Class and the Development of Trust in Police in Latin America*

Tara Slough† Carolina Torreblanca‡

April 3, 2024

Abstract

In the United States, trust in police is strongly positively correlated with socioeconomic status. We show that this is not the case in Latin America. Using 147 cross-country surveys across 20 countries and three panel surveys, we find that, if anything, trust is *negatively* correlated with socioeconomic status—a fact that neither regional experts nor subject-matter experts anticipated in forecasts conducted for this project. We consider three possible (non-rival) explanations: that the security-socioeconomic-status gradient is steeper in the United States than in Latin America; that citizens of different socioeconomic statuses translate security outcomes into signals about the police in different ways; or that people of different socioeconomic statuses update upon these signals in different ways. We conclude that the second mechanism—novel to this paper—is active: rich people are more likely to interpret negative experiences as signals of police (non)-trustworthiness.

*We thank Dorothy Kronick, Zara Riaz, and audiences at the Penn Development Research Initiative Crime and Policing in Comparative Perspective Conference and APSA for excellent comments.

†Assistant Professor, New York University, tara.slough@nyu.edu.

‡PhD Candidate, New York University, ctb370@nyu.edu.

A durable empirical regularity in the United States holds that citizens' race and class correlate with their beliefs about and attitudes toward the police. For example, in a 2021 Gallup survey, 56% of white respondents and 27% of Black respondents expressed "quite a lot" or "a great deal" of confidence in the police (Gallup, 2021). Such disparities in trust or confidence in the police, assessments of police service quality, and attitudes toward police have been observed for decades (Decker, 1981; Skogan, 2005; Macdonald and Stokes, 2006). Conventional wisdom holds that groups with worse experiences of the police—whether more abuse, worse quality of service, or poorer security outcomes—tend to trust the police less. In this paper, we document that this conventional wisdom about police service quality and trust from the United States does not travel to Latin America. In so doing, we propose a new account of the formation of trust in police.

Latin America and the Caribbean suffer the highest rates of crime and violence of any region (Muggah and Tobon, 2018; Vilalta, 2020). These high crime rates impose substantial welfare costs (Soares and Naritomi, 2010; Jaitman et al., 2015) and, in public opinion surveys, regularly register as important concerns of citizens (LAPOP, 2022). The primary institutional responses of the state to problems of crime and violence involve the police and the justice system. But these state agents do not operate in a vacuum, in isolation from the citizens they police and serve. In this paper, we seek to more accurately characterize one dimension of the relationships between citizens and police in order to generate new implications for states' abilities to address crime and violence.

We focus on citizen trust in the police. Refining general definitions by Hardin (2003) and Bhattacharya, Devinney, and Pollutla (1998), we conceptualize trust in the police as a citizen's belief that a representative police agent will take an action that produces a beneficial outcome for the citizen. Citizen trust in police is important because many security outcomes are co-produced between citizens and police officers. Indeed, Blair et al. (2021: p. 1) assert that by generating more trust, interventions like community policing can "build more effective police agencies in environments of low trust." In these environments, trust is posited to promote information sharing (e.g., crime tips), which should increase the ability of police to locate, stop, remedy, investigate, or preempt crimes or misdemeanors. To the extent that community-policing interventions that seek

to build trust in police have occupied a central role in policing in Latin America (as elsewhere), it is important to understand this key outcome—trust in police—across the region.

Drawing on a growing body of literature on policing in Latin America, we evaluate a widely-held assumption that trust in police is increasing in social class or socioeconomic status. Latin America is the world's most unequal region (Hoffman and Centeno, 2003; Gasparini and Lustig, 2011) where social class has long been considered a salient social cleavage.[1] Work on police abuse and repression suggests that these behaviors target poor and marginalized communities (Magaloni, Franco-Vivanco, and Melo, 2020; González, 2020; González and Mayka, 2022). Moreover, anti-poor bias is believed to be pervasive in many Latin American justice systems (O'Donnell, 1999; Brinks, 2007, 2019). Following the conventional logic from the United States, that poor treatment and bad security outcomes reduce trust in police, we would expect trust in police to be increasing in socioeconomic status.

Leveraging 147 surveys from 20 countries in the region, we do not find support for this hypothesis. Indeed, pooling the 235,230 individual responses from all countries, we estimate that the correlation between income and trust in police is -0.053, and the correlation between education, a proxy for social class with less missingness, is -0.084. The negative correlation suggests that low socioeconomic-status individuals trust the police slightly more than their high socioeconomic counterparts. Both correlations are statistically distinguishable from zero. Disaggregating across countries, we do not find a positive correlation that is statistically distinguishable from zero between either measure of socioeconomic status and trust in police in *any* of the 20 Latin American/Caribbean countries for which we have data. This represents a strong departure from decades of findings from the US, where income and the same trust question (in English) exhibit a *positive* correlation of 0.13.[2]

These results are surprising in light of conventional wisdom about the relationship between

[1]Other identity-based cleavages (e.g., ethnicity or race) in Latin America vary more substantially across countries.

[2]In the US, the racial disparity between Black and White respondents is approximately twice the magnitude of the difference between respondents in the top and bottom income decile.

policing outcomes and trust in police. We measure the degree to which two groups of experts anticipated these findings through a forecasting exercise (DellaVigna, Pope, and Vivalt, 2019). Our two groups of experts consisted of a sample of scholars of politics and public administration in Latin America and a sample of activists working on issues related to justice in Mexico. We show that the modal expert and the average forecast anticipated a positive correlation between socioeconomic status and trust in the police. Interestingly, inaccuracies in forecasts stem from underestimates of trust by poor respondents (those at the 10^{th} percentile) only. Respondents are quite accurate in their assessment of trust by median and rich respondents (those at the 50^{th} and 90^{th} percentiles, respectively). This exercise affirms that our findings challenge conventional wisdom, not only from the US but also from subject-matter and regional experts.

Why do our results depart so substantially from existing understandings of the correlates of trust in police derived from decades of research in the US? Comparing our findings to those from the well-studied US case where trust increases in socioeconomic status, we measure manifestations of multiple mechanisms that could produce correlations between class and trust in police. We thereby probe the external validity of the mechanisms thought to generate canonical findings from the US (Slough and Tyson, 2023).

Within our cognitive conceptualization of trust in police, citizen trust in police evolves through observation of police or security outputs. When citizens are victims of police abuse or crime, for example, they update negatively on the trustworthiness of the police. Good service, on the other hand, leads to positive updating. Data from cross-sectional and panel survey reports of crime victimization, exposure to police corruption, and general feelings of security affirm this assumption. The conventional wisdom from the US holds that rich citizens receive better service from the police—in other words, they are more likely to see "trustworthy" signals—and thus are more likely to update positively than poor citizens. As this process perpetuates, the rich come to trust the police at higher rates thereby inducing the observed positive correlation between class and trust in police.

This mechanism—the rich receive more positive signals—was cited (in some way) by many

experts in our elicitation exercise. We show that on the basis of self-reported survey data this is not the case: in contrast to the US, in Latin America, rich respondents self report *higher* rates of crime victimization and police corruption than poor respondents; perceptions of insecurity do not vary substantially in socioeconomic status. While these observations help to explain the negative correlation between trust and socioeconomic status observed in Latin America, they stand in contrast to administrative data on *violent* crime (e.g., homicide) and expert beliefs that the poor are disproportionate victims of crime, insecurity, and abuse by police.

We suggest that this apparent contradiction between conventional wisdom and survey data can be resolved through two observations. First, survey and administrative data suggest that police corruption, perceived insecurity, and (in general) crime rates are higher in Latin America than in the US, which correspond to a greater overall probability of observing a poor signal of police performance (i.e., an intercept shift). Second, we provide suggestive evidence that the translation of observed outcomes of police service into *perceived* signals varies by socioeconomic status. Here, we argue that the threshold for perceiving a negative signal of police performance is *lower* for rich than for poor citizens. This increases the exposure of the rich to signals of poor police performance, holding constant actual performance.[3] This mechanism may also be present in the US, but it should be less influential (and thereby harder to detect) given the lower rate of crime and insecurity.

We explore but find no evidence for several alternative explanations premised on the measurement of trust in police and conceptualization of trust. Using various bounding approaches and ancillary analyses related to questions of institutional trust, we do not recover evidence that systematic measurement error drives our findings. We then consider the possibility that our concept of trust is mischaracterized. We examine whether trust in institutions is a fixed trait rather than a belief by examining the intra-cluster correlation in trust across multiple institutional trust measures

[3]We show that conditional on perceiving such a signal, poor and rich citizens update in the same direction by a similar magnitude. Thus, one need not be highly educated in order to rationally update on police trustworthiness from these signals, unlike arguments in other domains (Weitz-Shapiro and Winters, 2017).

(e.g., Ojeda, 2016; Mondak et al., 2011). Further, we ask whether trust might be picking up preferences for the role of police by considering respondent preferences over *mano dura*, which may be correlated with income. Here, we find that, if anything, preferences for *mano dura* are stronger among the rich and predict *lower* trust in police across the income distribution.

This paper contributes to the extensive literature on institutional trust (Levi and Stoker, 2000; Citrin and Stoker, 2018). We build on the insight that trust is cognitive (Bhattacharya, Devinney, and Pollutla, 1998; Hardin, 2003) by using data to isolate different attributes of citizens' updating on police agent trustworthiness. We show strong evidence that citizens of all socioeconomic statuses update in a roughly Bayesian manner (on average) in response to *perceived* signals of low trustworthiness. Yet, we note that these signals of police performance are ultimately perceptions, which vary in the their fidelity to actual service provision. To the extent that this mapping between service provision and citizen perceptions varies in demographic features—like class—similar state outputs may beget very different levels of trust.

Our paper is primarily descriptive, responding to important calls for larger-scale quantitative description (Gerring, 2012; Munger, Guess, and Hargittai, 2021). While recent randomized interventions have sought to increase citizen trust in police (Blair et al., 2021; Karim, 2020; Peyton, Sierra-Arévalo, and Rand, 2019), our goal here is to try to understand baseline levels of trust in police in Latin America. Because causal effects of policing interventions are ultimately differences from these baseline levels, a richer descriptive understanding of this important outcome has two benefits. First, we can better interpret the causal and welfare effects of interventions. Second, for policymakers, police forces, and scholar intent on increasing trust in police, we show that these interventions are less easily targetable (on the basis of income) than in the US context, affirming an observation by Hanson, Kronick, and Slough (2022).

We make one further methodological innovation that serves to advance efforts to cumulate descriptive knowledge. We integrate expert forecasting to assess the state of knowledge of these outcomes. While recent literature advances the use of forecasting in experiments (DellaVigna, Pope, and Vivalt, 2019; DellaVigna, Otis, and Vivalt, 2020), we show how these tools can be used

in service of description. Using forecasting data for description disciplines a reliance on heuristics and allows us to identify blindspots among scholars.

1 Trust in Police: Concept and Context

1.1 Concept of institutional trust

Social scientists have devoted substantial attention to the measurement of citizens' trust in government or governmental institutions (Levi and Stoker, 2000; Citrin and Stoker, 2018). Nevertheless, the definition of "trust"—and, indeed, the possibility of institutional trust—remains contested (PytlikZillig and Kimbrough, 2016). Building upon the conceptualizations of Hardin (2003) and Gerbasi and Cook (2009), we conceive of institutional trust as being *cognitive* and *relational*.

By cognitive, we mean that trust is fundamentally a belief. Bhattacharya, Devinney, and Pollutla (1998: p. 465) characterize trust as "an expectancy of positive (or nonnegative) outcomes one can receive based on the expected action of another party in an interaction characterized by uncertainty." In other words, individuals hold a belief about how the other party is likely to act or behave. In contexts of policing, this could be an individual's expectation about how police will treat them or an expectation of whether and how police will respond to a tip about a crime or misdemeanor.

By relational, we emphasize interactions between two parties, citizens and police agents. This is implicit in the formulation of trust by Bhattacharya, Devinney, and Pollutla (1998). Within this conception of trust, beliefs can be changed (updated) by observation of the other party's behavior. In the context of citizen trust in police, thus, when citizens observe police behavior, they gain information about police trustworthiness. This signal can be used to update a citizen's belief about how the police might behave toward them in subsequent encounters. An unsavory encounter with a police agent, for example, can lead citizens to negatively update about the trustworthiness of police, in general, thereby reducing trust in police.

Hardin (2003) was skeptical of whether institutional trust is possible, largely because of limits to citizens' ability to form relationships with an institution. Whereas citizens may be able to inter-

act with individual police agents, Hardin (2003) argues, it is nonstandard to think of an institution as an actor with whom these interactions might take place. We argue that interactions/relationships between individual citizens and individual police agents shape trust in police. In this context, citizens can hold beliefs about whether an individual officer is trustworthy. Moreover, they can make assessments about the share of trustworthy officers on a police force (or in a given police unit). Institutional trust is, therefore, both cognitive and relational.

It is useful to clarify two alternatives distinct from our concept of institutional trust. First, some authors view or evaluate political trust as a trait (e.g., Ojeda, 2016; Mondak et al., 2011). Individuals from different groups may have different baseline propensities to trust other individuals or agents of institutions. If this were the case, environmental or genetic traits could confound the relationship between social class and trust. Alternatively, social/political trust may facilitate economic advancement, thereby increasing an individual's social class and generating a positive association between the two measures (Putnam, Leonardi, and Nanetti, 1993). Second, individuals undoubtedly hold varying preferences over what police should do or how the institution should function. The accounts of motivated reasoning or inference proposed by Kunda (1987) and Taber and Lodge (2006) suggest that these preferences may affect how citizens form beliefs about police trustworthiness. While it is, of course, possible that preferences condition updating processes (Little, Schnakenberg, and Turner, 2022), we contend that experiences with the police—good or bad—shape future expectations about police agents in the direction of the signal.

1.2 Policing and class in Latin America

Despite Latin America's regional turn towards democracy, police forces routinely engage in corrupt and abusive behavior (Macaulay, 2012; Magaloni, Franco-Vivanco, and Melo, 2020; Johnson, Mendelson Forman, and Bliss, 2012). Citizens' opinions of police in the region tend to be by and large negative (Malone and Dammert, 2021; Cao and Zhao, 2005). Yet, experiences with police are far from uniform: research indicates that police forces can behave repressively toward lower-income individuals and individuals from marginalized groups while being responsive to the demands of privileged community members (González, 2020; González and Mayka, 2022). Addi-

tionally, regional scholars have pointed out bias against poor, indigenous, and other marginalized communities in the region's justice systems that lingers even after recent reforms (O'Donnell, 1999; Brinks, 2007, 2019).

Socioeconomic status could be related to exposure to policing or the outcomes of policing through different channels. Crime victimization, for instance, is negatively associated with trust (Cruz, 2015; Gaviria and Pagés, 2002) and positively associated with support for "*mano dura*" policing (Visconti, 2020). Further, feelings of insecurity correlate positively with more support for spending on the police (Altamirano, Berens, and Ley, 2022).

To our knowledge, no study has systematically analyzed how support for police covaries with class across Latin American countries. Nevertheless, existing accounts support a common premise: socioeconomic status predicts an individual's exposure to policing or the outcomes of policing. These distinct experiences with police agents or crime outcomes should provide different opportunities for learning about the trustworthiness of the police.

2 Research design

Our primary research question is descriptive: how does trust in police vary in social class? Accordingly, we estimate the correlation between measures of socioeconomic status and reported trust in police. We view these correlations as important in characterizing citizen-police relationships in Latin America and elsewhere.

This quantity—correlation between socioeconomic status and trust and belief—is likely to capture information relevant to interactions between citizens and police. Levels of trust are not outwardly-observed characteristic, but individuals' level of trust in the police predicts at least some citizen behaviors toward the police. For example, Hanson, Kronick, and Slough (2022) show that citizen behavior toward police—in the context of community-police meetings—does vary in levels of trust in Medellín, Colombia. In the context of interactions with citizens, it is plausible that police officers may want to ascertain a citizen's level of trust when deciding how to engage. Because beliefs are unobserved, police may use observable characteristics to infer a citizen's trust

and their likely behavior, a form of statistical discrimination (Phelps, 1972). In contrast to levels of trust, in Latin America, socioeconomic status is typically easily observable through an individual's dress, way of speaking, comportment, and surroundings (Britto Ruiz and Ordóñez Valverde, 2005; Sabatini, 2006; Villarreal, 2010).

2.1 Data

Our principal data source is LAPOP's AmericasBarometer (LAPOP, 2022). The sample incorporates the responses of 241,626 individuals collected from 147 unique surveys in 20 Latin American countries between 2004 and 2019. Each survey round was designed to be representative of the country's voting-age population that year. Table A1 in the Appendix lists the years and countries where the surveys were collected.

We use respondents' self-reported income bracket to measure socioeconomic status or class. We supplement this measure with individuals' educational attainment, in years of schooling, to assuage concerns that systematic misreporting of income might drive any results. In each round, respondents are asked about their trust in several institutions. We measure trust in the police with the question "To what extent do you trust the Police?" Responses range from 1 (not at all) to 7 (a lot).[4]

We complement the (repeated) cross-sectional data from LAPOP with three panel surveys that measure trust in police: a five-wave nationally representative survey from Chile (COES, 2022), a two-wave representative survey from Medellín, Colombia conducted in 2018 and 2019 (Hanson, Kronick, and Slough, 2022), and a quarterly representative rolling panel from Mexican cities spanning 2017-2023 (INEGI, 2024). While these surveys cover just three settings (Chile, Medellín, and Mexico, respectively), they allow us to examine within-individual variation over time.[5] We focus on individual variation in exposure to poor police service provision—crime victimization, police corruption, and perceived insecurity.

[4]We report all survey questions and response scales in Table A5.

[5]While other existing longitudinal datasets cover other Latin American countries, we are unaware of others that include repeated questions about policing.

2.2 Estimation

We use ordinary least squares (OLS) to estimate the correlation between trust in police and socioeconomic status, as detailed in Equation 1. We regress individual i's self-reported trust in the police, Trust_i, on a self-reported measure of class (income or education), Class_i. Both trust and class measures are standardized within country-year to account for secular trends. We cluster standard errors at the level of each survey's primary sampling unit. Because the dependent and the independent variables are standardized, β is the estimator of the correlation coefficient.

$$\text{Trust}_i = \alpha + \beta \text{Class}_i + \epsilon_i \tag{1}$$

Correlation summarizes a linear relationship. We allow for non-linearities in the relationship between trust and socioeconomic status by binning measures of socioeconomic status by decile and estimating the following equation by OLS:

$$\text{Trust}_i = \sum_{d=1}^{10} \beta_d \mathbb{I}[\text{Class}_i = d] + \epsilon_i \tag{2}$$

In this expression, the β_d's are estimators of the average level of trust in police by respondents in each decile d.

Finally, when considering explanations for observed patterns of trust in police, we examine how individual experiences of police abuse or poor security outputs affects trust in police (both unconditionally and by a respondent's socioeconomic class). To do so, we rely on the panel surveys at the individual level to estimate the average treatment effect on the treated (ATT). To estimate the ATT, we report estimates from a standard two-way fixed effects estimators. We denote a binary observation of poor performance $S_{it} \in \{0, 1\}$ and estimate:

$$\text{Trust}_{it} = \delta S_{it} + \gamma_t + \psi_i + \epsilon_{it}, \tag{3}$$

where γ_t and ψ_i are time and unit fixed effects, respectively. In this specification, δ is our estimator

of the ATT. We also employ a more general fixed effects counterfactual estimator proposed by Liu, Wang, and Xu (2022) to ensure robustness of our ATT estimates to a variety of weighting problems that plague two-way fixed effects estimators.

3 Baseline results

We present estimates of the correlation between socioeconomic status and trust in the police for the entire pooled sample of respondents in Figure 1. In addition, we also plot country-specific correlations. Contrary to conventional wisdom, the overall pooled correlation and all but one country-specific result are close to zero and slightly negative. The estimated correlation for the pooled sample is -0.053 [95% CI: -0.059, -0.046] when class is operationalized as income and -.084 [95% CI: -0.089, -0.078] when we use education as a proxy. El Salvador has the most negative country-specific correlation, with an estimated correlation of -0.18 [95% CI: -0.20, -0.-0.16] when class is operationalized as education. For the rest of the countries, the correlation oscillates between -0.15 and .01. In the cases of Argentina, Chile, Costa Rica, Ecuador, Jamaica, and Peru, at least one correlation estimate is not statistically distinguishable from zero.

Correlation measures a linear relationship. Do the weak, negative correlations reported in Figure 1 mask a stronger, non-monotonic relationship between socioeconomic status and self-reported trust in the police? To explore the possibility, we divide respondents into class deciles and plot the mean level of trust in the police for the members of each decile. As before, we compute the means for the entire pooled sample and each country. Figure 2 reveals no evidence of a non-monotonic relationship between the two variables: the mean levels of trust are stable or decreasing only slightly in income or education in all countries and the pooled sample.

One limitation of our analysis is that we measure class using self-reported education and income measures, which are indirect measures of socioeconomic status. Two alternative measures may be preferable for different purposes. First, some countries have administrative classifications of class. For example, in Colombia, dwellings are categorized by *estrato* (socioeconomic stratum) to prorate public utility charges. Similarly, based on census data, the Mexican National Institute of

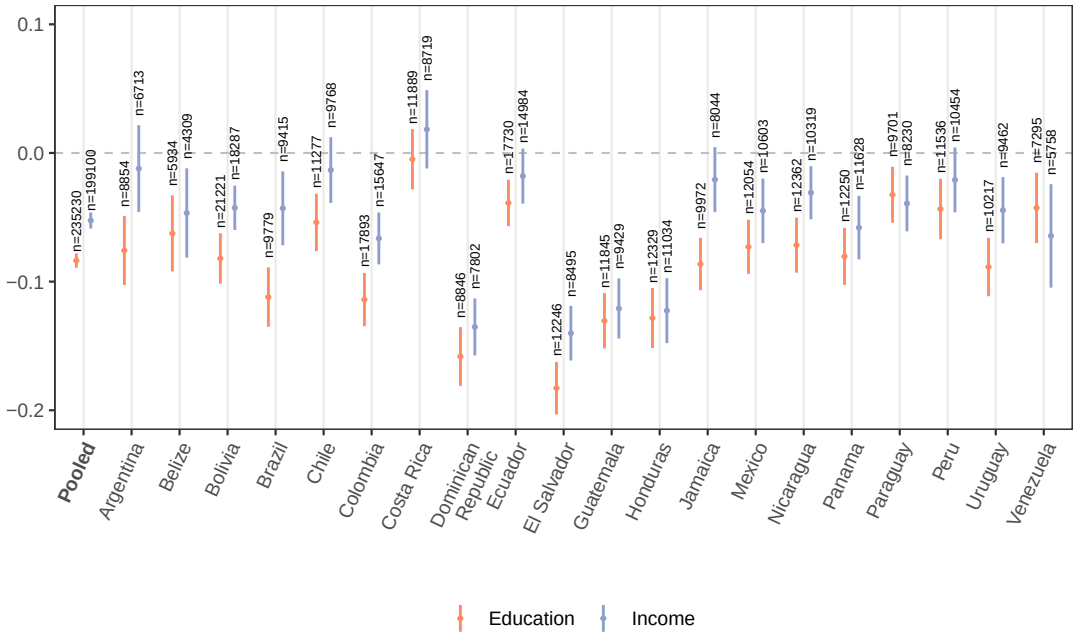

Figure 1: Correlation between LAPOP respondents' self-reported trust in police and two class measures: income (in blue) and education (in orange).

Statistics and Geography (INEGI) classifies primary sampling units of dwellings into sociodemographic *estratos*. Drawing upon the original panel survey in Medellín, we estimate a correlation between dwelling *estrato* and trust in the police of 0.068 [95% CI: 0.029, 0.107]. Using the rolling panel survey conducted in Mexican cities, we estimate a correlation between *estrato* and trust in the police of -0.063 [95% CI: -0.068, -0.059]. The correlation estimated using the Mexican panel data is similar to the LAPOP estimates based on self-reported education and income: -0.073 [95% CI: -0.093, -.052] and -0.045, [95% CI: -0.070, -.020], respectively. The correlation estimated using the Medellín panel survey is somewhat higher than the Colombian averages reported in Figure 1 (-0.068, 95% CI: [-0.108, -0.028] and 0.006, 95% CI: [-0.035, 0.046], respectively.) However, rescaling the national Colombia correlation estimates by a magnitude similar to the difference between estimates using self-reported class and the administrative *estrato* classification from the Medellín survey would preserve a weak negative correlation with trust in police.

Second, citizens may identify with a higher or lower class than their income or education would suggest. From panel data in Chile, in which class identification is measured subjectively, we es-

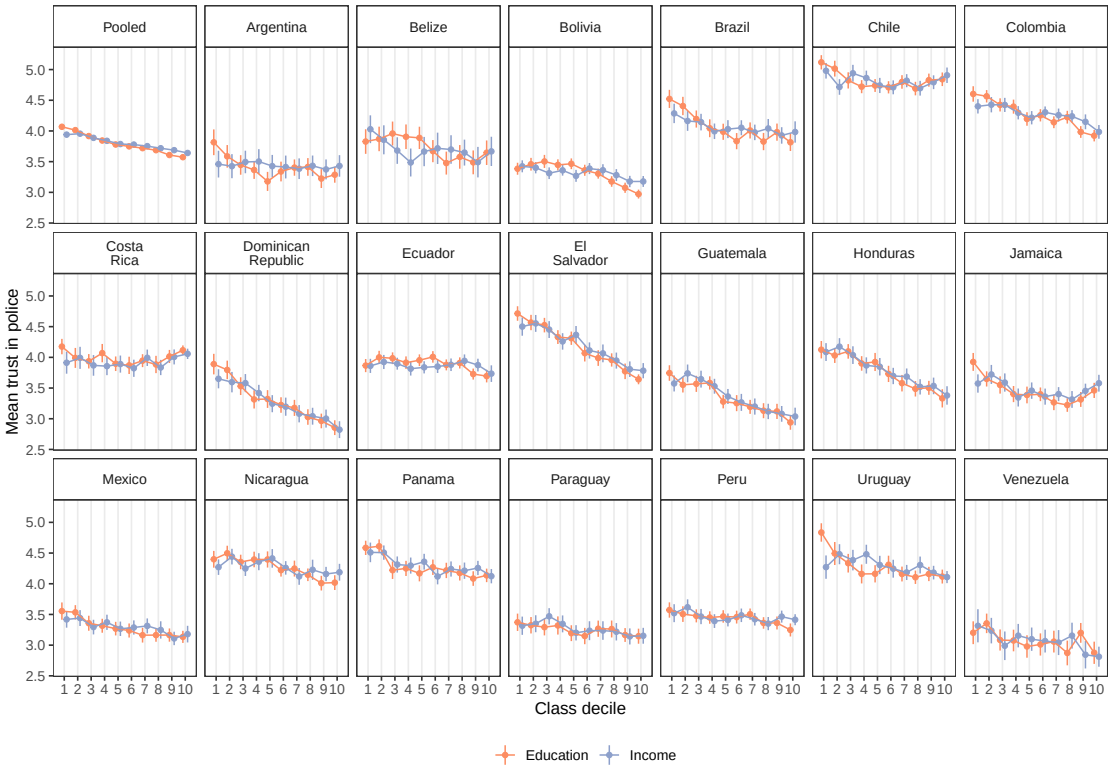

Figure 2: Mean trust in the police on a seven point scale (1-7) by decile of income (in blue) and education (in orange).

timate a correlation of 0.038 (95% CI: [0.003, 0.073]) between self-identified class and trust in police. While this is the only positive correlation that we find that is (statistically) distinguishable from zero, it remains substantially weaker than correlations between class and trust in police observed in the US. These ancillary surveys suggest that our findings from the widely-available LAPOP proxies of class do not substantially mislead relative to plausible alternative measures.

4 Expert Forecasts

The finding that trust in the police covaries weakly and, in general, negatively with socioeconomic status in Latin America was surprising to us. We conducted an expert forecast elicitation with two samples to assess whether our findings were similarly surprising to other experts. Expert forecasts are increasingly used to measure experts' prior beliefs about quantities of interest in social science research (DellaVigna, Pope, and Vivalt, 2019).

Our two expert samples are (i) scholars of Latin American politics and (ii) activists working on issues of human rights and policing. We used a recent program of a Latin American politics conference known for the participation of scholars from all regions as our sample frame for the academic sample. One of the authors identified a network of activists through past non-academic employment related to policing in Mexico. All participants were invited to share the survey with other interested colleagues and students. In sum, we collected 121 country-level forecasts from 101 unique experts. Table A7 in the Appendix shows the count of forecasts by country and respondent type.

We asked experts to provide at least one *forecast* for one country in Latin America (or the region as a whole). A forecast consists of three quantities: mean levels of trust—per the LAPOP survey question—at the 10th, 50th, and 90th percentiles of household income. As depicted in Appendix A2, our forecasting instrument contextualized the income range by reporting average income at each level. Respondents had access to the text of relevant LAPOP questions in English and Spanish.

Figure 3 shows that, on average, expert respondents expected a positive correlation between

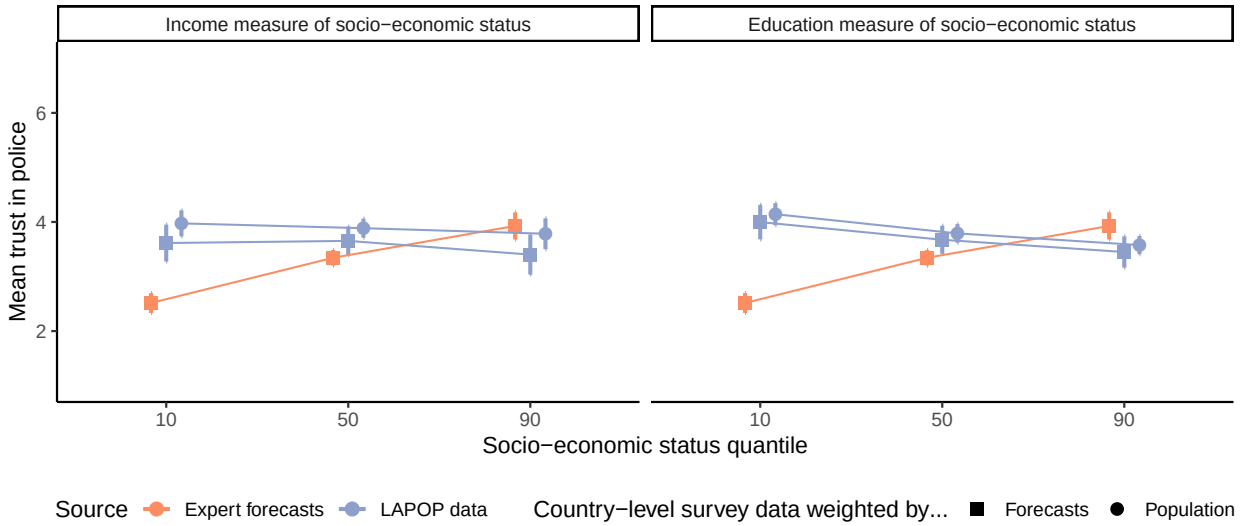

Figure 3: Divergence between average forecasts (in orange) and corresponding survey-based measures (in blue).

socioeconomic status and trust in police. The average expert forecast is monotonically increasing from 2.51 [95% CI: 2.32, 2.69] for citizens in the first decile of income to 3.34 [95% CI: 3.20, 3.49] for citizens of median income to 3.93 [95% CI: 3.68, 4.18] for citizens in the tenth decile of income. Because the forecast is an average over country-level forecasts, we report two measures from the survey data, drawn in blue. First, we weight country-level surveys by the prevalence of each country in the forecasts to ensure that the frequency with which countries are forecasted does not drive the divergence. We also weigh the survey data by each country's population, offering a more interpretable regional average. Both weighting schemes yield similar flat or slightly negative relationships between socioeconomic status and trust in police, starkly contrasting with the forecasts. The right panel of the graph shows that using education rather than self-reported income to measure class in the LAPOP survey data does not change our qualitative finding that the positive relationship between income and crime anticipated by experts —including us— is not borne out in the data.

Two further disaggregations of the forecast data help to clarify experts' expectations. First, Figure 4 disaggregates four types of forecasts. As in Figure 3, the modal prediction suggested a *positive* relationship between socioeconomic status and trust in police. Additionally, 30% of

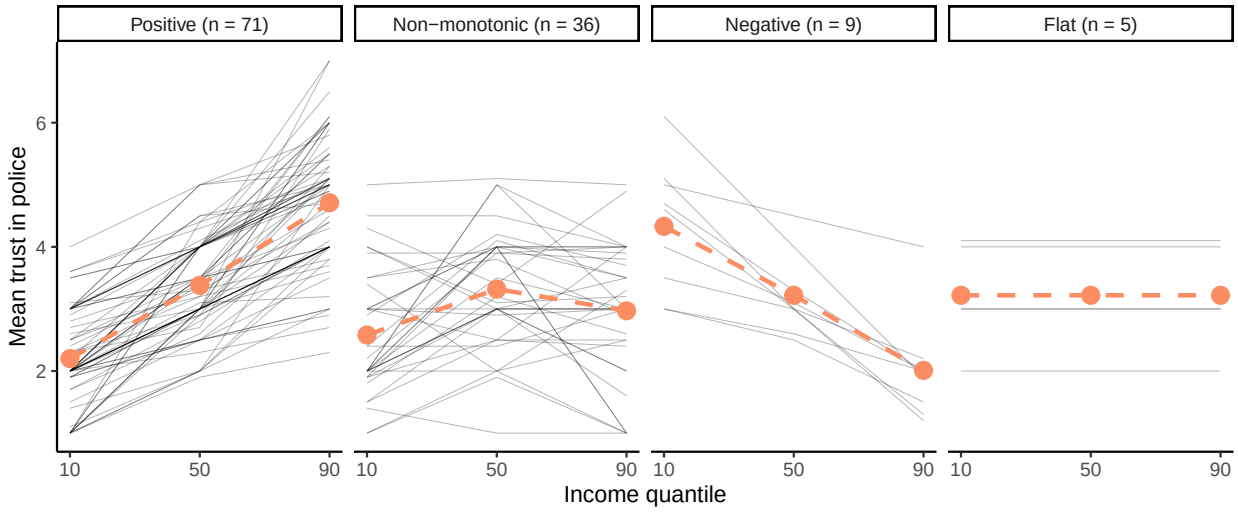

Figure 4: Classification of individual forecasts. Thin lines each represent individual forecasts. The points and dashed lines represent averages for each type of forecast.

forecasts posited a non-monotonic relationship. On balance, these forecasts suggested that a citizen with a median household income should trust the police more than the rich or the poor. Just 7.5% and 4% of forecasts posited a negative or flat relationship, respectively. Second, Figure A4 in the Appendix shows that the prediction of a positive relationship between income and trust in police is not specific to any country: we observe a similar pattern in all countries for which we have ≥ 8 predictions: Argentina, Brazil, Chile, Colombia, and Mexico.

5 Trust and experiences of police

To this point, we have established that, contrary to expert expectations and conventional wisdom based on evidence from the US, trust in police is generally weakly and negatively correlated with socioeconomic status in Latin America. We now seek to explain this discrepancy between expectations and empirical findings. To do so, we first examine additional data on respondent experiences with police and perceptions of security outcomes that fall within our conceptualization of trust in police as cognitive and relational. We then turn to alternative explanations that may explain our findings using alternative concepts of trust.

5.1 Signals of police trustworthiness

Our concept of trust in police centers on citizens updating their beliefs about police trustworthiness based on their experience of police behavior or observation of policing outcomes. To this end, our survey data identifies one experience—police solicitation of bribes (corruption)—and two policing outcomes—crime victimization and perceptions of safety—that allow for empirical investigation of these dynamics. While the experiences of crime victimization and feeling "unsafe" do not necessarily require active interaction with police, they at least imply that police failed to prevent crime or inspire a feeling of security. To the extent that these observations or experiences serve as signals of police trustworthiness, we are interested in how such signals vary in socioeconomic class. We first examine the frequency or degree to which citizens of different socioeconomic classes experience different signals of police trustworthiness. We then examine how these (self-reported) experiences influence respondent updating on police trustworthiness for respondents of different socioeconomic statuses.

In Figure 5, we plot the country-specific probability that a respondent from each socioeconomic decile reports each of the three binary signals of police trustworthiness, following (3). Given the prominence of the US case in setting expectations or conventional wisdom about trust in police, we use analogous LAPOP survey data from the US as a benchmark. The top three panels bin respondents into deciles using education, while the bottom three panels bin respondents using income. The first column panels in Figure 5 show that, for most Latin American countries,[6] the probability of self-reported crime victimization increases in socioeconomic status. Conversely, in the case of the US, the probability of reporting crime victimization does not change with class when it is measured with education, and it slightly *decreases* when it is measured with income.[7]

Whether the presence of a positive relationship between class and victimization in Latin Amer-

[6]The correlation between crime victimization and class is positive and statistically distinguishable from zero at the $\alpha = 0.05$ level for all Latin American countries and measures of socioeconomic status except for the income measure in Panama and Venezuela.

[7]The correlation between crime victimization and class in the US is negative and statistically distinguishable from zero at the $\alpha = 0.05$ level when class is measured with income, and positive but statistically indistinguishable from zero when measured with education.

ica is a surprising finding is unclear. Our evidence aligns with findings on urban property crime by Gaviria and Pagés (2002), albeit in a larger sample of urban/rural municipalities and with a broader range of crimes. On the other hand, poorer neighborhoods and municipalities are often distinguished by high rates of violent crime and insecurity. The panel survey from Medellín and administrative crime data for Medellín and Mexico City offer a potential reconciliation of these patterns by disaggregating crime victimization experience by type of crime. Figure A9 shows that self-reported exposure to violent crimes (especially homicide) is more common among the poor, whereas property crimes—which happen with higher frequency—disproportionately target the rich. Figures A11 and A10 show the same pattern when analyzing within-city variation in crime—as reported in administrative data—and income.

Despite the positive relationship between socioeconomic class and self-reported victimization identified in the data, the middle panels of Figure 5 suggest that the probability of feeling unsafe in a respondent's neighborhood generally does not vary in socioeconomic status for Latin American respondents. Conversely, for US respondents, the probability of feeling unsafe is uniformly lower than in any of the Latin American countries and decreases as socioeconomic class increases. As such, in contrast to the US case, the rate at which the poor and rich observe poor security outcomes in Latin America is ambiguous in the survey data. The disparities between survey and administrative data on violent crime in the case of Latin America leave open the possibility that what is perceived as an informational signal varies as a function of socioeconomic status.

The right panels of Figure 5 report the probability that respondents recall a police officer asking for a bribe in the last year. The positive relationship suggest that across countries and class measures, the rich are asked for bribes more frequently. From the perspective of rent maximization, these are the citizens from whom police may be able to extract larger sums. Conversely, the probability of reporting bribe solicitation is near zero at each decile for US respondents. This measure, of course, does not rule out police corruption in the US, but it does suggest that individual experiences of bribe solicitation are more highly circumscribed in the US. Collectively, the three panels of Figure 5 suggest that experiences with policing vary in social class in Latin America. Perhaps

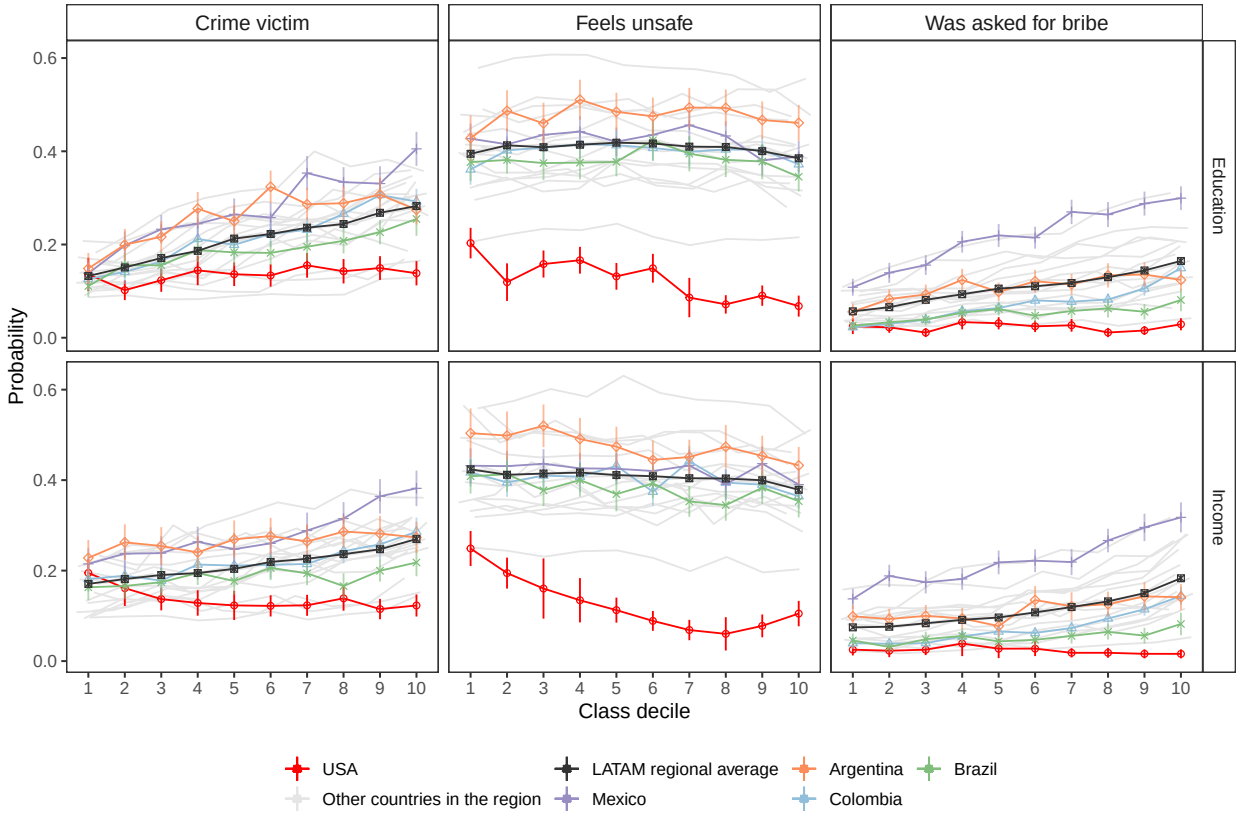

Figure 5: Lines show the estimated country-specific probability that a respondent from each education decile (top row) and income decile (bottom row) reports (from left to right): having been a victim of a crime during the past 12 months, perceiving the neighborhood as unsafe, and a police officer soliciting a bribe.

more surprising, the rich report more frequent experiences of crime victimization and bribe solicitation than the poor and assessments of neighborhood security outcomes do not substantially vary in socioeconomic status in the region.

5.2 Updating on police trustworthiness

We now show that exposure to each of these signals about the police is associated with the level of trust reported. Figure 6 plots the predicted level of trust in the police as a function of our three signals (denoted S_i) of police trustworthiness: (i) crime victimization in the past 12 months (left), (ii) feeling "unsafe" in their neighborhood (center), and (iii) whether a police officer asked for a bribe during the past 12 months (right). In each panel, the black line plots the mean level of trust, by decile of socioeconomic status, across the full sample. This line is very similar across all three vertical panels: the only (slight) differences come from variation in the presence of questions measuring the aforementioned signals across country-year survey waves. For each measure of socioeconomic status (the horizontal panels), trust decreases slightly and monotonically as income increases. We note that these means can be additively decomposed as follows:

$$E[\text{Trust}_i] = E[S_i = 1]E[\text{Trust}_i|S_i = 1] + E[S_i = 0]E[\text{Trust}_i|S_i = 0] \tag{4}$$

The orange and blue points and lines report our estimates of the conditional expectations in (4). The blue line ($E[\text{Trust}|S_i = 1]$) consistently falls below the orange line ($E[\text{Trust}|S_i = 0]$). This is consistent with our expectations—and the conventional wisdom—that poor security outcomes or treatment by police reduce trust in police. Importantly we see evidence consistent with updating—the difference in the orange and blue lines—for all deciles of socioeconomic status. The idea that poor treatment or outcomes reduce trust is consistent with many existing theoretical and empirical accounts of trust in government or government institutions (Hardin, 2003; Levi and Stoker, 2000). We do not find evidence that poor, middle-class, and rich Latin Americans update according to fundamentally different cognitive processes. This analysis does not, for example, support claims that citizen rationality (in this domain) varies in education or socioeconomic status.

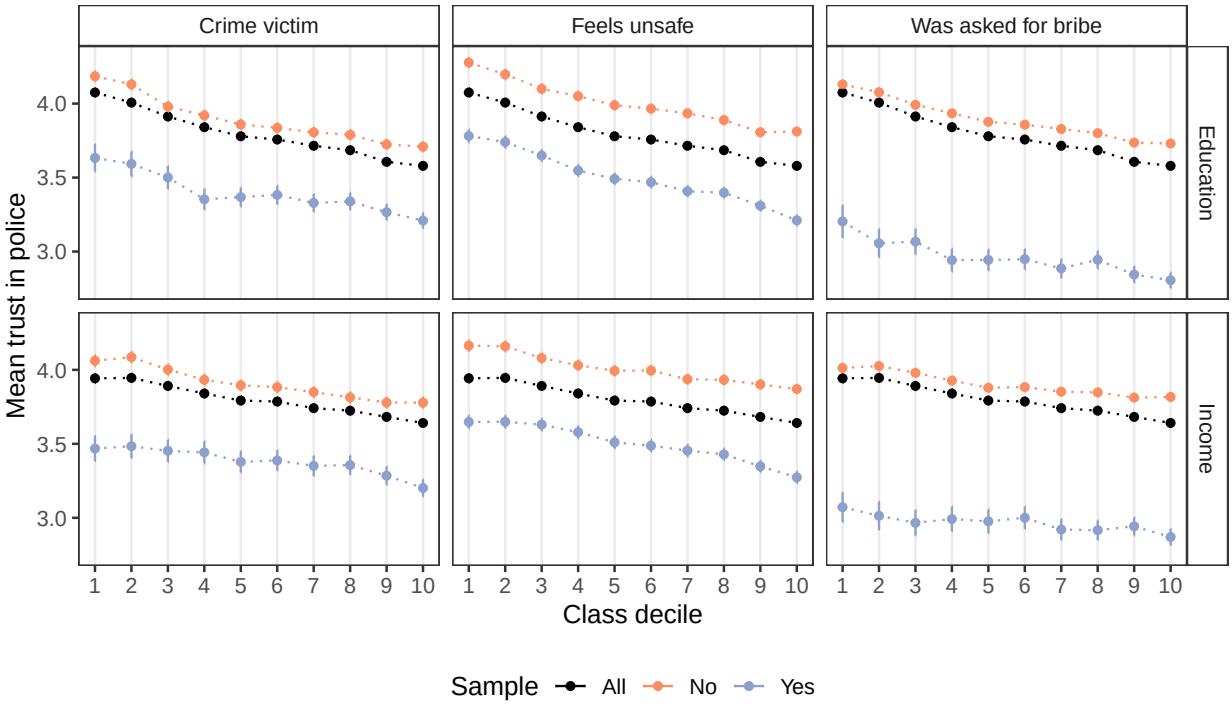

Figure 6: Predicted level of trust in police, by class decile, as a function of the following binary signals: crime victimization in the past 12 months (first panel), feeling unsafe in their neighborhood (second panel), and a police officer asked for a bribe during the past 12 months (third panel). Model was fit on pooled data from all country-waves.

The distance between the black lines and the orange/blue lines reflects the share of respondents that experience a given signal (e.g., $E[S_i = 1]$ from (4)). We see that crime victimization and especially bribe solicitation are *rare* at all levels of socioeconomic status (consistent with Figure 5). This is evident because the black line is much closer to the orange line, the conditional means for citizens who did not observe the signal in the last year. Indeed, in the full sample, only 21.0% and 10.6% of respondents reported crime victimization or bribe solicitation in the last year, respectively. The correlations in Figure 5 suggest that these outcomes of policing are increasing in socioeconomic status, which is evident from the growing distance between the black and orange lines as socioeconomic status increases. For example, moving from the lowest to the highest decile of education corresponds to (reported) increases from 13.5% to 28.1% in crime victimization and from 5.7% to 16.3% in bribe solicitation. While these differences in exposure to poor police behavior do increase the (negative) gradient of socioeconomic status and trust in police by pulling the black line toward the blue line, we note that these differences in isolation do not account for the negative gradient of the orange and blue lines.

Our language in this section has veered closer to causal language. Ideally, we would describe the updating in Figure 6 as the *effect* of different signals of police trustworthiness on trust in police at different income levels. However, we lack a research design capable of supporting such an inference with the LAPOP data. However, the panel surveys from Mexico, Medellín, and Chile permit estimation of average treatment effects on the treated (ATTs) for several closely related signals. We employ a two-way fixed-effect estimator and the fixed-effect counterfactual estimator proposed by Liu, Wang, and Xu (2022). From the latter, we report the ATT that weights respondents (units) equally. As in Figure 6, these signals are self-reported, though the questions vary slightly (as we report in Table A5). In Figure 7, we compare the estimated ATTs to associations (analogous to Figure 6) for the full sample of respondents. We show that all of the estimated ATTs are significant at the $\alpha = 0.05$ level and signed in the same direction as in the cross-sectional analyses from the Mexico, Medellín, Chile, and LAPOP samples: feeling unsafe, crime victimization, and viewing the police as corrupt reduces trust in police. However, the magnitudes of the ATTs relative to the

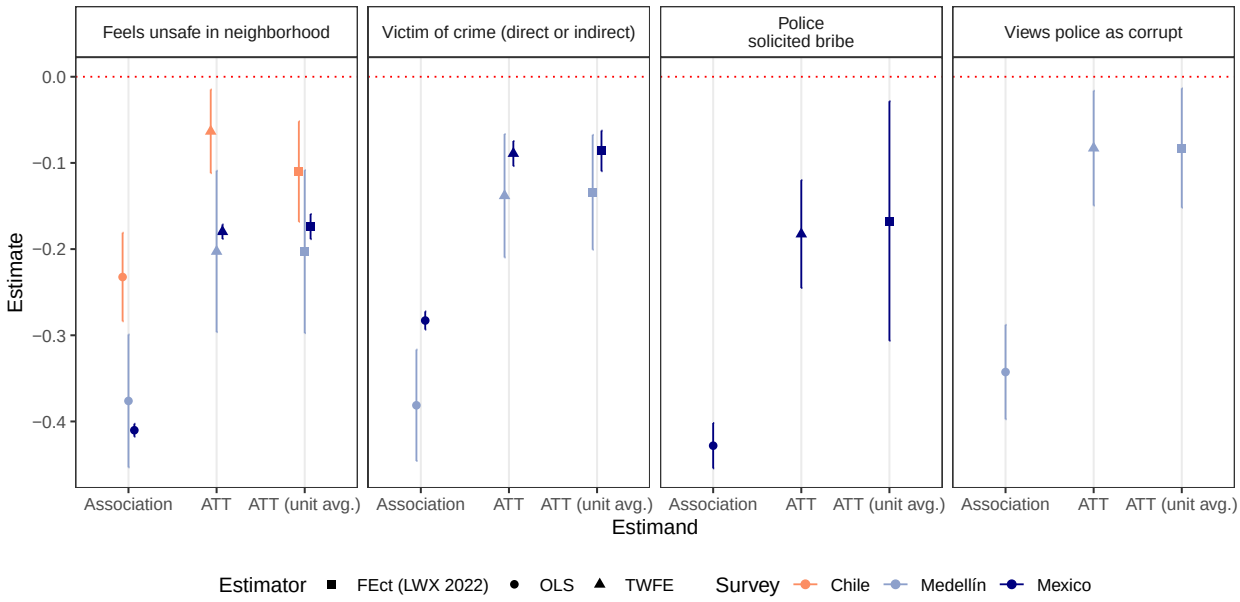

Figure 7: Estimates of pooled associations (across waves) to estimates of the average treatment effect (ATT) on the treated of signals analogous to those in Figure 6. LWX (2022) indicates the fixed effects counterfactual estimator proposed by Liu, Wang, and Xu (2022). 95% confidence intervals are calculated on standard errors clustered at the primary sampling unit.

pooled cross-sectional associations are reduced by 49-81%. This suggests that the magnitude of findings in Figure 6 is likely overstated, but that updating on the basis of poor security outcomes or abusive interactions with police leads to lower levels of trust in police.

5.3 Translating perceptions into signals of police trustworthiness

Our analysis of respondent updating on interactions with the police and security outcomes yields two main findings about the evolution of citizen trust in police. First, and consistent with conventional wisdom, poor experiences with police beget worse assessments of police trustworthiness. Importantly, both the direction and magnitude of updating do not appreciably vary in socioeconomic status. Second, and arguably more surprising, relative to poorer respondents, richer respondents self-report higher rates of crime victimization and bribe solicitation alongside similar assessments of the security situation in their neighborhoods.

We propose a theoretical framework in Appendix H to understand the sources of variation in citizen updating on police trustworthiness. Within our cognitive concept of trust, there are

three sources of such variation: different priors, different likelihoods, or different distributions of (positive or negative) signals about policing outcomes. In order for these sources of variation to explain differences in trust in policing by socioeconomic status, it must be the case that these attributes vary in socioeconomic status. Our evidence suggests that different distributions of signals are likely to account for the observed negative correlation between socioeconomic status and trust in police in Latin America.

Suppose first that a citizen has observed a signal of police trustworthiness—a poor security outcome or mistreatment by police—in time t. In this scenario, different citizens could emerge with different beliefs because (a) they had different priors; or (b) they update via different likelihoods. While we cannot directly measure priors or initial perceptions of the police (e.g., during a respondent's childhood), priors—or assessments of police trustworthiness at time $t = 0$—should play little role in characterizing trust in a cross-section of citizens when citizens are repeatedly observing signals of police trustworthiness over time. Analyzing trust in police by respondent age and socioeconomic status, as depicted in Figure A15, offers indirect evidence about how priors may vary in socioeconomic status. We see in this figure that older respondents, on average, express greater trust in the police. While intercepts vary by age group, the negative gradient of trust in police in socioeconomic status across age groups (different lines) is remarkably similar. If the negative correlation between socioeconomic status and trust in police were driven by differences in priors by socioeconomic status, we would expect the correlation to be *stronger* for younger respondents who have observed shorter histories of signals. We find no evidence that this is the case.

Figure 6 suggests that respondent likelihoods do not vary in socioeconomic status. The difference between the blue and orange lines—the conditional expectation of trust in police with and without a given signal—does not systematically or substantially vary in socioeconomic status (see Table A10). Further, Figure A12 in the Appendix shows results analogous to those of Figure 7 for the Mexican rotating panel, but disaggregated by administrative *estrato*. The estimated ATT of each signal of police trustworthiness on trust is remarkably similar for individuals of different

class *estratos*. These findings indicate that rich and poor respondents are not learning different things about police trustworthiness from the same signals. If this were the case, we would expect differences in posterior beliefs to vary systematically in socioeconomic status.

Instead, our evidence suggests that citizens of different socioeconomic statuses observe different signals at different rates. The rich report being more frequent victims of crime and bribe solicitation by police, while (on average) reporting similar rates of neighborhood insecurity. Here, there are two possibilities with regard to the relation between individual experiences and the rate of self-reported signals in Figure 5. First, self-reports may be a direct reflection of actual experience, meaning that the rich are subject to more police abuse and poor security outcomes than the poor. Second, poor and rich respondents may translate experiences with police into signals of police trustworthiness in different ways. This means that even if the rich were to experience crime victimization at equal rates and/or severity, they may be more likely to perceive themselves as victims of crime than poor respondents. Within our framework for updating, this means that the rich perceive negative signals of police trustworthiness at a higher rate than poor respondents, all else equal. This latter account introduces a *behavioral* component to our cognitive account of trust in police.

Directly distinguishing between these two possibilities—the rich receive worse service from the police than the poor versus a fixed set of experiences with the police yield different perceived signals—is quite challenging. On one hand, given well-known underreporting of crime in administrative data (Carr and Doleac, 2016; Jaitman and Anauati, 2020), it is very difficult to validate self-reports of crime victimization.[8] Even less data exist on bribes or objective security conditions in different communities, making the validation of other signals even more difficult. Further, we cannot know precisely how individual respondent experiences translate into self-reports in the LAPOP data. However, the aggregate data do provide indirect tests to discriminate between these two possibilities.

First, recall the disaggregation of offenses from Medellín in Figure A9 that suggests that rich

[8]Indeed, many researchers seek to measure victimization through surveys because of known limits to the quality of administrative crime data (Blair et al., 2021; Slough, 2023).

respondents report higher levels of property crimes (e.g., theft) whereas poor respondents report higher levels of violent crimes (e.g., homicide). Our binary measure of *any* crime victimization in the last year suggests similar levels of updating despite differences in the profile and severity of crimes that target these populations. Similarly, respondents across socioeconomic strata express similar perceptions of the security situation in their neighborhoods in Figure 5. Thus, despite higher rates of violence in poorer neighborhoods or communities, citizens express similar perceptions of safety in all neighborhoods. Both observations are consistent with rich citizens may applying a lower threshold for translating security outcomes into poor signals of police trustworthiness. Administrative data on homicide rates—the crime that is least likely to go unreported—from Medellín and Mexico City suggest that violent crime is decreasing in neighborhood income (see Figures A11 and A10). Thus, despite higher rates of violence in poorer neighborhoods or communities, citizens express similar perceptions of safety in all neighborhoods. Both observations are consistent with behavior in which rich citizens may apply lower threshold for translating security outcomes into poor signals of police trustworthiness.

Second, we consider measures of variation in trust in police by social class in Figure 8. This figure shows that, across both proxies of socioeconomic status, the standard deviation of trust in police is decreasing in socioeconomic status. This indicates that there is more variation in assessments of trust in police among poor than rich individuals. A similar pattern is evident in the US, though the standard deviation of trust in police is lower across the board than in most Latin American countries. This decreasing variance is consistent with a pattern of updating in which the rich observe more frequent signals of bad (or good) police performance than the poor. This is similarly consistent with a behavioral model in which the rich employ a lower threshold for translating good or bad experiences with the police into signals of police trustworthiness. Importantly, Figure 8 suggests that such a mechanism may be operative in the US as well as in Latin America.

Why, then, do we observe differences the gradient of trust in police in social class between the US and Latin America? If this behavioral mechanism predisposes the rich to negatively update more frequently than the poor, why is trust in police increasing in socioeconomic status in the US?

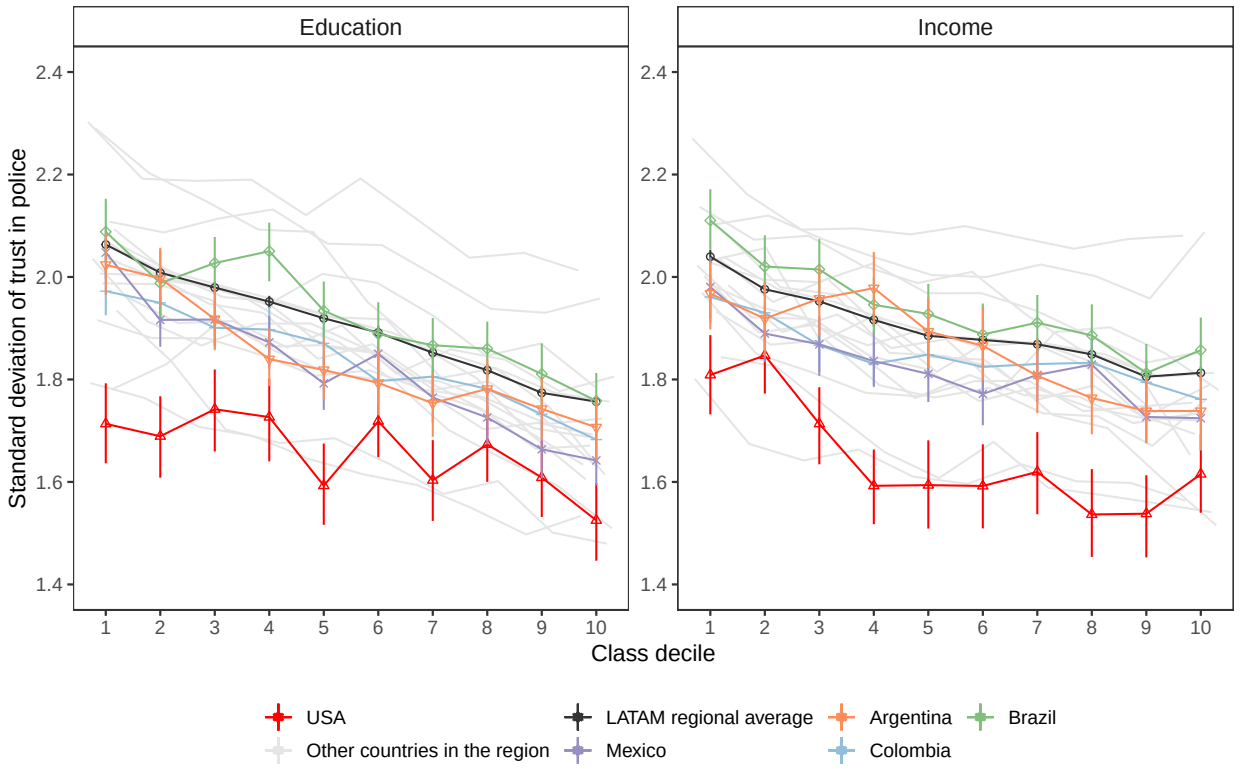

Figure 8: Figure shows the estimated standard deviation of trust in police by class decile for the country-specific and Latin America pooled samples and its the 95% bootstrapped confidence intervals.

Our framework suggests that other mechanisms may produce counterveiling effects on assessments of police trustworthiness. First, oft-documented discriminatory policing against racial minorities and/or poor citizens may be stronger in the US. This should produce a positive correlation between socioeconomic status and trust in police. Second, Figure 5 suggests that policing outcomes are generally *better* in the US: crime victimization rates are lower for almost all citizens, more citizens are secure in their neighborhoods, and bribe solicitation is nearly non-existent. If the levels of security outcomes are higher, a lower rate of the production of negative signals should attenuate differences in how these signals are perceived. Both possibilities should reduce the extent to which this behavioral mechanism yields a negative relationship between socioeconomic status and trust in police. Moreover, to the extent that better signals (lower crime, less petty corruption) are more common, we should also see lower variance in trust in the US context, as in Figure 8.

6 Alternative explanations

Thus far, we have marshaled empirical evidence from different data sources that align with our theoretical framework conceptualizing trust as relational and cognitive. However, in this section, we consider whether alternative explanations premised on the measurement of trust in police and its alternative conceptualizations could also account for the empirical patterns observed in the data. We consider three classes of alternative explanations: measurement concerns, institutional trust as a trait, and departures from Bayesian updating on police trustworthiness.

6.1 Artifacts of measurement

One possible cause of the divergence between experts' beliefs and what LAPOP survey data shows could be how class and trust are measured. Missingness may be correlated with socioeconomic status and trust in police. For example, if low-socioeconomic-status individuals who do not trust the police are less likely to respond or high-socioeconomic-status individuals who trust the police are more likely to respond to the relevant questions, we would underestimate the correlation between socioeconomic status and trust in police.

We conduct a bounding exercise in Appendix A6 to assess the maximum extent to which missingness could bias our conclusions. To generate worst-case bounds, note that correlation is bounded between -1 and 1. Since both trust in police and our measures of socioeconomic status are coded as Z-scores in (1), we can calculate these worst-case bounds for any respondent that reports at least one of the two measures. For this subset of missing observations (99.8% of all missing responses), we impute the observed Z-score for the missing Z-score (such that the imputed observation lies on the 45° line) to generate the worst-case upper bound. We then impute the negative of the observed Z-score (such that the imputed observation lies on the -45° line) to generate the worst-case lower bound. We estimate the correlation with all observed and imputed observations to generate worst-case bounds for the correlations reported in Figure 1. We estimate a worst-case bound using the education measure of [-0.11, -0.05], suggesting that missingness can have only a minimal effect on our conclusions. The worst-case bounds for the income measure of

social class are [-0.22, 0.13]. The width of these bounds is unsurprising given the different degrees of missingness. Nevertheless, it is useful to note that this worst-case upper bound on trust in the Latin American case is equivalent to the estimated correlation between social class and trust in the United States (reported above), showing how different these empirical patterns are in practice.

Alternatively, one may be worried about bias in the semantic content of respondents' opinions. Suppose LAPOP survey respondents have no prior opinion about how much they trust the police or other institutions. In that case, they might offer meaningless answers to questions about institutional trust when prompted. To test this possibility, we leverage changes in the ideological orientation of a country's subsequent government and all the questions that ask after trust in different political institutions. If respondents answer questions on institutional trust randomly, on average, their answers should not vary systematically between authorities within a single year, nor should they respond to changes in the political environment.

We first identify nine spells of right-wing presidencies experienced by countries in the LAPOP data. Next, using a difference-in-difference-inspired design (with repeated cross-sectional data), we compare within-country the correlation between trust in each authority and socioeconomic status in the survey immediately before the right-wing spell and in the subsequent two survey waves. Figure A7 in the Appendix shows the results. Having a right-wing president increases the correlation between trust in the president and income, as we would expect if respondents' self-reported evaluation of trust did respond to changes in who political actors *are*. Further, the correlation between trust in other political institutions, like parties, congress, the police, and the army, is also positively affected by the incoming of a right-wing president, but to a lesser degree than trust in the presidency itself. In contrast, trust in the Supreme Court, an institution less directly dependent on the president, remains more stable. The evidence points to respondents' assessment of trust in authorities being meaningful.

Finally, we draw upon a finding from the Colombian panel survey and associated field experiment to suggest that survey-measured trust in police predicts subsequent engagement with police. Hanson, Kronick, and Slough (2022) show that respondents that report the top category of trust

in the police (out of four categories) in a baseline survey are more than twice as likely to attend community-police meetings in beats assigned to treatment (18.4% vs. 8.4% of respondents). If survey measures of trust in the police were entirely random noise, we would not expect this alignment between reported trust and observed behavior.

6.2 Institutional trust as a trait

We have shown that in Latin America, trust in the police does not vary in class in the way most experts predicted. However, the alternative concepts of trust we described above may imply different predictions for the correlation between socioeconomic status and trust in police. If trust or propensity for trust were a trait rather than a relational expectation of police behavior, our findings could be explained by a weakly negative correlation between this predisposition and class. We note that a negative correlation between trust predispositions and socioeconomic status would cut against notions that higher trust or social capital promotes economic advancement (e.g., Putnam, Leonardi, and Nanetti, 1993).

To examine the possibility that our results are driven by stable individual differences in LAPOP respondents' *trust propensities*, we estimate the intra-class correlation between each respondent's trust in multiple institutions: the army, political parties, the sitting president, the supreme court, the national legislature, and the police. The intra-class correlation gives the ratio of between-respondent variance to the total variance in trust in these institutions. If the ICC were close to 1, it would suggest limited variance in an individual's assessment of multiple institutions, suggesting that institutional trust functions as a stable trait or predisposition. However, we do not observe a high ICC. Across all subjects in the LAPOP surveys, we estimate an ICC of .05 (95% CI: [.02,.23]). Disaggregating by country in Figure A8, we show similarly small estimates in all countries. It is therefore unlikely that a stable individual-level predisposition to trust that correlates with socioeconomic status can explain away our results.

6.3 Beliefs vs. Preferences

We have argued that trust should be characterized as a belief, but skeptical readers may argue that it is, instead, a manifestation of a preference about policing. Moreover, psychology and political psychology literature argues for the plausibility of an interaction between (prior) beliefs and preferences in the updating process (Kunda, 1987; Taber and Lodge, 2006; Little, Schnakenberg, and Turner, 2022). Specifically, the evolution of trust could be subject to motivated reasoning or directional motives, whereby trust becomes a function of an individual's prior preferences regarding policing (Ibid.). If this were the case, a respondent who prefers a policy that necessitates active police involvement may be motivated to hold more positive views of the police, thereby generating higher levels of trust in police, all else equal.

To gauge if respondents' beliefs about police trustworthiness may be shaped by their preferences over policing practices or policy, we characterize the relationship between socioeconomic status, self-described support for tough-on-crime or *mano dura* policing, and trust in police. Two expectations are worth clarifying. First, a motivated-reasoning or inference account of updating on police trustworthiness holds that pro-*mano dura* individuals have higher trust in police. Second, given the generally negative (if weak) correlations between socioeconomic status and trust in police reported in Figure 1, we would expect that the poor hold more favorable views of *mano dura* policies. The figure counters both expectations.

First, the left panel in Figure A14 in the Appendix shows a close-to-zero and *positive* correlation between income and support for tough-on-crime policing across all countries except Argentina and Uruguay. This finding is in line with recent research showing there is a positive relationship between (i) crime victimization and support for 'mano-dura' (Visconti, 2020) and (ii) income and urban property crime victimization (Gaviria and Pagés, 2002). The right panel in Figure A14 in the Appendix shows the predicted level of trust in police by class decile as a function of respondents' self-reported support for 'mano dura.' The black line plots the expected level of trust in police for respondents in each decile, and the blue line plots the conditional expectation for respondents in that decile who support 'mano dura.' In contrast, the orange line plots the conditional expectation

for respondents in that decile who are *unsupportive* of 'mano dura.' As we can see, the expected level of trust for individuals supportive of *mano dura* is lower than for individuals unsupportive of the measure across all income levels. Additionally, trust for both groups decreases at a similar rate. The results show the opposite empirical pattern we would expect to find if trust was largely driven by individuals' preferences, discounting the possibility of a motivated-reasoning explanation of our results.

7 Conclusion: Implications for the study of institutional trust

Conventional wisdom from the US suggests that racially and socioeconomically marginalized populations have lower trust in police than their advantaged counterparts. This apparent conventional wisdom has been extended to other contexts by police agencies that design interventions to increase trust (Blair et al., 2021) and, as our forecasting exercise shows, expert beliefs about trust in police. We show that descriptively, a distinct pattern obtains in Latin America, as a whole, and effectively all countries therein.

We argue that trust in the police should be viewed as a belief about the trustworthiness of the officer pool as a whole and that there is ample evidence that citizens update in a roughly Bayesian manner. We find suggestive evidence that low- and high-class citizens translate observations of police outputs or behavior into signals of trustworthiness in different ways. To our knowledge, this is a novel mechanism for explaining differences in trust in police. More research is needed to integrate this process for how citizens observe or interpret government outputs into an informational signal.

This mechanism has important implications for the design of policing interventions intended to increase trust or improve security outcomes. A central mechanism underpinning community policing, for example, is that greater exposure to local police officers increases citizen observation of these officers in a positive setting. But if trust is enhanced by *lack* of observation of, or translation of police outputs into, informational signals, community policing could well backfire. Moreover, if these meetings raise the expectations for police, it may have a similar effect of changing how

observations are converted into informational signals, reducing trust in police. This mechanism also suggests that improvements in police quality or reductions in police misbehavior/bias should have effects on trust that are *heterogeneous* in magnitude. In the Latin American context, these interventions may yield the highest trust increases among high-socioeconomic status populations, introducing a heretofore undocumented set of distributive considerations.

References

Altamirano, Melina, Sarah Berens, and Sandra Ley. 2022. "Security or Social Spending? Perceptions of Insecurity, Victimization, and Policy Priorities in Mexico and Brazil." *Political Studies* 1 (24): 00323217221096559.

Bhattacharya, Rajeev, Timothy M. Devinney, and Madan M. Pollutla. 1998. "A Formal Model of Trust Based on Outcomes." *The Academy of Managment Review* 23 (3): 459–472.

Blair, Graeme, Jeremy M. Weinstein, Fotini Christia, Eric Arias, Emile Badran, Robert A. Blair, Ali Cheema, Ahsan Farooqui, Thiemo Fetzer, Guy Grossman, Dotan Haim, Zulfiqar Hameed, Rebecca Hanson, Ali Hasanain, Dorothy Kronick, Benjamin S. Morse, Robert Muggah, Fatiq Nadeem, Lily L. Tsai, Matthew Nanes, Tara Slough, Nico Ravanilla, Jacob N. Shapiro, Barbara Silva, Pedro C. L. Souza, and Anna M. Wilke. 2021. "Community Policing Does Not Build Citizen Trust in Police or Reduce Crime in the Global South." *Science (New York, N.Y.)* 374 (6571): eabd3446.

Brinks, Daniel M. 2007. *The Judicial Response to Police Killings in Latin America: Inequality and the Rule of Law*. Cambridge: Cambridge University Press.

Brinks, Daniel M. 2019. "Access to What? Legal Agency and Access to Justice for Indigenous Peoples in Latin America." *The Journal of Development Studies* 55 (3): 348–365.

Britto Ruiz, Diana, and Jorge Ordóñez Valverde. 2005. "Las prácticas de distinción social." *Polis. Revista Latinoamericana* (11).

Cao, Liqun, and Jihong Solomon Zhao. 2005. "Confidence in the Police in Latin America." *Journal of Criminal Justice* 33 (5): 403–412.

Carr, Jillian, and Jennifer L. Doleac. 2016. "The Geography, Incidence, and Underreporting of Gun Violence: New Evidence Using Shotspotter Data.".

Citrin, Jack, and Laura Stoker. 2018. "Political Trust in a Cynical Age." *Annual Review of Political Science* 21: 49–70.

COES, Centre for Social Conflict and Cohesion Studies. 2022. "Estudio Longitudinal Social de Chile 2016-2021.".

Cruz, José Miguel. 2015. "Police Misconduct and Political Legitimacy in Central America." *Journal of Latin American Studies* 47 (2): 251–283.

Decker, Scott H. 1981. "Citizen Attitudes Toward the Police: A Review of Past Findings and Suggestions for Future Policy." *Journal of Police Science and Administration* 9: 80–87.

DellaVigna, Stefano, Devin Pope, and Eva Vivalt. 2019. "Predict Science to Improve Science." *Science (New York, N.Y.)* 366 (6464): 428–429.

DellaVigna, Stefano, Nicholas Otis, and Eva Vivalt. 2020. "Forecasting the Results of Experiments: Piloting an Elicitation Strategy." *AEA Papers and Proceedings* 110: 75–79.

Gallup, Inc. 2021. "In U.S., Black Confidence in Police Recovers From 2020 Low." https://news.gallup.com/poll/352304/black-confidence-police-recovers-2020-low.aspx.

Gasparini, Leonardo, and Nora Lustig. 2011. "The Rise and Fall of Income Inequality in Latin America." In *The Oxford Handbook of Latin American Economics*, ed. José Antonio Ocampo, and Jaime Ros. Oxford University Press p. 0.

Gaviria, Alejandro, and Carmen Pagés. 2002. "Patterns of Crime Victimization in Latin American Cities." *Journal of Development Economics* 67 (1): 181–203.

Gerbasi, Alexandra, and Karen Cook. 2009. "Trust.". Oxford University Press pp. 218–241.

Gerring, John. 2012. "Mere Description." *British Journal of Political Science* 42 (4): 721–746.

González, Yanilda, and Lindsay Mayka. 2022. "Policing, Democratic Participation, and the Reproduction of Asymmetric Citizenship." *Forthcoming, American Political Science Review* .

González, Yanilda María. 2020. *Authoritarian Police in Democracy: Contested Security in Latin America*. Cambridge University Press.

Hanson, Rebecca, Dorothy Kronick, and Tara Slough. 2022. "Preaching to the Choir: A Problem of Participatory Interventions.".

Hardin, Russel. 2003. "Trust in Government.". Russell Sage Foundation chapter 1, pp. 9–27.

Hoffman, Kelly, and Miguel Angel Centeno. 2003. "The Lopsided Continent: Inequality in Latin America." *Annual Review of Sociology* 29 (1): 363–390.

INEGI, Instituto Nacional de Estadística y Geografía. 2024. "Encuesta Nacional de Seguridad Pública Urbana (ENSU).".

Jaitman, Laura, Rodrigo Soares, Mauricio Olavarría-Gambi, and Roberto Guerrero Compeán. 2015. "The Welfare Costs of Crime and Violence in Latin America and the Caribbean.".

Jaitman, Laura, and Victoria Anauati. 2020. "The Dark Figure of Crime in Latin America and the Caribbean." *Journal of Economics, Race, and Policy* 3 (1): 76–95.

Johnson, Stephen, Johanna Mendelson Forman, and Katherine Bliss. 2012. Police Reform in Latin America: Implications for U.S. Polic. Technical report Center for Strategic and International Studies.

Karim, Sabrina. 2020. "Relational State Building in Areas of Limited Statehood: Experimental Evidence on the Attitudes of the Police." *American Political Science Review* 114 (2): 536–551.

Kunda, Ziva. 1987. "Motivated Inference: Self-serving Generation and Evaluation of Causal Theories." *Journal of Personality and Social Psychology* 53 (4): 636–647.

LAPOP, Lab. 2022. "The AmericasBarometer." www.vanderbilt.edu/lapop.

Levi, Margaret, and Laura Stoker. 2000. "Political Trust and Trustworthiness." *Annual Review of Political Science* 3: 457–507.

Little, Andrew T., Keith E. Schnakenberg, and Ian R. Turner. 2022. "Motivated Reasoning and Democratic Accountability." *American Political Science Review* 116 (2): 751–767.

Liu, Licheng, Ye Wang, and Yiqing Xu. 2022. "A Practical Guide to Counterfactual Estimators for Causal Inference with Time-Series Cross-Sectional Data." *American Journal of Political Science* n/a (n/a).

Macaulay, Fiona. 2012. "Cycles of Police Reform in Latin America." In *Policing in Africa*, ed. David J. Francis. New York: Palgrave Macmillan US pp. 165–190.

Macdonald, John, and Robert J. Stokes. 2006. "Race, Social Capital, and Trust in the Police." *Urban Affairs Review* 41 (3): 358–375.

Magaloni, Beatriz, Edgar Franco-Vivanco, and Vanessa Melo. 2020. "Killing in the Slums: Social Order, Criminal Governance, and Police Violence in Rio de Janeiro." *American Political Science Review* 114 (2): 552–572.

Malone, Mary Fran T., and Lucia Dammert. 2021. "The Police and the Public: Policing Practices and Public Trust in Latin America." *Policing and Society* 31 (4): 418–433.

Mondak, Jeffery J., Damarys Canache, Mitchell A. Seligson, and Matthew V. Hibbing. 2011. "The Participatory Personality: Evidence from Latin America." *British Journal of Political Science* 41 (1): 211–221.

Muggah, Robert, and Katherine Tobon. 2018. "Citizen Security in Latin America: Facts and Figures." *Security Research Hub Reports* .

Munger, Kevin, Andrew M. Guess, and Eszter Hargittai. 2021. "Quantitative Description of Digital Media: A Modest Proposal to Disrupt Academic Publishing." *Journal of Quantitative Description: Digital Media* 1: 1–13.

O'Donnell, Guillermo. 1999. "The (Un)Rule of Law and the Underprivileged in Latin America.". Notre Dame Press chapter Polyarchies and the (un) rule of law in Latin America: A Partial Conclusion.

Ojeda, Christopher. 2016. "The Effect of 9/11 on the Heritability of Political Trust." *Political Psychology* 37 (1): 73–88.

Peyton, Kyle, Michael Sierra-Arévalo, and David G. Rand. 2019. "A Field Experiment on Community Policing and Police Legitimacy." *Proceedings of the National Academy of Sciences* 116 (40): 19894–19898.

Phelps, Edmund. 1972. "The Statistical Theory of Racism and Sexism." *American Economic Review* 62: 659–61.

Putnam, Robert D., Robert Leonardi, and Raffaella Y. Nanetti. 1993. *Making Democracy Work : Civic Traditions in Modern Italy*. Princeton, UNITED STATES: Princeton University Press.

PytlikZillig, Lisa M., and Christopher D. Kimbrough. 2016. "Interdisciplinary Perspectives on Trust: Towards Theoretical and Methodological Integration.". New York: Springer chapter Consensus on Conceptualizations and Definitions of Trust: Are We There Yet?, pp. 49–70.

Sabatini, F. 2006. The Social Spatial Segregation in the Cities of Latin America.

Skogan, Wesley G. 2005. "Citizen Satisfaction with Police Encounters." *Police Quarterly* 8 (3): 298–321.

Slough, Tara. 2023. "Phantom Counterfactuals." *American Journal of Political Science* 67 (1): 137–153.

Slough, Tara, and Scott A. Tyson. 2023. "External Validity and Meta-Analysis." *American Journal of Political Science* 67 (2): 440–455.

Soares, Rodrigo R., and Joana Naritomi. 2010. "Understanding High Crime Rates in Latin America: The Role of Social and Policy Factors." In *The Economics of Crime: Lessons For and From Latin America*. University of Chicago Press pp. 19–55.

Taber, Charles S., and Milton Lodge. 2006. "Motivated Skepticism in the Evaluation of Political Beliefs." *American Journal of Political Science* 50 (3): 755–769.

Vilalta, Carlos. 2020. "Violence in Latin America: An Overview of Research and Issues." *Annual Review of Sociology* 46 (1): 693–706.

Villarreal, Andrés. 2010. "Stratification by Skin Color in Contemporary Mexico." *American Sociological Review* 75 (5): 652–678.

Visconti, Giancarlo. 2020. "Policy Preferences after Crime Victimization: Panel and Survey Evidence from Latin America." *British Journal of Political Science* 50 (4): 1481–1495.

Weitz-Shapiro, Rebecca, and Matthew S. Winters. 2017. "Can Citzens Discern? Information Credibility, Political Sophistication, and the Punishment of Corruption in Brazil." *Journal of Politics* 79 (1): 60–74.