Using Every Door Direct Mail Web Push Surveys and Multi-level modelling with Post Stratification to estimate Perceptions of Police at Small Geographies

Authors

Giovanni Circo, Gainwell Technologies

Andrew Wheeler, CRIME De-Coder LLC

Submission for the National Institute of Justice *Innovations in Measuring Community Perceptions Challenge*

Replication materials for MRP estimates available at https://github.com/gmcirco/nij_survey

Using Every Door Direct Mail Web Push Surveys and Multi-level modelling with Post Stratification to estimate Perceptions of Police at Small Geographies

Summary of Proposal

We propose using the United States Postal Service of *Every Door Direct Mail* (EDDM), in combination with push web surveys, as a cost-effective mechanism for police agencies to generate survey responses persistently over time with effective geographic coverage of small areas.

This is a **non-probability sampling** approach, in which a police department would submit a mailer (such as a postcard) along a particular postal route, with a link and/or QR code to provide responses to a web-based survey.

Postal routes are relatively small geographic areas, typically covering a few hundred residences, which can identify the location of the survey taker at small spatial scales. One can then use these micro level identifiers to build spatial models that can take into account differential response rates using *multi-level regression with post-stratification* (MRP).

We believe this is the *best* approach to meet the stated goals of the NIJ *Innovations in Measuring Community Perceptions Challenge*.

- This approach is ***representative***. Web based push surveys tend to have better response rates for younger individuals compared to direct telephone surveys (Dillman, 2017; Grubert, 2019; Rosenbaum et al., 2015).
- This approach is ***cost effective***. Our appendix estimates a cost of $7.25 per completed survey. It will not rely on needing to hire additional labor to conduct in-person canvases or conduct telephone surveys.

- This approach is ***accurate across micro-geographies***. Using MRP can take local estimates of demographic proportions, and reweight survey estimates to properly estimate average attitudes in small geographic areas (Gelman et al., 2017; Wheeler et al., 2020).
- This approach can be done ***frequently***. This approach only requires setting up an online survey and dropping off mailers to USPS (in bulk).
- This approach is ***scalable***. The EDDM service by the USPS is available in the entire United States. It requires very little upfront capital (print and postage). The labor to implement is small (just dropping off the mailers to the USPS).
- This approach ***protects privacy***. Because the mailers are delivered to everyone along a particular postal route, the sender does not know the name of individuals.

We further describe the necessity and challenge of measuring community perceptions of police at micro geographies, and why we believe our approach is the most rigorous and feasible compared to other methods.

The Necessity and Challenge of Measuring Perceptions at Small Geographies

Like hotspots of crime, there is micro-level variation in attitudes towards police (Wheeler et al., 2020). If one wishes to improve attitudes towards police in these areas, it will be necessary to use methods to accurately estimate those perceptions at small areas. If there are micro locations with negative sentiments towards police, averaging perceptions over larger areas, such as entire cities or larger neighborhoods, as is traditionally done in social science research (Raudenbush & Sampson, 1999), will ultimately result in misleading estimates.
The challenge comes amidst declining response rates to traditional polling methods, in particular telephone surveys (Dillman, 2017).

Using EDDM to gather a Non-Probability Sample

EDDM is a method to send mailers, such as postcards or letters, to every individual in a single mailing route (Grubert, 2019). It is commonly used by advertisers, because it is not required to know the name of the individual to whom the mail is addressed. Rather, one delivers to USPS a package of mailers to be delivered on a particular postage route. One can then have a QR code or a web-link on the mailer to respond to a web-based survey instrument. The web-based survey instrument can have additional checks to ensure that the same respondents do not answer multiple times, such as only using one response per IP address with similar demographics (allowing multiple individuals in the same household to respond to the survey). One may simply link to a different survey endpoint in each post-card mailer, so one can determine the particular postal route associated with each survey response.

Before discussing specifics of EDDM, it is worth briefly describing alternative methods to obtain survey measures and why we believe these are not viable options for accuracy (at micro-geographic areas) or cost-effectiveness. These include two different modalities – cell phone telephone surveys and in-person canvas surveys.

Telephone cell-phone surveys, in our experience, do not generate reliably accurate estimates of micro-level areas. Wheeler et al. (2020) found that for a random digit dialing survey in which they asked respondents to list their nearest street intersection, over 25% refused to answer the question, and an additional 13% provided either ungeocodeable responses or addresses not in the city of interest. While some firms offer polling at the zipcode level, given the mobility of individuals it is very difficult to tie cell phone numbers to specific geographic locations (We believe web-based advertisement driven targeting is likely to suffer

from a similar issue when one wants very micro-geographic responses). In addition to geographic accuracy problems, another issue relates to estimating the cost per individual in surveys that rely on human labor. As such, even without the data quality issues, telephone surveys can be very costly, typically taking several hours of interviewer time per completion of a single survey (Guterbock et al., 2018). Estimates of survey costs per survey completion for cell phone surveys, when including labor cost estimates to conduct the survey, are typically over $30.

While we are unaware of similar cost estimates for in-person canvasses, we suspect they will be *higher* than telephone surveys in terms of labor costs. For example, Fontaine et al. (2019) used teams of 8-12 people over two weeks to gather approximately 200 surveys over several sites. As an estimate, this suggests that a single canvasser generated approximately 25 survey responses over two weeks. While Fontaine et al. (2019) does not list time spent in the field for canvassers, it seems unlikely they would gather more than a few surveys per day.

As such, we consider both approaches non-starters for police departments wishing to conduct persistent surveys with widespread and accurate small area geographic coverage. Although prior academic projects have used individuals to conduct in person canvas interviews in hotspots of crime (Haberman et al., 2016; Kochel 2018; Koper et al., 2022; Weisburd et al., 2023), we are unaware of a single city using this approach over time (without one time funding from outside resources). It would require hiring multiple persistent positions to conduct the surveys and given salary and fringe for such full-time positions, it will ultimately be very costly to conduct a regular survey using either of those methods for police departments or other city agencies.

In comparison, we are familiar with cities enlisting outside firms to conduct semi-regular paper or web-based surveys. For one example, the city of Raleigh, North Carolina pays a firm approximately $60,000 every other year to conduct a paper-based general survey on community opinions of public services the city offers (see <https://cityofraleigh0drupal.blob.core.usgovcloudapi.net/drupal-prod/COR11/fy22-program-catalog.pdf>, page 122). For another example, Chicago contracts with a web-based firm that uses digital advertisements to generate web survey responses for an annual cost of $220,000 (<https://blockclubchicago.org/2020/12/11/the-city-is-polling-people-on-their-feelings-on-police-but-can-it-lead-to-change/>).

Our budget estimates $7.25 per completed survey using the EDDM approach (see Appendix). Therefore, with a budget of under $10,000 per year, a city can generate over 1,000 survey responses. Just as importantly, EDDM provides the ability for an agency to conduct surveys targeted at specific micro-geographic areas. This can be useful in ex-post evaluating particular interventions at specific locations but is also cheap enough to routinely conduct widespread general surveys of attitudes towards law enforcement across a jurisdiction.

Based on these requirements, we believe EDDM will be the *best* solution to obtaining geographically targeted, micro-level area estimates of attitudes towards police for several reasons. One is that given increases in non-response over time, researchers effectively need to target 100% of the population for a small area to be able to obtain reasonable coverage. Any approach that does not give 100% of individuals at a particular micro-geographic location an opportunity to respond will ultimately be less reliable and rely on more extrapolation. EDDM can ensure that everyone in a particular chosen area has the ability to respond.

The second benefit of EDDM is that it is an explicit way to target surveys at a *specific* geographic area. Instead of passive approaches (such as cell phone based or web advertisements) that hope to gather geographically diverse responses, EDDM can be explicitly geographically focused. Thus, it can be used in a similar fashion to micro-level canvas approaches, when very small areas need to be targeted at specific times. But is still much cheaper than hiring individuals to canvas neighborhoods.

The third benefit of EDDM is that it relies on very little capital investment. It requires only four steps: 1) set up a web-based survey, 2) identify the postal routes of interest (which can be easily done in the USPS web-based tool, < https://eddm.usps.com/eddm/select-routes.htm>), 3) print out mailers with a link to the survey, 4) drop the mailers off at USPS. We believe most police departments (or city governments) can use already capitalized labor (such as crime analysts) to generate surveys and physically drop off the mailers. It does not require hiring additional labor to conduct the surveys, nor does it rely on outsourcing to an external firm (such as via web-based digital advertising).

The fourth benefit we list here is that EDDM surveys have shown higher response rates for *younger* individuals (Grubert, 2019). One evaluation comparing web-based to telephone surveys on perceptions of policing showed there were very similar response distributions (Rosenbaum et al., 2015), so we do not believe the modality (web-based vs telephone) to intrinsically introduce any particular bias. In fact, we believe it will likely provide greater coverage, as it can provide opportunities for individuals living in apartments (who are typically not targeted in paper-based address mailing surveys).

While the approach we advocate does limit exposure to individuals with access to the internet (we suggest a mailer that has a convenient QR code link, as well as a printed web URL), we believe the cost-effectiveness of this approach is worth the tradeoff compared to lower response rates for purely paper-based surveys (Dillman, 2017; Grubert, 2019). It is possible to do traditional paper-based surveys using the same EDDM approach, although that would require additional printing costs and return postage costs, as well as labor to digitize the survey results. General access to the internet to complete web-based surveys is estimated to be over 90% (Cantor et al., 2023).

In the total survey error approach (Biemer, 2010), web-based push surveys design away many components of potential extra variance – there is no interviewer bias, and the modality of the survey is consistent for each taker. Being all digital, it does not rely on direct human labor, so analysis can be automated and delivered to end users in near real time. The next section will describe our suggested approach to correcting for differential response under this design, which one only needs estimates of the demographic composition for the area one wants to make inferences for, not survey weights for individual responses.

Using MRP to Estimate Accurate Measures at Small Geographic Areas

Multilevel regression with post stratification (MRP) is a commonly used technique to weight and adjust surveys for non-response bias as well as to generate small-area estimates. MRP is commonly used in public opinion and political polling where samples are often drawn from the state level, but inferences are desired at the county level. One can, however, use the same technique to generate reliable micro-level estimates that is of interest for this application.

Obtaining reliable estimates from surveys at smaller geographic areas present two problems: (1) responses to the survey are often not representative of the population of interest and (2) survey estimates for sparsely populated strata are highly variable or undefined using conventional weighting methods. MRP handles both issues by combining survey post stratification with Bayesian multi-level regression. In brief: the outcome of interest is first estimated using a multi-level regression that includes the demographic strata, auxiliary variables, and random effects for the strata of interest. This is useful because Bayesian methods allow more sparsely populated strata to "borrow" information from other strata via partial pooling (Gelman, et al., 2013). This has the distinct advantage of allowing for estimates for strata which are not even observed in the original survey and is a common technique for estimating small area effects (Buil-Gil et al., 2020).

We believe this is a crucial step in any approach that wants to make reliable inferences at small geographic areas. Micro-geographic areas, even in the unrealistic case with 100% coverage, will only have a small number of responses. Multi-level modelling allows one to estimate effects conditional on person and area level characteristics, while taking into account potential small sample sizes in particular areas.

In the second step the estimates from the regression model are post stratified onto ground-truth demographic counts. These typically are obtained from either the decennial census or one of the waves of the American Community Survey. Hence, MRP provides a statistically principled way to both adjust for non-response bias and obtain small-area estimates. We wish to emphasize that this is in contrast to more conventional survey weighting (e.g., using

inverse probability or raking weights) which are often unable to estimate stable weights for sparse population strata.

For a simplified example of MRP, imagine one estimated an equation, where higher values indicate more trust in the police. For simplicity we only include one demographic covariate, *Older*, and a random effect for each neighborhood *n*. We then have the final resulting equation:

$$\text{Trust} = 1.4 + 0.5 \cdot (\text{Older}) + \lambda_n$$

Where λ_n is the random effect for each estimated neighborhood. Now imagine that we are interested in estimating the average trust in police is a specific neighborhood, that has 20% older individuals and 80% younger individuals (based on external data, such as the American Community Survey), and the random effect estimate *for that specific neighborhood* is 0.1. One would then have the results:

Table 1: Example MRP Post-Stratification Estimate

Group	Equation	Expected Value	Proportion
Younger	1.4 + 0.1	1.5	80%
Older	1.4 + 0.5 + 0.1	2.0	20%
Weighted Expected Value		*1.6*	

Note that the final neighborhood post-stratified estimate does not rely on the proportion of obtained survey responses in either category. If the neighborhood initially had 10 older respondents and 5 younger respondents, older individuals would be given much more weight without some sort of adjustment. Additionally, the survey allows variation to be estimated for specific geographic areas, so one can estimate the spatial variability of attitudes towards the police at different geographic areas, conditional on that area's demographic composition (spatial variation is not entirely dictated by the local demographic composition).

Here there are two steps of adjustment: first is a regression equation to estimate the expected value per some strata, and the second is to properly aggregate those regression-based expected values to some geographic area of interest. The first step (the regression equation) can estimate specific neighborhood and/or micro-geographic level factors, the second step can provide more accurate interpolation to estimate aggregate attitudes over specific geographic areas.

MRP can be extended to more complicated scenarios. For instance, one can estimate random effects that spatially vary for demographics as well, e.g., allow the *Older* effect to vary across neighborhoods. It can also be extended to more complicated regression models, e.g., including more respondent demographic characteristics, such as race and gender, as well as other geographic characteristics, such as nearby crime (Wheeler et al., 2020). It is not limited to post-stratifying estimates to the same areas that are measured. One could have a random effect for postal routes, but then post-stratify to other arbitrary geographic areas using a dasymetric mapping approach (Kim, 2018). Or one could estimate a smooth underlying trend (Bader & Ailshire, 2014; Wheeler et al., 2020), instead of having random effects for discrete geographic areas.

This flexibility allows agencies to tailor the survey and the subsequent analyses to meet their needs. For example, if an agency is concerned about attitudes among a specific demographic (such as English as a second language immigrants), the survey distribution and the analysis can be specifically tailored to generate reliable estimates among that sub-population. Since EDDM targets everyone on a particular postal route, there is no need to have a different survey modality to collect responses among a particular subset of individuals. Web-based

surveys can trivially be constructed so they can be translated to different languages with the click of a button.

While MRP has often been applied in nation-wide surveys for county or sub-county estimates, it can be applied at any geographic level where sufficient auxiliary information is present. A forthcoming paper by Circo, Melde, & McGarrell (2023) used MRP to estimate block-group level estimates of attitudes toward the police in Detroit, MI based on about 800 survey responses.

For the current inquiry, we suggest either using the postal route as a geographic identifier (and have each route have a unique web-based survey endpoint), or if the agency wishes to have even more micro level geographic identifiers, directly collect that information in the survey instrument itself. Using either of these geographic pieces of information, one can estimate spatial variability at micro areas in a MRP framework. A preliminary proof-of-concept study using a community survey of 842 residents in Raleigh, NC showed that obtaining stable model estimates in all 107 block groups was trivial (see < https://github.com/gmcirco/nij_survey/blob/main/mrp_example/mrp.md>). Given the cost estimates above ($10,000/1,000 responses) a mid-to-large sized city could easily obtain small-area estimates of survey responses. This has the added benefit of helping to generate survey estimates for regions of a city which are typically underserved and often have lower response rates.

When considering MRP for this approach, it is important to consider specifically *why* this method is effective for DOJ's stated goals. First, MRP effectively uses existing population demographic data to adjust for differential non-response. While differential response is still a

threat under this design, empirical evidence often shows that response rates are not directly tied to the quality of different survey estimates (Pickett et al., 2018). We believe EDDM, even if response rates are lower than other canvas approaches, are the best approach that balances survey coverage and cost-effectiveness in the total survey error approach (Biemer, 2010).

Second, given the importance assigned to survey estimates at small levels of geography, MRP can effectively generate estimates for these areas within sparsely populated demographic strata (for example, 18–24-year-old Black Males). Historically it is these specific groups (young minority males) which are disproportionately underrepresented in surveys. This is the best use case for MRP because partial pooling can estimate spatial variability but still take into account sub-groups that have sparser responses. The final post-stratified estimates will likely be more representative of the overall perceptions in a micro-geographic area, giving groups historically under-represented via other survey modalities more voice in the final community-based output under our design.

Conclusion

Using EDDM to distribute postcards with a linked web-based survey, in combination with multi-level modelling and post-stratification, we believe is the approach that best meets the stated goals of the competition – providing a cost-effective and reliable approach to measure attitudes at specific geographic micro-areas. Other approaches are too labor intensive (person-based canvases), not geographically focused (cell phone targeting), or rely on outside firms to implement (web-based advertising). EDDM and MRP can be accomplished on a smaller budget with the resources police departments already have.

APPENDIX

Budget Estimates

These budget estimates are for using *Every Door Direct Mail* (EDDM) to submit mailers that contain a link to a *web-based push survey*. The main expense is printing and postage costs for the mailers. Setting up web-based push surveys are very cheap and can likely take advantage of current city web infrastructure or effectively free tools (like google forms).

Below is a hypothetical budget to mail 25,000 mailers. This number comes from an average response rate for EDDM in prior experiments of around 4% (Grubert, 2019). It is possible that EDDM with push-based surveys will result in higher response rates though, Rosenbaum et al. (2015) have an 11% web-based push survey response rate for post police contact survey. In an experiment testing a web-based version of the National Crime Victimization Survey, Cantor et al. (2023) have a 10% response rate. Grubert (2019) is a longer paper-based survey which requires mailing back the survey responses. So, these estimates are likely conservative in terms of response rates.

A 4% response rate would result in a total of 1,000 responses, which is a common benchmark for many city surveys, and appears to result in reasonable enough data to generate more micro level spatial variation (Wheeler et al., 2020). This results in an average cost of $7.25 *per completed survey*.

EDDM Items	**Approximate Unit Cost (Dollars)**	**Dollars per 25,000 mailers**
Printing Postcards	0.10	$2,500
USPS Mailing (Commercial Estimates)	0.19	$4,750
Total	*0.29*	*$7,250*

These cost estimates presume someone from the police department or city agency will bundle and drop-off the mailers directly to the post-office. So does presume a small labor effort on the part of the police agency. It is possible that the police department can also leverage already capitalized printers to further reduce printing costs as well, but we choose 10 cents as an estimate. Also note that EDDM has *non-profit* pricing as well (potentially down to 11 cents per mailing), but here we list the publicly available commercial pricing.

One can find similar batch printing estimates at several online sites, such as <https://www.postcardmania.com/> or <https://www.taradel.com/products/every-door-direct-mail>. These sites also offer entire end-to-end handling, although estimates for that are typically double that listed here (so over 60 cents per individual mailer).

The second cost are web-based push surveys. While there are paid tools to do this (such as SurveyMonkey or Qualtrics), it is fairly simple to set up a custom form on your own domain. We have provided an example application at <https://crimede-coder.com/graphs/survey?surv=se1> or <https://crimede-coder.com/graphs/survey?surv=se2> (this illustrates how using a simple change in the query string to the url can point to custom surveys, to differentiate between different postal routes).

For example, a custom domain (hosted on *Hostinger*) currently costs only $3 a month and comes with a backend database that can hold 3 gigabytes of data and serve over 20,000 web requests a day. If willing to forego a custom domain, a free google account could simple use google forms and google sheets (a single google sheet can have 40,000 rows).

Thus, we do not include cost estimates for the web-based push survey, as worst case these are trivial compared to the cost of postage and labor itself to conduct the survey analysis.

Biographical Sketches

Andrew Wheeler, PhD, is a Principal Data Scientist at Gainwell Technologies and an affiliated faculty member at Georgia State University in the Department of Criminal Justice and Criminology. He received his doctoral degree in criminal justice from the University at Albany SUNY. His research focuses on the applications of predictive policing and operations research within the criminal justice field.

Giovanni Circo, is a data scientist at Gainwell Technologies, holding a PhD in Criminal Justice from Michigan State University. His research focuses on causal inference, Bayesian statistics, and anomaly detection. He has performed large-scale research on evaluating CCTV programs, optimizing police staffing, and conducting evaluations of focused deterrence programs.

Bibliography

Bader, M. D., & Ailshire, J. A. (2014). Creating measures of theoretically relevant neighborhood attributes at multiple spatial scales. *Sociological Methodology*, 44(1), 322-368.

Biemer, P. P. (2010). Total survey error: Design, implementation, and evaluation. *Public Opinion Quarterly*, 74(5), 817-848.

Buil-Gil, D., Moretti, A., Shlomo, N., & Medina, J. (2020). Applying the spatial EBLUP to place-based policing. Simulation study and application to confidence in police work. *Applied Spatial Analysis and Policy*, 13(4), 901-924.

Cantor, D., Edwards, W. S., Giambo, P., Townsend, R., Yan, T., DeMatteis, J. (2023). NCVS Redesign – Comparison of Interviewer and Web Survey Modes. *Bureau of Justice Statistics*, Document No. NCJ 306998.

Circo, G., Melde, C., & Mcgarrell, E. F. (2019). Fear, victimization, and community characteristics on citizen satisfaction with the police. *Policing: An International Journal*, 42(2), 179-194.

Circo, G., Melde, C., & McGarell, E.F. (2023). Estimating Attitudes Toward the Police: A Small-Area Approach using Multilevel Regression with Post Stratification (MRP). *Under Review, Forthcoming*.

Dillman, D. A. (2017). The promise and challenge of pushing respondents to the web in mixed-mode surveys. *Survey Methodology*, 43(1), 3-31.

Fontaine, J., Esthappan, S., La Vigne, N., Lawrence, D. S., Jannetta, J., Dwivedi, A., Lynch, M., Paddock, E. & Vásquez-Noriega, C. (2019). *Views of the Police and Neighborhood Conditions*. Urban Justice Policy Center <https://www.urban.org/sites/default/files/2019/11/11/2019.11.11_changes_to_ni_community_survey_report.pdf>

Gelman, A., Goel, S., Rothschild, D., & Wang, W. (2017). High-frequency polling with non-representative data. *Political communication in real time: Theoretical and applied research approaches*, 89-105.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. CRC press.

Grubert, E. (2019). Every door direct mail in US survey research: an anonymous census approach to mail survey sampling. *Methodological Innovations*, 12(2), 2059799119862104.

Guterbock, T., Benson, G., & Lavrakas, P. (2018). The changing costs of random digital dial cell phone and landline interviewing. *Survey Practice*, 11(2).

Haberman, C. P., Groff, E. R., Ratcliffe, J. H., & Sorg, E. T. (2016). Satisfaction with police in violent crime hot spots: Using community surveys as a guide for selecting hot spots policing tactics. *Crime & Delinquency*, 62(4), 525-557.

Kim, Y. A. (2018). Examining the relationship between the structural characteristics of place and crime by imputing census block data in street segments: Is the pain worth the gain?. *Journal of Quantitative Criminology*, 34, 67-110.

Kochel, T. R. (2018). Police legitimacy and resident cooperation in crime hotspots: Effects of victimisation risk and collective efficacy. *Policing and Society*, 28(3), 251-270.

Koper, C. S., Taylor, B. G., Liu, W., & Wu, X. (2022). Police activities and community views of police in crime hot spots. *Justice Quarterly*, 39(7), 1400-1427.

Pickett, J., Cullen, F., Bushway, S. D., Chiricos, T., & Alpert, G. (2018). The response rate test: Nonresponse bias and the future of survey research in criminology and criminal justice. *The Criminologist*, 43(1), 7-11.

Raudenbush, S. W., & Sampson, R. J. (1999). Ecometrics: toward a science of assessing ecological settings, with application to the systematic social observation of neighborhoods. *Sociological Methodology*, 29(1), 1-41.

Rosenbaum, D. P., Lawrence, D. S., Hartnett, S. M., McDevitt, J., & Posick, C. (2015). Measuring procedural justice and legitimacy at the local level: the police–community interaction survey. *Journal of Experimental Criminology*, 11, 335-366.

Weisburd, D., V. Uding, C., Hinkle, J. C., & Kuen, K. (2023). Broken Windows and Community Social Control: Evidence from a Study of Street Segments. *Journal of Research in Crime and Delinquency*, Online First.

Wheeler, A. P., Silver, J. R., Worden, R. E., & Mclean, S. J. (2020). Mapping attitudes towards the police at micro places. *Journal of Quantitative Criminology*, 36, 877-906.