660	660	660	660

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are shown in parentheses.

Table 7: Effect of Other-State Profiles on Own-State Arrest Probability

	(1)	(2)	(3)
Violent Crime			
Other states' profiles	-0.0000		
	(0.0000)		
Nearby states' profiles		-0.0001	
		(0.0000)	
Distance-weighted profiles			-0.0033
			(0.0049)
Own qualifying offenders	-0.0000	-0.0001	-0.0001
	(0.0005)	(0.0006)	(0.0005)
2000 violent crime arrest probability	37.877		
Observations	2555344	2555344	2555344
Property Crime			
Other states' profiles	0.0000		
	(0.0000)		
Nearby states' profiles		-0.0000	
		(0.0000)	
Distance-weighted profiles			0.0021
			(0.0027)
Own qualifying offenders	-0.0000	-0.0001	-0.0000
	(0.0003)	(0.0003)	(0.0003)
2000 property crime arrest probability	12.099		
Observations	20307252	20307252	20307252

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are shown in parentheses.