Flypaper or Attractive Recycling?

The Department of Defense 1033 program and local government spending*

Donald J. Bruce, Celeste K. Carruthers, Matthew C. Harris †, Matthew N. Murray and Jinseong Park ‡

University of Tennessee§

February 17, 2017

Abstract

This paper examines the local public finance impact of one of the largest grant-in-kind program in U.S. history, the 1033 program, which allowed local police departments to receive decommissioned military goods. Several features of this program are unique among intergovernmental grants. Unlike virtually all other intergovernmental grants, these goods (rather than cash) are not fungible, oversight by non law-enforcement officials may be non-existent, and may not be included in local public budgets and capital accounts. While previous research shows that intergovernmental grants result in either perfect crowd-out or a partial flypaper effect, we find no evidence of crowd-out and some evidence for crowding-in. The features of this program may therefore be useful when designing future grants to increase local spending in a targeted category, but welfare is likely tempered by the absence of local oversight.

JEL classification: TBD

Keywords: Grant in-kind, Flypaper, Crowding Out

---

*We sincerely thank the Defense Logistics Agency, including Carlos Torres, Ken MacNevin, Michelle McCaskill and Susan Lowe, for their generosity with their time and for educating us on the institutional details and context of the 1033 Program necessary for identification. All errors are our own.

†Corresponding Author: 722A Stokely Management Center, 916 Volunteer Boulevard, Knoxville, TN 37996-0570; mharris@utk.edu, (865) 974-5591

‡Corresponding Author: 720A Stokely Management Center, 916 Volunteer Boulevard, Knoxville, TN 37996-0570; jpark61@utk.edu, (865) 974-5592

§Haslam College of Business, Department of Economics and Boyd Center for Business and Economic Research.

I Introduction

Intergovernmental transfers are a key source of funds to local governments. The National League of Cities reports that approximately 5 percent of municipal budgets comes through the federal government, while an additional 30-40 percent of local government revenues comes from states. The motivation for and structure of grant programs are both important because these grants will affect the allocation of local public resources and thus welfare. Grants may be used to redistribute resources from high-income to low-income areas, in turn addressing differentials in local tax capacity. In other instances, local government provision of a public good may be preferable to state or federal provision, either due to dis-economies of scale or heterogeneity in preferences between local jurisdictions. In this case, providing grants to local governments for the provision of a particular service may generate more welfare than the federal government providing the service directly. Two other rationales for grants are of particular interest to our research question. First, grants are also widely used to address interjurisdictional spillovers, as with transportation infrastructure. Second, there may be paternalistic preferences on the part of higher-levels of government in a federalist system. If local preferences lead to too little spending on a public good, the higher-level government may choose to encourage more provision through grants.

The underlying motivation for a grant matters because it determines how the grant should be structured to maximize welfare. For example, if the goal is simply to redistribute income, a lump sum grant would yield the largest welfare gain to local citizens. However, such a grant would result in near-perfect crowding-out through the local budget process. On the other hand, if the problem is under-provision of a public good or the need to address spillovers, the grant should be designed to promote stickiness. In this case, crowding out is highly undesirable and may defeat the welfare objectives of the grant program.

The empirical literature shows that no matter how grant funds are earmarked, they typically crowd out recipient spending to some extent. At one end of the spectrum, there is perfect crowding out. A seminal work by Bradford and Oates (1971) details the conditions under which grants are equivalent to a lump-sum tax reduction for the residents of the recipient government. For example, Lutz (2010) examines the fiscal consequences of New Hampshire's school finance reform in 1999 and shows that 90 cents per grant dollar is used to lower the tax burden of local residents. Lutz attributes this finding to unique features around the reform, including the state's direct democracy that allows for perfect reflection of median voter preferences, perfect information on the part of the voters regarding the reform, fiscal autonomy of the state, and fungibility of the grants. Most of the early empirical literature on intergovernmental grants showed that some significant share of grant proceeds tend to stick where they hit, a result dubbed the 'flypaper effect' (Hines and Thaler, 1995). Recent empirical literature employing better identification strategies or exploiting quasi-experimental settings provides evidence of substantial, but incomplete, crowding-out. The extent to which grants stick or crowd out local spending has been studied in a number of settings, including Title 1 grants for education (Gordon, 2004; Cascio et al., 2013), health care (Baicker and Staiger, 2005), highway funding (Knight, 2002), and law enforcement (Baicker and Jacobson, 2007; Evans and Owens, 2007). There is also a literature on the governmental or grant features that make income equivalence unlikely, usually contrasting institutional features that differ from the conditions considered by Bradford and Oates (Filimon et al., 1982; Strumpf, 1998; Payne, 2009; Glaeser, 2012; Brooks et al., 2011).

While most economic studies of intergovernmental transfers have centered on monetary transfers, grants in-kind are of particular interest because of their uniqueness and potential to have unique impacts on recipient government spending. There are numerous examples of in-kind transfers from the public sector directly to individuals, as with food stamps and housing vouchers. There are also examples of in-kind transfer from one level of government to another level of government. Examples include the Morrill Act that established land-grant universities and emergency response equipment provided to local governments in response to natural disasters. However, we are aware of no empirical research that has explored these in-kind grants from government to government.

We contribute to the literature on the crowding-out effects of grants by exploring the impact of the federal 1033 program, which provided surplus military gear to local governments, on county-level law enforcement spending. Since 1997, Section 1033 of the National Defense Authorization Act has allowed for the transfer of surplus U.S. military equipment to local law enforcement agencies (LEA) at no cost. Originally, the program was intended as a 'recycling program' whereby decommissioned and paid-for capital initially used to provide national defense was re-purposed for the provision of public safety. From 1997 through 2014, the department of defense transferred over $5.2 billion dollars in equipment to local law enforcement agencies, making it the largest federal-to-local grant-in-kind of capital goods of which we are aware.

The 1033 program offers a unique context in which to examine the effects of intergovernmental grants because it has the near-opposite features of the New Hampshire school finance reform noted above Lutz (2010). First, items received through the 1033 program are not fungible; they cannot be sold or transferred to other local governments. Our application is the first to our knowledge that investigates the fiscal impacts of intergovernmental grants that take the form of goods rather than income. Additionally, decisions to acquire items through the 1033 program are made solely by police chiefs, without any local institutional oversight, public input or even a signature from a city or county government official. The preferences of law enforcement executives are thus pivotal and may or may not reflect those of the voting populace. As a result, the median voter may be irrelevant. The opacity of the process implies that city or county officials with presumptive budgetary authority may have incomplete information about the amount of equipment that police chiefs acquire through the 1033 program. These features of the 1033 program are likely to result in less crowding out than lump sum or even matching monetary grants.[1]

To empirically evaluate the relationship between 1033 program receipts and local public spending on police protection, we use panel data on county expenditure accounts from the Annual Survey of Governments and data on the value of 1033 equipment transfers from the Defense Logistics Agency. Exploiting within-county variation in the value of 1033 goods received over time, we find that the value of 1033 receipts either has null effects on local spending (i.e., it sticks where it hits and there is no evidence of crowding out) or some evidence of a crowd-in of additional funding to local law enforcement. Results from a level-level specification indicate that a dollar increase in the value of items received through the 1033 program leads to an approximate \$1.50 increase in police spending in the following year. This result stands in sharp contrast to nearly all of the findings from prior literature which find evidence of at least some crowding out.[2] However, in no specification do we find any statistically significant results that indicate any crowding out. Results from our log-log specification are positive, but just below the bar for statistical significance.

[1] Evidence from the literature on Title I grants shows that even matching grants exhibit considerable crowding out (Gordon, 2004; Cascio et al., 2013).

[2] A back-of-the-envelope calculation by Brooks and Phillips (2008) suggests that the share of the Community Development Block Grant money spent is well over 100% in all subsequent years. However, the authors are reluctant to put too much weight on the point estimates.

While grants through the 1033 program are much less likely to reflect endogenous voter preferences than those sought by elected officials, we acknowledge the concerns of Knight (2002) about the endogeneity of grant receipts. To address concerns about omitted variables such as endogenous voter preferences, or time-varying heterogeneity in law enforcement leadership driving both police budgets and 1033 receipts, we exploit exogenous variation in time and transaction costs faced by departments when acquiring these items as in Harris et al. (2017). Results from our log-log fixed effects instrumental variables model indicate that a one percent increase in the value of receipts from the 1033 program leads to a 0.53 percent increase in additional funds to local police in the following year. While the log specification is our preferred model, we also report results from a level-level FE-IV specification for completeness. However, in the level-level specification our instruments are borderline weak.[3] Nevertheless, our level-level FE-IV results, while imprecise are positive and larger in magnitude than without instruments, providing no evidence of crowding out.

To further explore the crowding-in effect, we examine heterogeneous effects from 1033 receipts. Note first that crowding-in may occur if the items received through the 1033 program require complementary inputs. Suppose, for example, that a tactical truck may help police departments provide public safety, but only if there is a place to store the vehicle and there are trained personnel who know how to maintain it. Of the various types of equipment considered, vehicles are the most likely to require complementary inputs. Splitting the 1033 receipts into vehicle and non-vehicle items, we find that while the associative crowding in effects are attributable to vehicles, the IV estimates indicate that both vehicle and non-vehicle acquisitions lead to crowding in. Additionally, we might expect crowding-in to be more pronounced in areas which place relatively high value on 'law and order.' To evaluate this possibility, we split the counties in our sample into those which voted Democrat/Republican in the 2008 presidential election, operating under the stereotype that conservative, Republican counties place more value on law enforcement. In both the FE and the FE-IV specification, our estimates are positive and marginally significant for the Republican subsample, not but for the Democratic counties.

We believe these novel results are attributable to the unique features of the 1033 program, rather than special features of law enforcement. Both Evans and Owens (2007) and Baicker and Jacobson (2007) examine intergovernmental transfers to law enforcement and find significant, but

[3]First-stage F-statistics on our exclusion restrictions range between 7.0 and 9.7 for our primary sample.

partial crowding out. Our findings suggest policy implications for designing intergovernmental transfers that minimize crowd out. Specifically, if the Federal government wants to increase local resources available for other key areas (e.g., public education, libraries, parks and recreation) a distribution network of supplies available upon request directly to local program administrators may lead to less crowding out than even the most restrictive earmarking practices. Unfortunately, the welfare implications of such a policy remain unclear. Developing the conditions under which such an in-kind alternative would be preferable to matching monetary grants would depend on local preferences and the design of the aid-in-kind program, and is an area for future work.

II Relevant Background on 1033 Program

The 1033 Program was created as a part of the National Defense Authorization Act of 1997. The stated purpose of section 1033 was to enable the Department of Defense to transfer military equipment no longer in use to local LEAs to assist in drug interdiction. While the 1033 program has garnered considerable attention for transferring tactical equipment (e.g., assault rifles and armored personnel carriers) to local law enforcement, it has also facilitated transfers of clothing, ice chests, first aid kits, flashlights, etc. In actuality, over 70 percent of the items transferred were of a non-tactical nature. The 1033 program was not designed to be a grant-in-kind program, but rather a well-intentioned recycling program that takes items that are no longer useful for national defense and reallocates them to the production of public safety.

Several features of the 1033 program stand in sharp contrast to most intergovernmental transfers. First, the items transferred are non-fungible goods of a very specific nature. The Defense Logistics Agency (DLA) forbids transferring items acquired through the 1033 program to other agencies.[4] Second, it is impossible to overstate the administrative ease of participating in this program. There is a simple two-page form to sign up. There is then a one page form to request non-tactical items, a one page form to request an armored vehicle, a one-page form to request an aircraft, etc. The highest-ranking signature blank on these forms is that for the chief of police or a law enforcement officer of similar rank. Unlike Title 1 grants, for example, these transfers are not overseen or administered by the state. While each state does have a 'coordinator', the function of

[4]For example, a police department cannot request utility trucks and gift them to the county ambulance service.

this role is to facilitate communication rather than control of coordinate the use of funds or gather data. Similarly, there is no oversight at the local level. Therefore the acquisition process does not operate under the auspices of county or municipal governments who make funding decisions, nor is there any other form of public oversight. Rather, these acquisitions may occur with or without the knowledge of the local budgetary authority. While federal monetary grants to local governments must be 'accounted for' and therefore included in the budgetary process, that is not the case with the 1033 program.

These features create conditions under which transfers can have an ambiguous effect on local budgets for law enforcement. If local voters and elected officials are aware of equipment transfers through the 1033 program, they may choose to reduce police budgets in subsequent years as is the case with civil asset forfeitures (Baicker and Jacobson, 2007). This would be especially likely for transfers that are a close substitute for items that could be acquired from private sector vendors. Alternatively, if voters and elected officials are (relatively) unaware of 1033 transfers, these transfers will likely have no effect on police budgets. However, there are at least two mechanisms by which transfers through the 1033 program may actually increase police budgets, or lead to crowding in. First, many state programs contain clauses that equipment must be used in some form, or returned to the DLA. While these clauses are essentially toothless (no burden of proof is required), they would give law enforcement officials leverage to request budget increases. Second, for certain sorts of equipment there may be a 'complementary inputs' effect, where the items cannot be engaged to produce public safety without additional inputs. Therefore, the receipt of items through the 1033 program increases the marginal returns to discretionary resources allocated to police departments, thus justifying the increased spending until net marginal benefits equalize across funding categories.[5]

II.A Field Activity Centers and Program Logistics

Due to the unique structure of the 1033 program, we believe receipts through this program are less vulnerable to the sorts of endogeneity concerns first addressed by Knight (2002). Nevertheless,

[5]For example, if a police department receives a large military vehicle, they may need a storage facility and a diesel mechanic. With these inputs in place, the vehicle may contribute substantially toward public safety; without the complementary inputs, the vehicle will simply sit idle. The same analogy can be drawn for guns requiring ammunition and training time, or non-tactical items requiring storage space.

we perform the fixed effects instrumental variable (FEIV) estimation using the same identification and instrumental variables approach as Harris et al. (2017) as a check on our finding of no crowd-out.

Our identification strategy relies on two sources of exogenous variation. First, the amount of equipment available for distribution through the 1033 Program varies exogenously over time. Specifically, as documented in Harris et al. (2017), distributions of tactical items increased sharply in 2006 when the M-16 was replaced by the M-4 carbine as the standard issue weapon for the Army and Marine Corps. Additionally, the draw down from Iraq and Afghanistan exogenously increased the amount of equipment of all types available by an order of magnitude. Second, law enforcement agencies face time-invariant, location-varying exogenous differences in the time and logistical costs of acquiring said items. The interaction of these two factors (variation in availability and transaction costs) yields variables that affect the cost of acquiring goods through the 1033 program but are uncorrelated with bureaucratic and voter preferences for public safety or other unobservable factors that determine police budgets. While we refer the reader to Harris et al. (2017) for an in-depth discussion of these instruments, we present an abridged version below to frame this application.

According to the DLA, most items transferred through the 1033 Program are routed through one of 18 Field Activity Centers (FACs) starred in Figure 1. From an agency perspective, the DoD and DLA have incentives to place their disposition centers in cities that will minimize their own operating and transportation costs. The location of FACs is not influenced by local preferences for public safety. The DLA has operated since 1961, and most FACs have been in their current location since the 1970s. Additionally, police departments are not the sole, or even primary customers for the DLA Disposition Services Reutilization program, of which 1033 is a specific category. Determining which FAC receives decommissioned items is done solely on the basis of proximity to that unit. Once an FAC has received a piece of equipment, police departments have a very short (14 day) window to make a claim on that item and 21 days to take possession of any items claimed. The relatively short claim and pickup windows heighten the importance of the accessibility aspect of proximity. Therefore, proximity of a county to FACs is entirely a function of DoD cost minimization and institutional structure. The distribution of equipment between FACs is also driven by logistics and locations of military units. The amount of equipment made available through the 1033 program is determined by military needs. Finally, the distance of a police department to multiple FACs affects acquisition costs as the DoD does not ship between FACs to reduce costs for 1033 customers. These costs are determined by factors unrelated to local unobserved heterogeneity in local preferences for public safety.

The DLA also states that, particularly for vehicles, preference will be given to jurisdictions with larger land areas. Land area is a blank on the vehicle request form for the 1033 program. Ceteris paribus, counties whose law enforcement officers have more ground to cover will likely gain greater benefit from the use of military vehicles. This is an important source of variation, given that vehicles account for approximately half of the total value of goods distributed through the 1033 program. While DLA told us that they were able to meet the needs of LEA customers over time (meaning there was no long-term two-way selection problem), land area was a consideration in determining who got priority in making claims at a given point in time.

The interaction of land area or proximity variables with total value of equipment distributed through the 1033 program in a given year serves as our main instrument in the empirical specification. We enrich this approach based on an interview with a DLA spokesman who indicated that the DoD encourages localities with large land areas to participate in the 1033 program. Thus, we also interact total value of equipment with the log of land area of a county and use the interaction as another instrument in FEIV model.

III Data

The main source of local finance data is the Census of Governments and the Annual Survey of Government Finances (ASG) from 2005 to 2012, collected by the U.S. Census Bureau. While the Census Bureau canvassed the universe of government units in 2007 and 2012, data for other years are based on voluntary-response surveys of the same government units. The data provide detailed information on the revenues and expenditures of different levels of government. We primarily focus on police expenditures of county governments since this is the finest unit of government for which all necessary data are available. Police protection expenditures comprise spending on "police patrols and communications, crime prevention activities, detention and custody of persons awaiting trial, traffic safety, and vehicular inspection."[6]

[6] This is the definition provided by the Census (https://www.census.gov/govs/local/definitions.html).

Although the ASG data are widely used in the public finance literature, there are some known issues that must be acknowledged. First, any zero values in the ASG data can either represent a true value or a missing value; differentiating between the two is not possible. Fortunately, only 1.3 percent of the sample reported zero expenditure for police protection. Assuming that the data from counties that consistently reported zeros for police protection are valid, those dubious zeros account for less than one percent of the observations. Second, the county-year panel in the ASG data is not balanced. If survey participation decisions of governments are not random conditional on the control variables, our estimates would be biased. We therefore use the balanced panel as our primary estimation sample.

Data on transfers of military equipment to local law enforcement agencies come from the DLA. The data contain information on agencies to which transfers were made, item name and corresponding National Stock Number (NSN), shipping date, the quantity and the acquisition value of the item. The data that we use in the analysis were last updated in September, 2015 and coded at the individual agency level. Thus, we infer county information from the agency name and the state to which the agency belongs by combining the data with the Law Enforcement Agency Identifiers Crosswalk file created by the Bureau of Justice Statistics (BJS) and the National Archive of Criminal Justice Data (NACJD). We first combine two datasets using character-merge and then manually match observations that have low matching scores or cannot be matched automatically. Observations from the agencies whose county information cannot be identified (less than 5 percent of all agencies or 351 of 7,323 agencies), are dropped from the sample.

We also collect information on county characteristics from other sources. Data on crime rates are obtained from the county-level Uniform Crime Reports (UCR), published by the Federal Bureau of Investigation (FBI) and reproduced by the NACJD. Using agency-level data provided by the FBI, NACJD imputes missing data and aggregates the data by county.[7] We use aggregate counts of arrests for murder, rape, robbery, aggravated assault, burglary, and other assaults. The Census provides Intercensal Population Estimates, which includes demographic information for counties. Using these estimates, we calculate the share of the population that is male, the share of individuals aged 15 to 24, and an index of racial diversity similar to Alesina et al. (1999). The

[7]For more details regarding the imputation, refer to the NACJDs Uniform Crime Reporting Program Resource Guide (https://www.icpsr.umich.edu/icpsrweb/content/NACJD/guides/ucr.html).

racial diversity measure is defined as $diversity_{jt} = 1 - \sum_k race_{kjt}$, where $race_{kjt}$ represents the population share of a particular race k, k consists of white, black or African American, American Indian and Alaska Native, Asian and Native Hawaiian/Other Pacific Islander, two or more races, in county j in year t. Data on household median income and unemployment rate are obtained from the Small Area Income and Poverty Estimates of the Census and the Bureau of Labor Statistics, respectively. Finally, we collect data on the 2008 U.S. presidential election from the Guardian.[8]. The data contains the number of votes that each presidential candidate received by county.

Table 1 lists summary statistics for both our primary estimation balanced sample and an unbalanced panel. Note that lagged 1033 item values and police expenditures are expressed in per capita terms and crime rates are expressed per 100,000 population. All monetary values are adjusted for inflation using CPI-U-RS and given in thousands of 2012 U.S. dollars. The first column shows the mean and the standard deviation of each variable for the balanced panel and the second column shows the same statistics for the unbalanced panel.

The unbalanced panel contains about 24% more observations than the balanced panel. While the statistics of all other variables look quite similar across the panels, the average population size of the unbalanced panel is substantially smaller than the average population size of the balanced panel. Since small county governments show greater participation in the primary census years (2005 and 2012), the difference is not surprising. In the balanced panel, the average county spends about $91.4 per capita for police protection and receives 1033 items that amount to 8 cents per capita. While that number appears small, it is an artifact of irregular participation in the 1033 program. While most counties participated in the 1033 program in at least one year, most counties did not participate in the 1033 program in a given year. Conditional on receipt, the average item value per capita jumps to 52 cents, or 0.56 percent of total police spending.

[8] The data can be obtained from the following webpage (https://www.theguardian.com/news/datablog/2009/mar/02/us-elections-2008).

Figure 1: Heatmap: Log Tactical Items Received and Proximity to Disposition Center

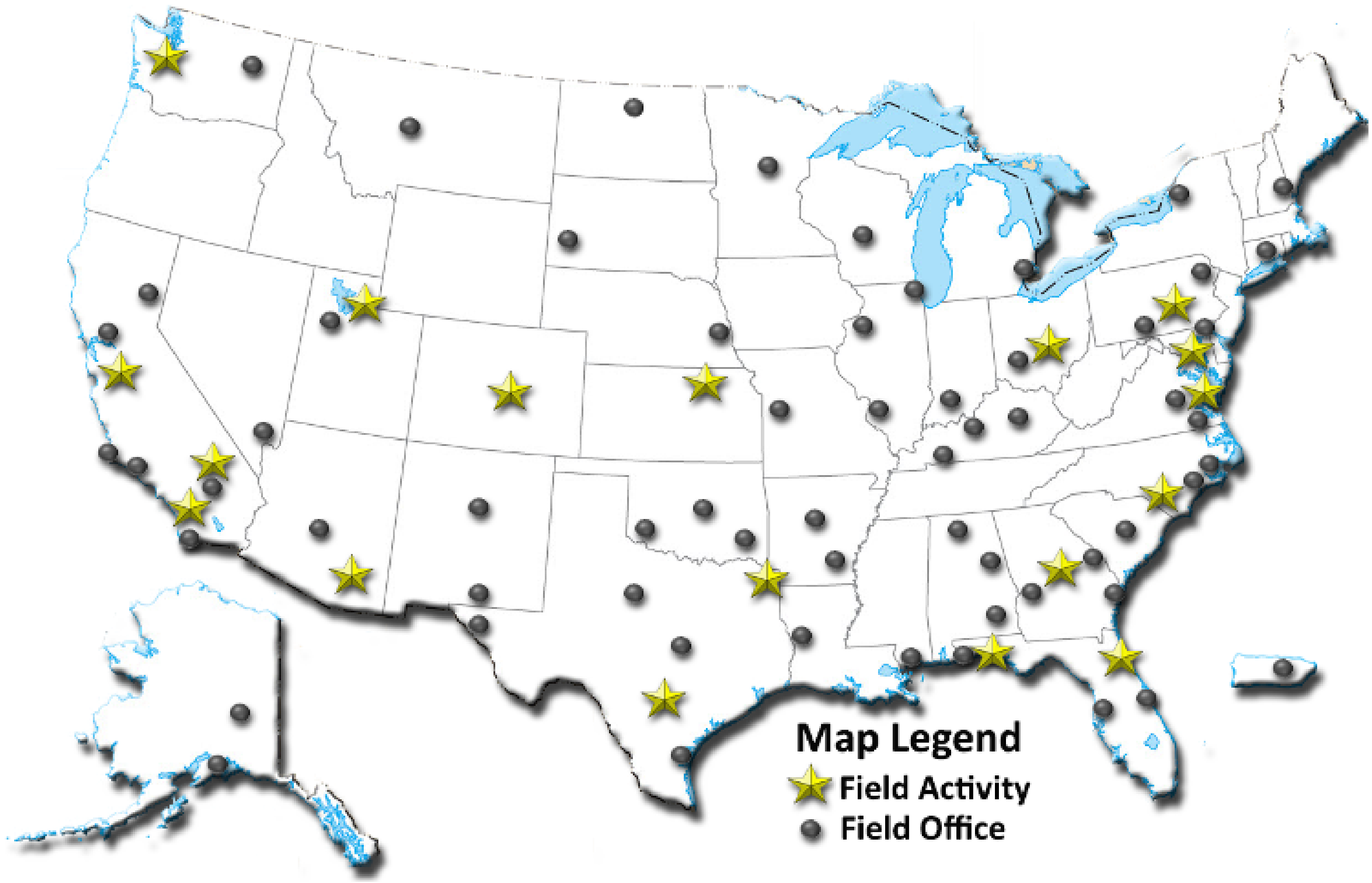

Table 1: Summary statistics

	Balanced		Unbalanced	
	Mean	Std. Errors	Mean	Std. Errors
Lagged 1033 item value per capita ($)	0.08	(0.63)	0.08	(0.69)
Police expenditure per capita ($)	93.57	(84.72)	96.47	(116.74)
Population	171714	(434478)	143063	(394533)
Median income	46.63	(12.20)	45.51	(11.86)
Male (%)	49.57	(1.55)	49.69	(1.69)
Age 15 to 24 (%)	13.81	(3.57)	13.58	(3.50)
Diversity index (% point)	22.39	(15.95)	21.22	(15.99)
Poverty (%)	15.14	(6.08)	15.46	(6.28)
Unemployment rate (%)	7.12	(3.04)	7.09	(3.04)
Lagged murder per 100,000 population	1.15	(3.38)	1.17	(3.66)
Lagged rape per 100,000 population	2.28	(5.20)	2.38	(5.81)
Lagged robbery per 100,000 population	3.96	(9.71)	3.67	(9.25)
Lagged aggravated per 100,000 population	35.03	(70.75)	35.45	(69.81)
Lagged burglary per 100,000 population	29.37	(46.01)	30.23	(47.11)
Lagged assault per 100,000 population	120.00	(179.75)	120.79	(175.02)
Observations	9471		11749	

Note: The balanced panel consists of county governments that fully participated in the Annual Survey of State and Local Government Finances (the Census of Governments in year 2005 and 2007) from 2005 to 2012 and the unbalanced panel consists of all county governments that at least once participated in the survey. Both balanced and unbalanced sample consist of county-level law enforcement agencies that are matched to county governments. Data section provides detailed description of each variable. All dollars are inflation-adjusted using CPI-U-RS and expressed in 2012 U.S. dollars.

IV Empirical Model

We estimate the following equation to evaluate the effect of 1033 items acquisition on police protection expenditures of county-level LEAs:

$$Protection_{j,s,t} = \beta_0 + \beta_1 Values_{j,t-1} + \mathbf{\Gamma} \boldsymbol{X_{j,t}} + \theta_j + \theta_s t + \varepsilon_{j,t} \tag{1}$$

where $Protection_{j,t}$ represents police protection expenditure of county j in year t and $Values_{j,t-1}$ represents a monetary value of items that the law enforcement agency of the government of county j received through the 1033 program in year $t-1$. Note that both $Protection_{j,t}$ and $Values_{j,t-1}$ are normalized to the county population and expressed in per capita terms. $\boldsymbol{X_{j,t}}$ is a vector of time-varying county-level characteristics including median household income, the share of male population, the share of population aged 15 to 24, racial diversity index, poverty rate, and unemployment rate. θ_j and $\theta_s t$ are county fixed effects and state-specific time trends, respectively.[9] Standard errors are clustered at the county level.

The fixed effects (FE) model presented in equation 1 relies on within-agency variation in the 1033 item values across time. We are mainly interested in the estimate for β_1, which would capture how police protection expenditures of counties on average respond to the receipt of 1033 items. Under the assumption that the error term $\varepsilon_{j,t}$ is not systematically correlated with the 1033 item values after controlling for other variables, we can obtain an unbiased estimate for β_1. However, this assumption might not hold for several reasons. For example, Knight (2002) shows that a positive correlation between preferences for public goods and grants receipts can bias the estimate upward. In general, any correlation between time-varying unobserved factors that affect both public expenditures on public safety and the value of receipts through the 1033 program (e.g., motivation by law enforcement leadership to court additional resources) can lead to biased estimates. Recent work in the public finance literature has addressed this endogeneity very carefully.

To address any concerns about endogeneity, we estimate the FEIV model using a set of instrumental variables that are also employed by Harris et al. (2017). The first instrument is the interaction between the lagged total value of 1033 items available nationwide in a given year,

[9] We also show that our results are robust to inclusion of year fixed effects in lieu of state-specific time trends.

$\sum_j^J Value_{j,t-1}$, and the distance from the centroid of a county to the nearest FAC (value-distance interaction). The second instrument is the interaction between the total item value and the log of land area of a county (value-land interaction). We argue that these two instruments are exogenous, based on the institutional details we provide in the Section II, but also empirically evaluate whether our instruments satisfy both the relevance and validity conditions.

V Results

We report results from both the FE and FE-IV models, for both log and level specifications. While the log specification is our preferred model due to the skewed distribution of receipt values, we also report the level specification as it is more common in the literature. We then investigate heterogeneity of treatment effects by equipment type, county politics, and county size using our log specification. Results from the split samples using the level specification are available in the appendix.

Tables 2 and 3 contain the results of our baseline FE model with log and level specifications, respectively. While the estimated coefficients are positive in both specifications, they are significant in the level specification but just below the significance threshold in the log specifications. In each table, Column (1) presents our baseline specification presented in equation 1. Although the estimates are positive in both specifications, they yield different interpretations on the responsiveness of local public spending on police to 1033 receipts. The coefficient in column (1) from the log-log specification in Table 2 implies that a one percent increase in the value of receipts from the 1033 program is associated with a statistically significant 0.025 percent increase in local police spending in the following year. By contrast, the results in Table 3 indicate that the response is greater than one-for-one: a one dollar increase in 1033 receipts is associated with a statistically significant $1.35 increase in the following year. Given that existing studies on intergovernmental grants mostly find evidence of crowding-out, this crowding-in result is noteworthy. The most important implication, however, is that unlike virtually all recently-studied grant programs, we find no evidence of crowding out.

Table 2: The effects of 1033 program on police spending: FE model, log-log specification

	(1)	(2)	(3)	(4)	(5)	(6)
$ln(Item\ value\ per\ capita_{\ t-1})$	0.025	0.026	0.029		0.020	0.023
	(0.017)	(0.017)	(0.020)		(0.018)	(0.017)
$ln(Item\ value\ per\ capita_{\ t-2})$			0.012			
			(0.022)			
$ln(Item\ value\ per\ capita_{\ t-3})$			0.045**			
			(0.023)			
$ln(Item\ value\ per\ capita_{\ t+1})$				-0.019		
				(0.014)		
County characteristics	✓	✓	✓	✓	✓	✓
Crime controls		✓	✓	✓	✓	✓
State-specific linear time trends	✓	✓	✓	✓		
Year Fixed Effects					✓	
National nonlinear time trends						✓
Observations	9471	9471	6755	8119	9471	9471
R^2	0.158	0.159	0.111	0.156	0.075	0.073

Note: This analysis uses the balanced panel described in Table 1. The dependent variable is logged police protection expenditure of county governments per capita. We control for county-level time-varying characteristics such as median household incomes, male shares, shares of the population aged from 15 to 24, racial diversity, poverty rates, and unemployment rates in all specifications. In column (2) to column (6), we add arrest rates for the following crime types: murder, rape, robbery, aggravated assault, burglary, and assault. All dollars are inflation-adjusted using CPI-U-RS and expressed in 2012 U.S. dollars. Standard errors are clustered at the government level and given in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In both tables, Column (2) shows that our results get slightly stronger when we include crime controls.[10] Because the items received through the 1033 program are durable capital goods, they may have budgetary effects beyond the next year, particularly if the estimated crowding-in effect is due to the need for complementary inputs.[11] Column (3) evaluates this possibility by examining whether items received in period $t-2$ or $t-3$ affect contemporaneous spending on public safety. While the results for all three lags are at least marginally significant in the level-level specification in Table 3, only the third lag is significant in the log-log results. However, none of the estimated coefficients are negative, reinforcing the absence of crowding out. Column (4) contains a falsification test for the observed crowding-in effect. We exclude the first lag of the item values from the model and include a lead of the item values instead. The estimate for the lead term is insignificant, providing evidence that pre-existing trends are not of primary concern in our setting. Columns (5) and (6) demonstrate that our results are robust to alternative ways of controlling for time-varying heterogeneity, such as year fixed effects and non-linearities in the national time trend.

While the institutional features of the 1033 program make the sources of endogeneity discussed by Knight (2002) less likely, they do not eliminate all concerns about correlations in the unobservables that affect local spending on local police and acquisition of items through the 1033 program. For example, if chiefs of police who participate in the 1033 program aggressively court resources from all sources, including the local government, our results will suffer from omitted variable bias. We therefore also implement a fixed effects instrumental variables approach as described in Harris et al. (2017), the results of which are found in Tables 4 and 5. As our instruments perform better in a log-log specification, we focus our discussion on these results, although we report and briefly discuss results from the level-level specification for completeness. Both specifications are useful. Ex ante, it is unclear whether or not the relationship between 1033 receipts and local police spending is linear. Also, the distribution of the value of 1033 receipts is highly skewed. The

[10] While adding crime variables to the baseline specification improves the precision of the estimates, it is unclear whether the 'best' specification includes or excludes crime rates. While the arguments for including crime rates as control variables are obvious, studies have also shown that police protection expenditures are endogenous to crime (Levitt, 1997; Di Tella and Schargrodsky, 2004). Therefore, the results in (1) and (2) should not be viewed as a horse-race, but do show that the result that implicates crowding in is robust to the inclusion of crime rates as control variables.

[11] For example, if a department acquires assault rifles, it will need additional funding for ammunition to train/qualify officers beyond the next calendar year.

Table 3: The effects of 1033 program on police spending: FE model, level-level specification

	(1)	(2)	(3)	(4)	(5)	(6)
Item value per capita $_{t-1}$	1.353*	1.495**	1.384*		1.500**	1.552**
	(0.694)	(0.687)	(0.828)		(0.756)	(0.756)
Item value per capita $_{t-2}$			2.288**			
			(1.144)			
Item value per capita $_{t-3}$			2.289*			
			(1.246)			
Item value per capita $_{t+1}$				-0.375		
				(0.373)		
County characteristics	✓	✓	✓	✓	✓	✓
Crime controls		✓	✓	✓	✓	✓
State-specific linear time trends	✓	✓	✓	✓		
Year Fixed Effects					✓	
National nonlinear time trends						✓
Observations	9471	9471	6755	8119	9471	9471
R^2	0.063	0.084	0.057	0.112	0.072	0.071

Note: This analysis uses the balanced panel described in Table 1. The dependent variable is police protection expenditure of county governments per capita. We control for county-level time-varying characteristics such as median household incomes, male shares, shares of the population aged from 15 to 24, racial diversity, poverty rates, and unemployment rates in all specifications. In column (2) to column (6), we add arrest rates for the following crime types: murder, rape, robbery, aggravated assault, burglary, and assault. All dollars are inflation-adjusted using CPI-U-RS and expressed in 2012 U.S. dollars. Standard errors are clustered at the government level and given in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

log-log specification therefore addresses disproportionate influence from jurisdictions that receive exceptionally large values of goods from the 1033 program.

Panel A of Table 4 contains the results of our first-stage estimation, and Panel B contains our FE-IV estimates of the effect of 1033 receipts on local police spending. In each specification, we use land area as a time invariant characteristic, interacted with the value of goods released through the 1033 program as an instrument for the value of 1033 receipts. As noted above, the DLA relayed that land area was a determinant used to resolve competing requests for items.[12] In column (3), we also use distance to the nearest FAC as an instrument. We then test the validity of these two instruments. For the log-log specification, all of our first stage F-statistics on our exclusion restrictions are greater than 20, which far exceeds the benchmark value of 10 (Staiger and Stock, 1997). Additionally, in column (3), the p-value from Hansen's test of validity does not reject the null that our instruments are valid.

In Panel B, Columns (1) and (2) in Table 4 employ the same specification as columns (1) and (2) in Table 2, but instrument for possible endogeneity in 1033 receipts. Estimated coefficients are much larger in magnitude and statistically significant. A one percent increase in 1033 receipts leads to a 0.53 percent increase in department funding in the following year. This causal finding of crowing in holds in our over-identified specification and also when we allow for non-linearities in the national time trend. When we employ year fixed effects as in column (5), we still find evidence of crowding in. However, as our instruments are the product of location-specific time-invariant factors and time-specific location-invariant factors, the weak LM statistic reflects that including county and year fixed effects weakens our identification.

Table 5 contains analogous results for a level-level specification. The statistics from these first stage results indicate that our instruments are borderline weak, but still close to traditional significance levels. The oft-cited rule of thumb for strong instruments is an F-statistic of greater than ten, as proposed by Staiger and Stock (1997). However, that value of ten is not arbitrary, but chosen because it "corresponds to a 5% level test that the maximum size is no more than 15%" in a 2SLS setting Stock and Yogo (2005). The critical values for tests of strength depend on the number of instruments and the particular estimation method. In our analysis, we follow the

[12] We again refer readers to Harris et al. (2017) for additional supporting evidence on the validity of these instruments.

recommendation of Angrist and Pischke (2009) and estimate our FEIV specification with LIML as it reduces bias from marginally weak instruments. As developed by Stock and Yogo (2005), the critical values for a weak instrument test with a maximal size 15% in a LIML estimator with one endogenous regressor and one (two) instrument are 8.96 (5.33) respectively. While Columns (1) and (2) have F-statistics below the threshold, columns (3) and (4) have values that exceed the critical values developed in Stock and Yogo (2005). While our F-statistics are not larger than Staiger and Stock's rule of thumb, they do exceed the critical value that indicates strength roughly equivalent to an F-statistic of 10 in a 2SLS setting.

Panel B contains the second-stage results from our fixed effects instrumental variables specification. Compared to the results in Table 3 they are larger in magnitude and far less precise. While none of the estimates are statistically significant, they do provide two additional pieces of evidence. First, all of the estimates are positive, as were the fixed effects results in Table 2. Second, and consistent with the baseline specification, we do not find evidence that receipts from the 1033 program crowd out any local police spending. In column (3), our overidentified specification, we are able to test the validity of our exclusion restrictions. We cannot reject the null that our exclusions are valid with 5 percent significance.

Table 6 subjects our FEIV results from the log-log specification to two robustness checks on the definition of our sample. First, column (1) reports the estimate from the unbalanced panel to address any concerns about the endogeneity of consistent reporting or the representativeness of the balanced panel. We still find statistically significant evidence of crowding in, although the magnitude is insignificantly smaller than our main sample. Columns (2) and (3) address concerns about whether our result is driven by counties that are highly likely to participate or not at the extensive margin. To that end, we estimated a propensity score for a county's participation in the 1033 program. We drop counties with a less than 10 or greater than 90 percent estimated propensity to participate in column (2) and restrict the sample further in column (3). Although that restriction reduces our sample size by approximately 20 percent and 40 percent in columns (2) and (3), we still find evidence of crowding in.[13]

[13]Level-level results are available in the appendix.

Table 4: The effects of 1033 program on police spending: FE-IV model, log-log specification

	(1)	(2)	(3)	(4)	(5)
Panel A: First-stage					
$ln(Value_{t-1} \times Land_j)$	0.021***	0.021***	0.032***	0.014***	0.021***
	(0.003)	(0.003)	(0.008)	(0.003)	(0.003)
$ln(Value_{t-1} \times Distance_j)$			-0.017		
			(0.012)		
Panel B: Second-stage					
$\ln(Item\ value\ per\ capita\ _{t-1})$	0.536**	0.529**	0.558**	0.449	2.187***
	(0.246)	(0.247)	(0.236)	(0.338)	(0.526)
Kleibergen-Paap LM statistic	43.792	44.011	49.344	20.614	0.922
Kleibergen-Paap Wald F statistic	47.008	47.156	26.743	20.915	45.746
J-Statistic p-value	0.000	0.000	0.294	0.000	0.000
Endogeneity test	0.030	0.032	0.013	0.189	0.270
County characteristics	✓	✓	✓	✓	✓
Crime controls		✓	✓	✓	✓
State-specific linear time trends	✓	✓	✓		
Year Fixed Effects					✓
National nonlinear time trends				✓	
Observations	9436	9436	9436	9436	9436

Note: This analysis uses the balanced panel described in Table 1. Panel A summarizes the first-stage estimates where the the dependent variable is the value of items that a county-level agency has received through 1033 program. $Value_{t-1}$ is the total value of 1033 items released by the Department of Defense in a year $t-1$. $Land_j$ is the land area of county j. $Distant_j$ measures a distance between the centroid of county j and the closest Field Activity Center from the county. Panel B summarizes the second-stage results where the dependent variable is logged police protection expenditure of county governments per capita. The Kleibergen-Paap Wald F statistic shows the relevance of the instruments. The Kleibergen-Paap LM statistic is used to test whether a model is under-identified. In column (3), we present J-statistics to test the validity of over-identifying restrictions. We also test the null hypothesis that the item value can actually be treated as exogenous using a GMM distance test proposed by Baum et al. (2007) and present the results. In order to reduce potential bias from weak instruments, we estimate just-identified models using only the value-land interaction term, except column (3). All dollars are inflation-adjusted using CPI-U-RS and expressed in 2012 U.S. dollars. Standard errors are clustered at the government level and given in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: The effects of 1033 program on police spending: FE-IV model, level-level specification

	(1)	(2)	(3)	(4)	(5)
Panel A: First-stage					
$Value_{t-1} \times log(Land_j)$	0.044***	0.044***	0.126***	0.312***	0.011
	(0.017)	(0.016)	(0.037)	(0.100)	(0.013)
$Value_{t-1} \times Distance_j$			-0.085***		
			(0.033)		
Panel B: Second-stage					
$Item\ value\ per\ capita\ _{t-1}$	6.577	5.288	2.460	4.284	-52.614
	(16.672)	(16.417)	(17.690)	(14.176)	(91.813)
Kleibergen-Paap LM statistic	6.833	6.928	12.583	9.737	0.679
Kleibergen-Paap Wald F statistic	7.029	7.122	6.513	9.784	0.678
J-Statistic p-value	0.000	0.000	0.064	0.000	0.000
Endogeneity test	0.751	0.816	0.857	0.843	0.427
County characteristics	✓	✓	✓	✓	✓
Crime controls		✓	✓	✓	✓
State-specific linear time trends	✓	✓	✓		
Year Fixed Effects				✓	
National nonlinear time trends					✓
Observations	9436	9436	9436	9436	9436

Note: This analysis uses the balanced panel described in Table 1. Panel A summarizes the first-stage estimates where the dependent variable is the value of items that a county-level agency has received through 1033 program. $Value_{t-1}$ is the total value of 1033 items released by the Department of Defense in a year $t-1$. $Land_j$ is the land area of county j. $Distant_j$ measures a distance between the centroid of county j and the closest Field Activity Center from the county. Panel B summarizes the second-stage results where the dependent variable is police protection expenditure of county governments per capita. The Kleibergen-Paap Wald F statistic shows the relevance of the instruments. The Kleibergen-Paap LM statistic is used to test whether a model is under-identified. In column (3), we present J-statistics to test the validity of over-identifying restrictions. We also test the null hypothesis that the item value can actually be treated as exogenous using a GMM distance test proposed by Baum et al. (2007) and present the results. In order to reduce potential bias from weak instruments, we estimate just-identified models using only the value-land interaction term, except column (3). All dollars are inflation-adjusted using CPI-U-RS and expressed in 2012 U.S. dollars. Standard errors are clustered at the government level and given in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: The effects of 1033 program on police spending: FEIV robustness checks

	(1)	(2)	(3)
ln(*Item value per capita* $_{t-1}$)	0.423*	0.594***	0.346*
	(0.220)	(0.229)	(0.177)
Kleibergen-Paap LM statistic	57.308	46.917	43.597
Kleibergen-Paap Wald F statistic	61.814	50.228	46.495
Endogeneity test	0.053	0.007	0.064
Sample	Unbalanced	$0.1 \leq p \leq 0.9$	$0.15 \leq p \leq 0.85$
County characteristics	✓	✓	✓
Crime controls	✓	✓	✓
State-specific linear time trends	✓	✓	✓
Observations	11709	7478	5854

Note: This analysis examines the robustness of our estimates to different samples. Column (1) uses the unbalanced panel of county governments described in Table 1. From column (2) to column (3), we use the balanced panel with sample restrictions based on the likelihood of receiving any items through the 1033 program, p. The dependent variable is logged police protection expenditure of county governments per capita. All columns present estimates from the just-identified model where the value-land interaction term is used to instruments for the item value. See the note in Table 4 for descriptions of test statistics. All dollars are inflation-adjusted using CPI-U-RS and expressed in 2012 U.S. dollars. Standard errors are clustered at the government level and given in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

V.A Heterogeneous effects of the 1033 program

We show in the previous section that grants-in-kind through the 1033 program do not lead to crowd-out, and depending on the specification, may lead to crowd-in of local police spending. While the mechanism is unclear, heterogeneity in the estimated effects may provide some insight. For example, if crowd-in is attributable to a complementary inputs story, one might expect the strength of the crowding in effect to be most concentrated among vehicle receipts, which would require additional funds for maintenance and operation. Alternatively, if acquiring resources through the 1033 program gives chiefs additional political capital when making budget requests, we might expect the crowding-in effect to be more pronounced in counties with a more 'law and order' leaning.

To investigate the complementary inputs hypothesis, we split the value of receipts into two categories: vehicle and non-vehicle. While it is not always clear which items will require complementary inputs, vehicles clearly will. In the absence of additional funds for fuel, maintenance, and possibly personnel who can operate/maintain military vehicles, these items will not be productive inputs for public safety.[14]

[14] We define an item as a vehicle if its FSG is 15 (aircraft and airframe structural components), 16 (aircraft components and accessories), 17 (aircraft launching, landing, and ground handling equipment), 19 (ships, small craft,

Table 7: The heterogeneous effects of 1033 program on police spending: by item types

	Vehicle		Non-vehicle	
	FE	FEIV	FE	FEIV
$\ln(Vehicle\ value\ per\ capita\ _{t-1})$	0.031*	0.854**		
	(0.017)	(0.424)		
$\ln(Nonvehicle\ value\ per\ capita\ _{t-1})$			-0.019	1.275**
			(0.058)	(0.589)
Kleibergen-Paap LM statistic		17.947		76.117
Kleibergen-Paap Wald F statistic		18.638		79.118
Endogeneity test		0.031		0.022
County characteristics	✓	✓	✓	✓
Crime controls	✓	✓	✓	✓
State-specific linear time trends	✓	✓	✓	✓
Observations	9471	9436	9471	9436

Note: We use the balanced panel and estimate the baseline equation (1) using both FE and FEIV models. In column (1) and column (2), the dependent variable is the logged value of vehicle items that a county has received through 1033 program. In column (3) and column (4), the dependent variable is the logged value of non-vehicle items. The FEIV models use the value-land interaction to instrument for the item value. See the note in Table 4 for descriptions of test statistics. All dollars are inflation-adjusted using CPI-U-RS and expressed in 2012 U.S. dollars. Standard errors are clustered at the government level and given in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7 summarizes the estimation results. We still estimate the baseline model presented in equation 1, except that the key covariate in this analysis is either the lagged values of vehicle items or the lagged values of non-vehicle items. Column (1) presents the estimate from the FE model. It suggests that a one percent increase in the lagged 1033 vehicle values raises local police spending by 0.03 percent, and the estimate is statistically significant at the 10% level. The associative estimate for non-vehicle receipts is negative, but statistically insignificant. However, the FEIV estimates for both vehicles and non-vehicles, shown in columns (2) and (4) are both positive and significant. Contrary to what we would expect if vehicles are, in fact, more in need of complementary inputs, the estimated causal crowding in effect is larger for non-vehicle receipts than vehicles.

In Table 8, we question whether the fiscal impact of the 1033 program on local police expenditures depends on the political preferences of counties. We divide the sample into two groups, Democrat and Republican, based on the 2008 U.S. presidential election votes. If the number of votes for the candidate from the Democratic party is above the median among U.S. counties, the

pontoons, and floating docks), 20 (ship and marine equipment), 23 (ground effect vehicles, motor vehicles, trailers, and cycles), or 24 (tractors). Non-vehicle item values are obtained by simply subtracting the vehicle item values from the total item values at the agency level.

Table 8: The heterogeneous effects of 1033 program on police spending: by political preferences

	Republican		Democrat	
	FE	FEIV	FE	FEIV
ln(*Item value per capita* $_{t-1}$)	0.041*	0.553*	-0.005	0.404
	(0.023)	(0.290)	(0.022)	(0.456)
Kleibergen-Paap LM statistic		33.809		12.496
Kleibergen-Paap Wald F statistic		36.514		13.088
Endogeneity test		0.060		0.362
County characteristics	✓	✓	✓	✓
Crime controls	✓	✓	✓	✓
State-specific linear time trends	✓	✓	✓	✓
Observations	4738	4738	4733	4698

Note: We split the balanced panel into two groups based on political preferences of counties and estimate the baseline equation (1) using both FE and FEIV models. The sample is restricted to counties whose Democrat vote shares for 2008 U.S. presidential election are less than or equal to the median of the balanced panel ("Republican") in the first two columns and to counties whose Democrat vote shares are greater than the median ("Democrat") in the last two columns. The dependent variable is logged police protection expenditure of county governments per capita. The FEIV models use the value-land interaction to instrument for the item value. See the note in Table 4 for descriptions of test statistics. All dollars are inflation-adjusted using CPI-U-RS and expressed in 2012 U.S. dollars. Standard errors are clustered at the government level and given in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

county is considered as Democrat. Otherwise, the county is considered as Republican. We then estimate equation 1 using the restricted sample. Under both a FE and FEIV model and a log-log specification, we find statistically significant evidence of crowding in only for Republican counties. While the IV estimate for the Democratic counties is positive, it is not statistically significant. This result, in conjunction with results from Table 7, is more consistent with crowding in resulting from increased political capital rather than complementary inputs. However, we must stress that we cannot identify the specific mechanism behind the estimated crowding in.

Finally, we evaluate whether the heterogeneity in the size of counties matters in determining the effect of the 1033 receipts. We split the sample into large counties and small counties based on the population size and estimate FE and FEIV specifications. The results are summarized in Table 9. Although insignificant, the FE estimates given in column (1) and column (3) reinforce the lack of evidence for any crowding out. The FEIV results show that the crowding in effects are concentrated entirely in the smaller counties. This finding lends some credence to a complementary inputs story. Larger counties, if complementary inputs are needed, find it easier to move money between line

Table 9: The heterogeneous effects of 1033 program on police spending: by population size

	Small		Large	
	FE	FEIV	FE	FEIV
ln(*Item value per capita* $_{t-1}$)	0.031	0.757**	0.016	0.075
	(0.025)	(0.357)	(0.018)	(0.325)
Kleibergen-Paap LM statistic		27.402		19.296
Kleibergen-Paap Wald F statistic		29.133		20.335
Endogeneity test		0.025		0.859
Crime controls	✓	✓	✓	✓
State-specific linear time trends	✓	✓	✓	✓
Observations	4706	4692	4765	4744

Note: We split the balanced panel into two groups based on the population size of counties and estimate the baseline equation (1) using both FE and FEIV models. The sample is restricted to counties whose population sizes are less than or equal to the median of the balanced panel ("Small") in the first two columns and to counties whose population sizes are greater than the median ("Large") in the last two columns. The dependent variable is logged police protection expenditure of county governments per capita. The FEIV models use the value-land interaction to instrument for the item value. See the note in Table 4 for descriptions of test statistics. All dollars are inflation-adjusted using CPI-U-RS and expressed in 2012 U.S. dollars. Standard errors are clustered at the government level and given in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

items because of the breadth and size of the overall budget. Small counties are more likely to face sharper tradeoffs in reallocating resources, and are therefore more likely to seek additional money from the county government.

VI Discussion

This paper investigates the effect of aid-in-kind from the 1033 program on local expenditures on police protection. In sharp contrast to virtually all recent empirical literature on intergovernmental transfers, we find that that the value of 1033 receipts sticks completely (a perfect flypaper effect) or crowds-in additional funding to the recipient county The crowding-in effect is more pronounced in smaller counties and Republican counties.

We believe that the stickiness or crowding-in effect of the 1033 grants is attributable to unique features of the program. Unlike other intergovernmental aid programs, 1033 grants are provided in the form of non-fungible, non-transferable capital goods and the take-up decision is made by a chief of police, rather than local voters or budgetary personnel. The opacity of the process means that there is little if any public oversight associated with gear acquisition. Crowd-in effects are likely driven, in part, by the need for complementary inputs associated with some transferred gear, in particular vehicles, including armored personnel carriers, diesel trucks, airplanes and helicopters. While the stickiness and crowd-in effects of the grant program may be deemed attractive, the lack of local oversight raises questions regarding the welfare properties of the 1033 program.

Due to data limitations, we leave for future work the question of which specific features of the program are responsible for the perfect stickiness and crowding in. Understanding which features (lack of transparency, non-fungibility, application/receipt below the level of budget authority, etc.) leads to crowd in would be of great value in designing future grants in contexts where crowd out is highly undesirable. In education, for example, Title I might be augmented with a network of warehouses that contain ceiling tiles, desks, office supplies, and computers. Principals facing binding budget constraints could make claims on these items. With, for example, a complementary input of a resource person, a school without the funding for a technology class might be able to make one work. When either paternalistic motivations or the desire to enhance positive externalities is the primary motivation for a grant, programs that share some structural or administrative features of the 1033 program may be more effective than earmarked monetary grants.

References

Alesina, A., R. Baqir, and W. Easterly (1999). Public goods and ethnic divisions. *Quarterly Journal of Economics 114:4*, 1243–1284.

Angrist, J. D. and J.-S. Pischke (2009). *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton, NJ: Princeton University Press.

Baicker, K. and M. Jacobson (2007). Finders keepers: Forfeiture laws, policing incentives and local budgets. *Journal of Public Economics 91*, 2113–2136.

Baicker, K. and D. Staiger (2005). Fiscal shenanigans, targeted federal health care funds, and patient mortality. *Quarterly Journal of Economics 120(1)*, 345–386.

Baum, C. F., M. E. Schaffer, and S. Stillman (2007). Enhanced routines for instrumental variables/generalized method of moments estimation and testing. *Stata Journal 7(4)*, 465–506.

Bradford, D. and W. E. Oates (1971). The analysis of revenue sharing in a new approach to collective fiscal decisions. *Quarterly Journal of Economics 85(3)*, 416–439.

Brooks, L. and J. H. Phillips (2008). An institutional explanation for the stickiness of federal grants. *Journal of Law, Economics, and Organization 26(1)*, 263–284.

Brooks, L., J. H. Phillips, and M. Sinitsyn (2011). The cabals of a few or the confusion of a multitude: The institutional trade-off between representation and governance. *American Economic Journal: Economic Policy 3(1)*, 1–24.

Cascio, E. U., N. Gordon, and S. Reber (2013). Local Responses to Federal Grants: Evidence from the Introduction of Title I in the South. *American Economic Journal: Economic Policy 5(3)*, 126–159.

Di Tella, R. and E. Schargrodsky (2004). Do police reduce crime? estimates using the allocation of police forces after a terrorist attach. *American Economic Review 94(1)*, 115–133.

Evans, W. N. and E. G. Owens (2007). C.O.P.S. and crime. *Journal of Public Economics 91*, 181–201.

Filimon, R., T. Romer, and H. Rosenthal (1982). Asymmetric information and agenda control: The bases of monopoly power in public spending. *Journal of Public Economics 17(1)*, 51–70.

Glaeser, E. L. (2012). Urban public finance. *NBER Working Paper Series w18244*.

Gordon, N. (2004). Do federal grants boost school spending? Evidence from Title I. *Journal of Public Economics 88*, 1771–1792.

Harris, M. C., J. Park, D. J. Bruce, and M. N. Murray (2017). Peacekeeping force: Effects of providing tactical equipment to local law enforcement. *American Economic Journal: Economic Policy Forthcoming*.

Hines, J. R. and R. H. Thaler (1995). The flypaper effect. *Journal of Economic Perspectives 9(4)*, 217–226.

Knight, B. (2002). Endogenous federal grants and crowd-out government spending: Theory and evidence from the federal highway aid program. *American Economic Review 92(1)*, 71–92.

Levitt, S. D. (1997). Using electoral cycles in police hiring to estimate the effect of police on crime. *American Economic Review 87(3)*, 270–290.

Lundqvist, H. (2015). Granting public or private consumption? effects of grants on local public spending and income taxes. *International Journal of Tax and Public Finance 22(1)*, 41–72.

Lutz, B. (2010). Taxation with representation: Intergovernmental grants in a plebiscite democracy. *The Review of Economics and Statistics 92(2)*, 316–332.

Nunn, N. and N. Qian (2011). The potato's contribution to population and urbanization: Evidence from a historical experiment. *Quarterly Journal of Economics 126(2)*, 593–650.

Payne, A. A. (2009). Does government funding change behavior? an empirical analysis of crowd out. In J. R. Brown and J. M. Poterba (Eds.), *Tax Policy and the Economy, Volume 23*. University of Chicago Press.

Singhal, M. (2008). Special interest groups and the allocation of public funds. *Journal of Public Economics 92*, 548–564.

Staiger, D. and J. H. Stock (1997). Instrumental variables regression with weak instruments. *Econometrica 65(3)*, 557–586.

Stock, J. H. and M. Yogo (2005). Testing for weak instruments in linear iv regression. In D. W. K. Andrews and J. H. Stock (Eds.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, New York, NY. Cambridge University Press.

Strumpf, K. S. (1998). A predictive index for the flypaper effect. *Journal of Public Economics 69*, 389–412.

A Supplementary Level Results

Table A1: The effects of 1033 program on police spending: FEIV robustness checks

	(1)	(2)	(3)
Item value per capita $_{t-1}$	22.484	15.678	-0.902
	(37.594)	(16.370)	(10.887)
Kleibergen-Paap LM statistic	8.717	7.522	7.827
Kleibergen-Paap Wald F statistic	8.976	7.745	8.031
Endogeneity test	0.556	0.345	0.834
Sample	Unbalanced	$0.1 \leq p \leq 0.9$	$0.15 \leq p \leq 0.85$
County characteristics	✓	✓	✓
Crime controls	✓	✓	✓
State-specific linear time trends	✓	✓	✓
Observations	11709	7478	5854

Note: This analysis examines the robustness of our estimates to different samples. Column (1) uses the unbalanced panel of county governments described in Table 1. From column (2) to column (3), we use the balanced panel with sample restrictions based on the likelihood of receiving any items through the 1033 program, p. The dependent variable is police protection expenditure of county governments per capita. All columns present estimates from the just-identified model where the value-land interaction term is used to instruments for the item value. See the note in Table 4 for descriptions of test statistics. All dollars are inflation-adjusted using CPI-U-RS and expressed in 2012 U.S. dollars. Standard errors are clustered at the government level and given in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A2: The heterogeneous effects of 1033 program on police spending: by item types

	Vehicle		Non-vehicle	
	FE	FEIV	FE	FEIV
Vehicle value per capita $_{t-1}$	1.646**	6.148		
	(0.714)	(19.110)		
Nonvehicle value per capita $_{t-1}$			-3.351	37.821
			(3.147)	(117.773)
Kleibergen-Paap LM statistic		6.114		4.278
Kleibergen-Paap Wald F statistic		6.273		4.303
Endogeneity test		0.812		0.725
County characteristics	✓	✓	✓	✓
Crime controls	✓	✓	✓	✓
State-specific linear time trends	✓	✓	✓	✓
Observations	9471	9436	9471	9436

Note: We use the balanced panel and estimate the baseline equation (1) using both FE and FEIV models. In column (1) and column (2), the dependent variable is the value of vehicle items that a county has received through 1033 program. In column (3) and column (4), the dependent variable is the value of non-vehicle items. The FEIV models use the value-land interaction to instrument for the item value. See the note in Table 4 for descriptions of test statistics. All dollars are inflation-adjusted using CPI-U-RS and expressed in 2012 U.S. dollars. Standard errors are clustered at the government level and given in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: The heterogeneous effects of 1033 program on police spending: by political preferences

	Republican		Democrat	
	FE	FEIV	FE	FEIV
Item value per capita $_{t-1}$	1.677*	17.004	0.925*	-32.790
	(0.966)	(20.070)	(0.537)	(48.062)
Kleibergen-Paap LM statistic		4.685		3.224
Kleibergen-Paap Wald F statistic		4.828		3.267
Endogeneity test		0.405		0.434
County characteristics	✓	✓	✓	✓
Crime controls	✓	✓	✓	✓
State-specific linear time trends	✓	✓	✓	✓
Observations	4738	4738	4733	4698

Note: We split the balanced panel into two groups based on political preferences of counties and estimate the baseline equation (1) using both FE and FEIV models. The sample is restricted to counties whose Democrat vote shares for 2008 U.S. presidential election are less than or equal to the median of the balanced panel ("Republican") in the first two columns and to counties whose Democrat vote shares are greater than the median ("Democrat") in the last two columns. The dependent variable is police protection expenditure of county governments per capita. The FEIV models use the value-land interaction to instrument for the item value. See the note in Table 4 for descriptions of test statistics. All dollars are inflation-adjusted using CPI-U-RS and expressed in 2012 U.S. dollars. Standard errors are clustered at the government level and given in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: The heterogeneous effects of 1033 program on police spending: by population size

	Small		Large	
	FE	FEIV	FE	FEIV
Item value per capita $_{t-1}$	1.594**	21.807	0.909	-580.957
	(0.798)	(22.113)	(0.855)	(4488.828)
Kleibergen-Paap LM statistic		7.234		0.017
Kleibergen-Paap Wald F statistic		7.458		0.017
Endogeneity test		0.326		0.123
County characteristics	✓	✓	✓	✓
Crime controls	✓	✓	✓	✓
State-specific linear time trends	✓	✓	✓	✓
Observations	4706	4692	4765	4744
R^2	0.096	-0.219	0.113	-127.723

Note: We split the balanced panel into two groups based on the population size of counties and estimate the baseline equation (1) using both FE and FEIV models. The sample is restricted to counties whose population sizes are less than or equal to the median of the balanced panel ("Small") in the first two columns and to counties whose population sizes are greater than the median ("Large") in the last two columns. The dependent variable is police protection expenditure of county governments per capita. The FEIV models use the value-land interaction to instrument for the item value. See the note in Table 4 for descriptions of test statistics. All dollars are inflation-adjusted using CPI-U-RS and expressed in 2012 U.S. dollars. Standard errors are clustered at the government level and given in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.