Review of Economic Studies (2024) **00**, 1–37 doi:10.1093/restud/rdae112

Advance access publication 24 December 2024

Consumer Surplus of Alternative Payment Methods

Fernando Alvarez
University of Chicago, USA

and

David Argente
Yale University, USA

First version received June 2022*; Editorial decision August* 2024*; Accepted December* 2024 *(Eds.)*

This paper estimates the consumer surplus from using alternative payment methods. We use evidence from Uber rides in Mexico, where riders have the option to use cash or cards to pay for rides. We design and conduct three large-scale field experiments, which involved approximately 400,000 riders. We also build a structural model which, disciplined by our new experimental data, allows us to estimate the loss of private benefits for riders when a ban on cash payments is implemented. We find that Uber riders who use cash as means of payment either sometimes or exclusively suffer an average loss of approximately 40–50% of their total trip expenditures paid in cash before the ban. The magnitude of these estimates reflects the intensity with which cash is used in the application, the shape of the demand curve for Uber rides, and the imperfect substitutability across means of payments. Welfare losses fall mostly on the least-advantaged households, who rely more heavily on the cash payment option.

Key words: Macro development policy, Means of payment, Ban on cash

JEL codes: O1, O2, E4, E5

1. INTRODUCTION

Some academics and policymakers have recently advocated for a cashless economy to address the prevalence of criminal activities and tax avoidance (*e.g.* Rogoff 2017). For example, in India, a 2016 demonetization plan was enacted to, in part, remove certain large-denomination bills from circulation.[1] In Mexico, until a November 2018 ruling by the Supreme Court disallowing cash bans, several of the country's largest cities banned cash payments for app-based ride-sharing firms like Uber.[2] Has the time come to phase out cash? This question is particularly relevant

1. See Chodorow-Reich *et al*. (2020) for a description and evaluation of its macroeconomic effects.
2. See the decision of the "Suprema Corte de Justicia de la Nacion" in the case of "Ley de Movilidad Sustentable pare el Estado de Colima" in October of 2018.

The editor in charge of this paper was Nir Jaimovich.

for low- and middle-income countries, where the least-advantaged households tend to use cash much more, and where policies that restrict the use of cash could limit economic access for the poor and could have important distributional consequences.

Despite the renewed attention to the role of cash in the economy, the debate over the consequences of phasing out cash is far from settled. In fact, it is challenging to estimate the private costs of policies limiting the use of cash. Such estimates would require detailed information on cash transfers and on people without access to banking services, especially in developing countries given their prevalence. They would also require both variation in prices, to estimate the elasticity of demand, and crucially, information about the functional form of the demand curve, since consumer surplus estimates are sensitive to the magnitude of choke prices. An accurate estimate of consumer surplus must also account for the costs of adopting cashless payment methods like debit or credit cards.

This paper builds a structural model and combines it with three large-scale field experiments to overcome these challenges and estimate the private benefits from using alternative payment methods. To do so, we use evidence from the effects of allowing for different payment methods in a ride-sharing app. In more than 400 cities worldwide, Uber allows its riders to select cash as a payment method—in the same way that their app allows riders to set more than one payment card as a means of payment. However, the use of cash to pay for Uber, in Mexico and other countries such as Panama and Uruguay, has encountered severe restrictions. Cash was originally not allowed in several cities in Mexico (for example, in Mexico City or Querétaro) and was later banned in other cities, such as Puebla and San Luis Potosí. Motivated by these recent policies, we estimate the consumer surplus loss caused by banning cash as a payment method in a city where it was available.

We develop a model of an Uber rider in a city where she can purchase Uber trips paid in cash, Uber trips paid in card, an arbitrary number of other goods that might be complements or substitutes of Uber, and an outside good.[3] We assume that the utility function is quasi-linear in the outside good, an assumption that we test and is not rejected by the data. We assume weak separable preferences so that we can define the demand for "composite Uber trips," an aggregate of both types of trips. Furthermore, we model both the extensive-margin choice of registering a card to gain access to both payment methods, and the intensive-margin choice of how many trips to take with each of the available payment methods. Thus, we distinguish between a ban's effect on riders that use both payment methods (mixed riders), and the effect on riders that do not register a payment card in the app (pure cash riders). We also allow for heterogeneity among riders in their preferences for paying with cash or card, and in the cost of registering a card in the application.

The calculation of consumer surplus is guided by a general result in demand theory: when prices of alternative options are held fixed, as documented in previous studies for this market after a ban or introduction of cash (*e.g.* Alvarez and Argente 2022), cross-price elasticities and quantities demanded for these alternatives do not influence calculation.[4] This result enables us to estimate consumer surplus without the need to measure the quantities. The riders' consumer surplus from paying Uber in cash can be obtained by integrating the area under the demand curve, starting with the current price and up to the choke price at which the demand reaches zero. In theory, one could estimate the demand for Uber paid in cash by imposing increasingly higher

3. Since users can use either a credit or debit card to pay for Uber rides, in the rest of the paper, we refer to card payments as those conducted with either a debit or a credit card.

4. In other words, the supply of drivers is elastic at the relevant time horizon. According to Hall *et al.* (2023), following a base fare increase, driver hours worked increase on both the extensive and intensive margins. After about 8 weeks, there is no clear difference in the driver's gross average hourly earnings rate.

prices until reaching the choke price. In practice, however, this exercise is almost impossible to implement using exclusively field experiments.[5] We overcome this challenge by using our theory to inform the design and implementation of three large-scale field experiments and a survey, involving over 400,000 riders in the State of Mexico. Our model allows us to extrapolate the demand curve from the variation in demand that we observed as prices were reduced.[6]

In the first experiment, we targeted *mixed riders* to estimate the elasticity of substitution between paying for trips with cash or with a card. We varied prices using discounts for trips paid in cash or discounts for trips paid with a card. The experiment had a total of six treatment groups, each with about 20,000 riders who had registered a card with Uber. These riders received discounts of either 10% or 20%. Some of them received discounts for paying with cash, some received discounts for paying with a card, while others received discounts regardless of their payment method. A control group of approximately 90,000 riders received no discounts. We estimate the elasticity of substitution to be about three. We also use the price discounts given regardless of the payment method to estimate the price elasticity for Uber rides for *mixed users*, which can be as large as 1.1, evaluated at current prices.

We combine these findings with our structural model to produce theoretically based estimates of the consumer surplus for mixed riders. This is equivalent to increasing the price in cash from its current value to infinity—or to the choke price at which there will be no more trips paid in cash. The effect of this increase can be decomposed into two parts. The first part is the change in the choice of payment for a given number of trips, which depends on the elasticity of substitution between payment methods, as well as the share of trips paid in cash. The second part is given by the change in the ideal price index for Uber trips caused by the cash ban, which depends on the price elasticity of Uber trips. Integrating across all types of mixed users, we find that the consumer surplus of an Uber ride falls by more than 25% of mixed users' expenditures on Uber when cash payments are banned. These users represent approximately 50% of the total Uber customers in the State of Mexico.

Our second and third field experiments were intended to estimate the consumer surplus of *pure cash riders*, who account for about 25% of Uber's user base in the State of Mexico. In the event of a ban on cash payments, pure cash users must either cease to use Uber, and lose the entire consumer surplus of using the app, or register a card at some cost. The second experiment allows us to estimate the first part of this consumer surplus loss. We randomize the amount of the discount offered to pure cash riders, measuring the effect on purchases in terms of the length and number of trips. We use four treatment groups of 23,000 riders, each with discounts of 10%, 15%, 20%, and 25%, and a control group of 56,000 riders. The four treatments cover several price points so that we can learn about the shape of the demand curve. From this experiment, we find that the price elasticity for *pure cash riders* is about 1.3, evaluated at current prices.

We use the third experiment to estimate the distribution of fixed costs of adopting a card, in order to adjust the consumer surplus for pure cash users that decide to remain in the application in the event of a ban on cash. Pure cash users were offered a small reward of credit for future trips, contingent on registering a card in the application. This experiment included six treatment groups of about 20,000 riders each. In return for registering a payment card, we offered rewards equivalent to about three, six, or nine times a user's average weekly expenditure on Uber.

5. While Uber allowed us to implement discounts in experiments, its policies did not permit an increase in absolute prices. Raising prices until demand reaches zero would likely harm its customer base. However, Uber did permit us to experiment with increasing relative prices. To the best of our knowledge, there is no research that conducts experiments by raising the price of a specific good or service as a treatment.

6. Other recent examples of work integrating structural models with experimental evidence are Kaboski and Townsend (2011) and Buera *et al*. (2021a).

The same reward was offered to one group of riders for registering a card in less than a week, and to another group of riders with a time limit of 6 weeks. We consider these two-time frames to test for the hypothesis that riders may not register a card in the application *even if they have one*, because it is difficult to obtain a card in Mexico within 1 week, but reasonable within six. The temporal migration patterns across user types (*e.g.* pure cash riders becoming mixed riders) inform us about whether the likely margin of response is to register a card that the riders already have, or to obtain a new card. We find that the smallest incentives *double* the rate at which riders register a card, compared with the control group. We also find that only slightly more pure cash users register a payment card with a 6-week window, as most excess migration to cards occurs in the first week. The latter finding suggests that migration from cash to card induced by small rewards is mostly driven by riders registering payment cards that they already own.

With the results of the second and third experiments targeting pure cash users and the elasticity of substitution between payment methods, we calculate the consumer surplus for pure cash users. This estimation also requires an estimate of the rate at which pure cash users *return* to the application after a ban on cash payments. We draw this estimate from a case study of such a ban in the city of Puebla, Mexico conducted by Alvarez and Argente (2022). In the most-extreme case, if no pure cash users register a payment card, then the effect of a cash-payment ban is to erase the entire consumer surplus these users enjoy when purchasing Uber rides, which we estimate to be at least as much as 47% of a user's *total expenditures* on Uber. We also know that roughly 30% of pure cash riders switch to using cards after a ban. With this figure and the above results, we estimate that a previously pure cash user who registers a card will see her consumer surplus decrease by about 44% of her Uber expenditures. Aggregating both groups, we find a ban on cash leads to a large loss in consumer surplus for pure cash riders, equal to about 46% of their total Uber expenditures. Given that lower-income households are more-likely to rely on cash as the primary mode of payment, the private costs of a ban on cash will fall disproportionately on such households.

As a complement to our estimate of Uber trip price elasticity, we considered two *additional* independent price experiments conducted by Uber. These experiments were not designed for our specific purposes, yet the price elasticities observed align closely with our field experiment findings. Notably, one of these experiments involved longer-lasting discounts, which better approximate permanent price changes and yielded similar elasticities. To further validate our findings, we explored a quasi-natural experiment in Uber Panama, where prices increased significantly due to changes in costs and licensing requirements, resulting in a substantial reduction in driver supply. This event allowed us to estimate price elasticity for Uber trips with significant *price increases*, and our results aligned with those from our model. Additionally, we employed a survey instrument to gather evidence on users' choke prices. Over 6,000 users responded to this survey, sent almost a year after the experiments. They were asked about their responses to various price changes, including substantial increases. We found that the reported elasticities in the survey closely matched the revealed preference elasticities, and when comparing these reported choke prices to our model's estimates, they exhibited remarkable consistency, further validating our structural assumptions.[7]

Taken together, our results show that the loss in consumer surplus due to a ban in cash is large, at least 50% of total expenditures on Uber paid in cash or approximately 0.8% of

7. Our choice of the functional form for demand, featuring a constant semi-elasticity, *aligns* with the local convexity observed in the relationship between trips and prices, as well as with the presence of a finite choke price in the experimental and survey data.

annual per capita income in the State of Mexico.[8] The magnitude of our estimate reflects the following: First, we argue that any effect on riders who exclusively use cards before the ban on cash is likely to be small. Second, in the State of Mexico, pure cash riders account for 20% of total expenditures, and 50% of total expenditures are from mixed riders, who pay about 42% of their fares in cash. Third, while riders that use both means of payment do react to changes in the relative prices of the two payment methods, they view the payment methods as imperfect substitutes. Fourth, while riders without registered cards react to incentives, a significant fraction of them face large costs for registering a card. Fifth, we find that the demand for Uber trips is relatively inelastic, regardless of payment method. Importantly, consumer surplus losses fall mostly on the least-advantaged households, who rely more heavily on the cash payment option.

2. RELATED LITERATURE AND CONTRIBUTION

While the literature has used *observational* data to document how the option for cash payment affects rides, prices, and the use of other payment methods (Alvarez and Argente 2022), this paper is the first one to conduct consumer surplus evaluation. Consumer surplus evaluation is complex, as it must incorporate mixed users and pure cash users, and both intensive and extensive margins. The analysis of consumer surplus also demands building new structure and producing estimates from *experimental* data, which are at the core of this paper. In particular, to estimate the consumer surplus losses of pure cash users who drop from the application after the ban and lose the entire consumer surplus of using the service, we need *variation in prices* to estimate the elasticity of demand for Uber. To calculate the consumer surplus of mixed users, we need variation in prices by payment method to estimate the relevant elasticity of substitution. We also need variation to approximate the distribution of the cost to register a card in the application, as well as information of the appropriate *functional* form of the demand curve, since consumer surplus estimates are very sensitive to the magnitude of choke prices. All the former are obtained from the experimental undertaking in this paper, and are not present in any previous work.

The approach of this paper is closely related to recent studies that assess the welfare implications of policies in developing economies by combining randomized controlled trials and natural experiments with structural modelling (*e.g.* Kaboski and Townsend 2011; Buera *et al*. 2021a). Buera *et al*. (2021b) review this literature and Townsend (2020) highlight how macroeconomic theory and empirical research can complement one another to improve macro development policy in payment systems. Our experimental design represents an innovation in *tailoring* such large-scale field experiments with a structural model in mind. This approach differs from the previous related work, which relied solely on structural models (*e.g.* Alvarez and Lippi 2017; Briglevics and Schuh 2020; Alvarez *et al*. 2022), and it is closer in spirit to Chodorow-Reich *et al*. (2020), who use the Indian demonetization as a natural experiment.

The paper also contributes to the literature on money demand, with its focus on the effects of availability and optimal choices of means of payment. Examples of earlier theoretical studies on the choice of payment are the cash-credit model in Lucas (1987), the model of multiple payment methods in Prescott (1987), and studies that followed: Whitesell (1989), Lacker and Schreft (1996), Freeman and Kydland (2000), Lucas and Nicolini (2015), Koulayev *et al*. (2016),

8. Cohen *et al*. (2016) use a discontinuity design based on the rounding of prices dictated by the surge algorithm to estimate the consumer surplus of Uber for three large U.S. cities and find it to be about 1.6 of the expenditure of Uber riders. This difference is in large part explained by the different elasticity that they estimate for U.S. riders versus users in the State of Mexico. In our case. the price elasticity at the current equilibrium values is 1.3 for pure cash users, 1.1, for mixed users, and 0.7 for pure card users. In Cohen *et al*. (2016), the price elasticity is below 0.55.

and Stokey (2019).[9] Several mechanisms might explain the relatively inelastic substitutability between cash and cards. For instance, the popularity of paying for other goods with cash in Mexico encourages consumers to use cash for Uber rides, even those that own payment cards, as in Deviatov and Wallace (2014) and Alvarez and Lippi (2017).[10]

Our well-identified estimate of the elasticity of substitution between cash and card payments for a given good is, in itself, a contribution to the empirical study of money demand. To the best of our knowledge, ours is the first estimate of this parameter using experimental data.[11] Furthermore, the experimental variation used to estimate the elasticity of demand for Uber rides in this paper also allows us to draw upon the quasi-natural experiments in Alvarez and Argente (2022), which provide information on the long-run elasticity of substitution across payment methods, given the estimates of the elasticity of demand obtained in this paper. This alternative estimate for the elasticity of substitution complements the estimates from experimental data, enabling us to provide a long-run estimate of the consumer surplus lost after a ban on cash payments, amounting to 38% of total cash expenditures.

Our paper is also related to research studying the adoption of debit and payment cards (*e.g.* Borzekowski *et al*. 2008; Yang and Ching 2014), which has focused on identifying the determinants of consumers' adoption decisions. Our work contributes to this literature with experimental data about the distribution of adoption costs among consumers.

In summary, this paper develops and estimates a structural model suitable for the evaluation of different margins of adjustment across users and provides the first welfare estimates of alternative payment methods. Additionally, this paper conducts large-scale field experiments tailored to generate variation which is essential to estimate the relevant parameters for this calculation. These are the first experimental estimates of the elasticity of substitution across payment methods. We also provide estimates of the fixed cost of adopting or registering a payment-card. All these estimates can be used beyond our application for the analysis of policies attempting to encourage or discourage payment methods.[12] We also provide evidence that reported elasticities in a survey are informative about the revealed preference elasticities, which contributes to the recent literature examining the external validity of survey instruments as low-cost alternatives to experimental evidence.[13]

3. INSTITUTIONAL BACKGROUND

At its launch in 2010, Uber was notable for offering users the ability to easily hail a car and pay for the ride with a credit or debit card registered in a mobile-phone app. As Uber expanded to

9. Other related work is the search-theoretical literature considering money as a payment method, largely started by Kiyotaki and Wright (1989), which incorporates credit payments as in Kocherlakota (1998), Lagos and Wright (2005), or Wang *et al*. (2017).

10. According to the 2018 National Survey of Financial Inclusion (ENIF), cash is the primary payment method in Mexico. Around 95% of all transactions below 25 USD and 87% of transactions above 25 USD are conducted in cash. The share of transactions paid in cash is above 90% for most goods in the economy.

11. Very few studies have considered the behaviour of households when faced with a differential cost in means of payment. Klee (2008) estimates the transaction times needed for different payment methods in grocery stores using data from time-stamped cash registers; this study observed no variation in prices. Humphrey *et al*. (2001) use aggregate semiannual time series from Norway during the 1990s and observed variations price across payment methods to estimate patterns of substitution between cash, checks, and debit cards. Ching and Hayashi (2010) estimate the effects of payment-card rewards on consumer choice of payment methods in retail stores. Amromin *et al*. (2006) use a one-time change in toll booth prices on a Chicago highway, which depended on whether payment is made using cash or a transponder.

12. Supplementary Material, Section H analyses a ban on *card* payments in Argentina.

13. Examples include Karlan *et al*. (2016), Parker and Souleles (2019), Méndez and Van Patten (2021), and Hainmueller *et al*. (2015).

cities across the globe, it began to accept cash payment during a 2015 pilot program in Hyderabad, India. This pilot program expanded Uber's user base by opening access to consumers who prefer to use cash, because they have no access to a bank or card or because they prefer not to register a card with Uber. Following the success of that pilot, Uber extended the option to four more cities in India. By the end of 2016, the cash-payment option was made available in over 150 cities (including Mexico City); by 2018, Uber users could pay using cash in 400 cities and 60 countries. Most Latin American countries are included in this list, including Brazil and Mexico, the two largest in terms of population.[14]

Uber began operations in Mexico in 2013, beginning with the Greater Mexico City area, which is composed of Mexico City and its adjacent municipalities in the State of Mexico. As of 2018, Uber operated in more than 40 of Mexico's cities. Greater Mexico City is one of the firm's top 10 most-active cities in the world, in terms of rides taken. In August 2018, when our experiments took place, Uber had almost the entire market share in Mexico; Cabify, a Spanish ride-sharing company, had a very low market share and Didi, the Chinese ride-hailing company, was not yet active.

Uber users can select the cash option in the payment tab of the application (*e.g.* Figure 1a). Drivers accept both cash and card payments and do not know the payment method chosen by the rider when a trip is requested. At the end of the trip, the customer hands over the amount shown in the application directly to the driver.[15] Figure 1b shows the share of trips and fares paid in cash in the cities where Uber was available in October of 2017. The figure shows that in the cities in which Uber accepts cash payment, the option is used heavily; almost half of the trips taken are paid for in cash and half of all fares are collected in cash.[16] In the State of Mexico, where we executed the experiments reported below, approximately 25% of users (approximately 30% of fares) only use card payments, 25% of users (20% of fares) only pay in cash, and 50% of users (50% of fares) pay with cash and card. The relevance of riders that use both payment methods actively informed the distinction between mixed users and pure cash users in our model and experiments.

Although Uber is a service mostly consumed by middle- to high-income consumers, the cash option is largely used by low-income consumers. Figure 9a shows the share of cash fares by income per capita at the municipality-level. In the State of Mexico, around 60% of fares in municipalities with low-income per capita are paid with cash (*e.g.* Teoloyucan, Coyotepec), while less than 20% of fares in municipalities with high income per capita are paid with cash (*e.g.* Naucalpan de Juárez, Huixquilucan). Alvarez and Argente (2022) show, using demographic information from the 2010 Mexican Census, that this pattern holds for other variables correlated with income, such as education. A greater share of trips are paid for in cash in municipalities that have less access to banking services, as measured by debit cards per capita, credit cards per capita, bank branches per capita, or ATMs per capita.[17] The share of cash trips is also larger in

14. Uber has been progressively expanding its cash-payment program, recently adding several high-income countries to the list: Germany, Spain, France, Czech Republic, Greece, Poland, Turkey, and Chile.

15. If the user cancels a trip and is charged a cancellation fee, this amount is added to her next trip's fare, which can also be paid using cash.

16. Cash-fare trips are shorter on average. As a result, the share of fares paid in cash is slightly lower than the share of trips paid in cash.

17. In Mexico, the use of debit cards is much more prevalent than that of credit cards. Supplementary Material, Figure I6a shows that, conditional on using cards as the most frequent payment method, more than 80% of households report using debit cards for payments below 20 USD, which covers the majority of Uber rides. Supplementary Material, Figure I6b shows the number of debit cards per capita and the number of credit cards per capita in the State of Mexico. The figure shows that the number of debit cards per capita exceeds the number of credit cards per capita, especially in municipalities where more cards are available, which coincides with those with more active Uber riders.

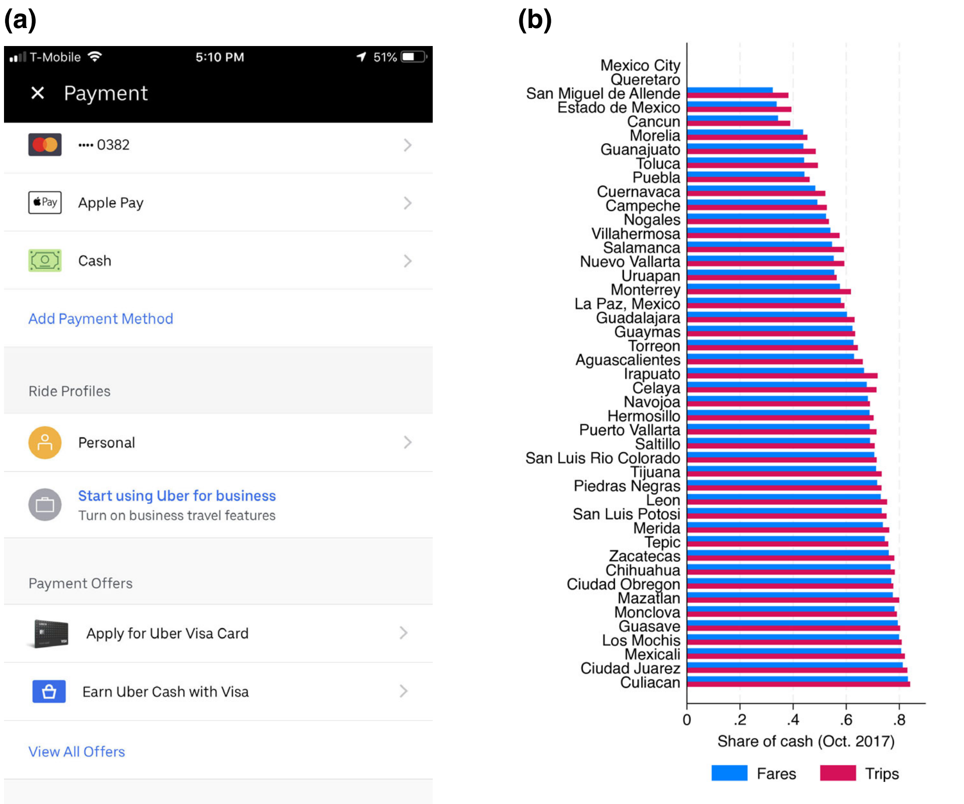


FIGURE 1

Uber Mexico (a) Paying with Uber with cash and (b) share of cash by City

Notes: Panel (a) illustrates how users select the cash option in the payment tab of Uber's mobile application. Panel (b) shows share of trips and fares paid in cash in different cities in Mexico. The lower bars show the fraction of trips and the upper bars the share of fares paid in cash. The sample of cities are those that were active in October 2017.

suburban regions of the State of Mexico and in municipalities with less-developed infrastructure, as measured by the availability of street lights, pavement, or whether the municipality has access to public transport.

Several local governments initially prohibited cash payments for Uber rides in response to complaints from traditional taxi drivers, who considered it to be unfair competition. Cash fares were not allowed within the city limits of Mexico City, whose local government prohibited drivers from receiving any payments in cash, non-banking pre-paid cards, or payment systems hosted by convenience stores through electronic wallets. Queretaro, a mid-size city close to Mexico City, enacted similar policy. In Puebla, ride-hailing fares were limited to electronic payments, but the government did not enforce the policy until a young student was allegedly murdered by a driver working under the auspices of Cabify, another ride-hailing firm. Puebla banned cash payments for ride-hailing services in December of 2017.[18] This decision was also motivated by the taxi-drivers' union's lobbying of the state government, complaining that cash fares for

18. Unfortunately, we do not have access to data of crimes committed by riders and/or drivers and cannot quantify the social benefits of such policy. Alvarez and Argente (2022) do not find evidence that the cash option in Uber affected city-level crime levels. Alvarez *et al*. (2022) measure the social benefits of restricting the use of cash on crime in Mexico. They find that the private costs of heavily taxing the use of cash outweigh the social benefits.

Uber rides represented unfair competition with traditional taxi services.[19] In fact, during the ban on cash, the local government launched its own ride-hailing application "Pro-taxi," which connected mobile-phone users with traditional taxis; cash payments were *allowed* for that app.[20] In November of 2018, the Mexican Supreme Court struck down a state ban on cash fares for ride-hailing firms, setting a national precedent that allows Uber and other ride-hailing firms to accept cash payments. By a vote of 8-3, the court ruled that the small western state of Colima's ban on cash fares was unconstitutional. After the court's decision, Uber began accepting cash payments in Mexico City, Querétaro, and Puebla.

4. RIDER'S MODEL AND CONSUMER SURPLUS

The model is centred on a general utility function for $n+1$ goods where good 1 is "composite Uber trips," goods 2, .., N are close substitutes for Uber, and good $n+1$ represents all other goods with constant marginal utility, such that utility is quasi-linear. Uber rides paid in cash and those paid in card are distinguished as distinct sub-goods that together comprise composite Uber trips. This intensive-margin choice is complemented with the choice of choosing to register a payment card, which we assume is subject to a fixed cost, so that agents have access to Uber trips paid in card only if they pay a fixed cost.

We assume that while a ban on cash payments is in effect, the prices for all goods remain constant. This assumption simplifies the problem and is documented extensively for the case of Mexico and Panama by Alvarez and Argente (2022). They find that the availability of cash payments has a substantial effect on the quantities of rides, but no effects on prices including: prices of Uber rides, average surge multiplier, waiting times for Uber rides, price of taxis, waiting times for taxis, prices of other ride-hailing companies, waiting times for other ride-hailing companies, and time to location of public transport.[21] These findings hold for both the entry and ban of the cash option, and suggest that the supply of drivers is elastic at the relevant time horizon.[22] The lack of effect on prices allows us to ignore the effects of a cash-payment ban on pure card riders and on drivers' producer surplus, since both effects are likely to be small. Moreover, this evidence allows us to apply general results from demand theory to estimate the consumer surplus of cash Uber fares without needing to measure the quantities of other goods. The model, therefore, focuses on the choice faced by consumers who potentially encounter different prices for Uber rides depending on the payment method, holding prices for other goods constant.

We consider the welfare cost for riders in the case of a ban on cash as means of payment for Uber rides. Before the ban on cash payments, riders face the same price for Uber rides paid in cash and Uber rides paid in card. Facing equal prices, heterogeneous riders then face the choice of whether to register a payment card or not. We then estimate the change in a rider's welfare, in dollars, if the price cash-fare for Uber rides increases to infinity (*i.e.* a ban on cash payments). This welfare loss equals the area under the demand curve for cash-fare Uber rides. This measure takes both the intensive and extensive margins into account.

19. According to the ENIF, approximately 90% of respondents, who own a debit or a credit card, report cash as their most frequently used payment method for transportation services (*e.g.* taxi, bus).

20. Pro-taxi was launched almost 10 months after the ban on cash in the city of Puebla. It became inactive 3 months later for lack of financial resources.

21. Supplementary Material, Section J shows that cancellation rates did not change significantly either.

22. Drivers' income per hour was unchanged. For an in-depth analysis of Uber drivers' labour supply with varying compensation schemes, see Angrist *et al*. (2021).

4.1. *Intensive-margin choice*

We assume that a rider's utility function is given by $u(x_1, x_2, \ldots, x_n; \phi) + x_{n+1}$, where x_1 are composite Uber rides and the goods or services $x_2, x_3, \ldots, x_n$ are close substitutes for and/or complements to Uber (*e.g.* taxis). The good x_{n+1} represents the rest of the goods and services. Preferences are quasi-linear, with the marginal utility of income normalized to one. We assume that $u(\cdot; \phi)$ is strictly concave and increasing in its n arguments. We let ϕ index the preferences of different riders, and let K be the distribution of ϕ across riders. We use ϕ to refer to types defined by variables that we can observe.

Assuming a quasi-linear utility function subject to idiosyncratic shocks at the rider level is reasonable, given the small share of each consumer's overall expenditures allocated to Uber rides (Vives 1987). These preferences offer two key advantages. First, they significantly simplify the analysis, as equivalent and compensated variations coincide. Second, they aggregate to a quasi-linear utility for a group of ex-ante identical riders with the same observable characteristics. Consequently, we can test all the restrictions implied by our experimental data on that aggregate utility function, with the null hypothesis being that the experimental data was generated by some quasi-linear utility function at the aggregate level. In Supplementary Material, Section A.2, we apply the test proposed by Allen and Rehbeck (2018) and confirm that all restrictions hold for the two price experiments used to quantify Uber riders' consumer surplus.[23]

Composite Uber rides, x_1, follow a constant returns-to-scale function, such as constant elasticity of substitution (CES), represented as $x_1 = H(a, c; \phi)$, where a denotes Uber rides paid in cash, and c Uber rides paid with a card. This framework, in line with Lucas (1987), provides a tractable framework for a welfare analysis of restrictions on cash usage. The function H captures consumer preferences between paying with cash or card.

It is convenient to have a specific notation for the price of Uber rides paid in cash, for which we use p_a, and Uber rides paid with card, for which we use p_c. We let $p_2, \ldots, p_n$ denote the prices of the rest of the goods. Thus, the intensive-margin problem for the rider is

$$\upsilon(p_a, p_c, p_2, \ldots, p_n; \phi) = \max_{a,c,x_2,\ldots,x_{n+1}} u\left(H\left(a, c; \phi\right)\right), x_2, \ldots, x_n) + x_{n+1} \tag{1}$$

$$\text{subject to} \quad p_a a + p_c c + \sum_{i=2}^{n} p_i x_i + x_{n+1} = I,$$

where we assume that the total income of rider I is large enough so that consumption of good $n+1$ is always positive. We have normalized $p_{n+1} = 1$, so that we can interpret the numeraire as dollars (or Mexican pesos). The indirect utility function v is the focus of our theory, since we will use it to estimate consumer surplus. We omit the prices $\{p_2, \ldots, p_n\}$ from most expressions since we keep them fixed in our applications given the evidence discussed above.

Our weakly separable specification allows us to isolate the choice between means of payment from the overall demand for Uber rides. Given the assumption that H is homogeneous of degree one, a rider's choice to pay for an Uber trip with cash depends only the rider's type ϕ and the ratio of cash and card prices p_a/p_c, but it does not depend on the rider's income I or any feature of the utility function u. On the other hand, if prices are equal for riders that have access to both means of payment, $p_a = p_c = P$, the demand for composite Uber rides depends only on the common price P and on the utility function u; demand is independent of the function H. We can

23. To the best of our knowledge, our work is the first to apply this type of statistical test to experimental data with price variation across individuals.

use H to define the ideal price for one composite Uber ride

$$\mathbb{P}(p_a, p_c; \phi) = \min_{a,c} p_a a + p_c c \text{ subject to } H(a, c; \phi) = 1 \tag{2}$$

We normalize the units of $H(\cdot; \phi)$ so that $H(p, p; \phi) = p$ for any $p > 0$.[24] We assume that H is such that $\mathbb{P}(\infty, 1; \phi)$ and $\mathbb{P}(1, \infty; \phi)$ are both finite (*i.e.* allows for finite choke prices for pure cash users and pure card users). For instance, since H is given by a CES function, we require the elasticity of substitution to be greater than one.

4.2. *Extensive-margin choice*

We assume a rider can pay with a card only after incurring a fixed cost $\psi \geq 0$.[25] We use the vector $\theta = (\psi, \phi)$ to fully specify the rider's type. The complete problem for the rider is

$$\mathcal{V}(p_a, p_c; \theta) \equiv \max \{v(p_a, p_c; \phi) - \psi, v(p_a, \infty; \phi)\}. \tag{3}$$

The first option is to pay the fixed cost ψ and face ride prices (p_a, p_c). The rider can also save the fixed cost ψ, but will then only have access to the cash price; we represent this limited access by setting the card price to infinity: $p_c = \infty$.

Let $\tilde{a}$ and $\tilde{c}$ be the demand functions for Uber rides paid in cash and cards, respectively. Also, let $1_c(p_a, p_c; \theta) \in 0, 1$ be an indicator that equals one if the optimal decision in equation (3) is to register a payment card with Uber and zero otherwise.[26] We can now define the rider's demands for cash or card Uber rides for any type of rider $\theta = (\psi, \phi)$, taking the intensive and extensive margins into account

$$\left(a^*(p_a, p_c; \theta), c^*(p_a, p_c; \theta)\right) = \begin{cases} (\tilde{a}(p_a, p_c; \phi), \tilde{c}(p_a, p_c; \phi)) & \text{if } 1_c(p_a, p_c; \theta) = 1 \\ (\tilde{a}(p_a, \infty; \phi), 0) & \text{if } 1_c(p_a, p_c; \theta) = 0. \end{cases}$$

We use the cumulative distribution functions G and K to describe the distribution of fixed costs conditional on ϕ and the distribution of ϕ, respectively. We let $\psi \sim G(\cdot \mid \phi)$ and $\phi \sim K(\cdot)$ describe the cross-sectional distribution of $\theta = (\psi, \phi)$. We assume that the distribution of ψ conditional ϕ has continuous density $g(\psi \mid \phi) = G'(\psi \mid \phi)$ for all (ψ, ϕ). We use F for the implied distribution of types θ.

4.3. *Welfare costs and consumer surplus*

Given our assumption of quasi-linearity, we can aggregate the riders' welfare level and measure it in units of the numeraire. We normalize the units that quantify trips so that the price of a trip is 1 when both means of payment are available, *i.e.* we normalize the length of each ride so that the cash and card prices are $p_a = p_c = 1$. We denote the consumer surplus lost in the ban of

24. We let $a(p_a, p_c)$ and $c(p_a, p_c)$ be the choices that attain the minimum in equation (2) so that $\mathbb{P}(p_a, p_c) = p_a a(p_a, p_c) + p_c c(p_a, p_c)$. The functions a and c are homogeneous of degree zero in (p_a, p_c) while $\mathbb{P}$ is homogeneous of degree one in (p_a, p_c). The ideal price index is given by $\mathbb{P}(p_a, p_c)$, and is increasing in the convex function of (p_a, p_c).

25. We express the fixed cost with its equivalent-flow value. This notation converts the fixed cost to units comparable with $v(p_a, p_c; \phi)$. Later on, we introduce a discount rate ρ which converts the flows into stocks. The discount rate ρ incorporates pure-time discounting and the expected duration for the registration of the payment card and/or the expected duration of the Uber service.

26. See Supplementary Material, Section A.5 for a full characterization of the demand functions.

cash by CS_{ban}, which we define as follows. We assume that riders have access to both cash and a payment card before the ban and that they have already made their optimal choice regarding registering a card by solving the problem in equation (1). The prior decision about registering a card is summarized by $1_c(1, 1; \theta)$ and the distribution of types F. The consumer surplus cost of the ban is, therefore:

$$\mathcal{CS}_{\text{ban}} = \int 1_c\,(1, 1; \theta)\left[\underbrace{\upsilon(1, 1; \phi)}_{\text{mixed}} - \underbrace{\upsilon(\infty, 1; \phi)}_{\text{pure card}}\right] \mathrm{d}F(\theta)$$
$$+ \int [1 - 1_c\,(1, 1; \theta)]\left[\underbrace{\upsilon(1, \infty; \phi)}_{\text{pure cash}} - \underbrace{\mathcal{V}(\infty, 1; \theta)}_{\text{pure card vs no Uber}}\right] \mathrm{d}F(\theta). \tag{4}$$

The first term counts the riders that registered a card before the ban, denoted by the indicator $1_c(1, 1; \theta)$. These riders are either pure card users or mixed users. Before the ban, their net utility flow is $\upsilon(1, 1; \phi)$. These users have paid the fixed cost to register a card, which is a sunk cost. After the ban, these riders face a much higher cash price for Uber ride, *i.e.* their utility flow value is $\upsilon(\infty, 1; \phi)$. The second term counts the riders that were pure cash users before the ban. Their utility-function flow before the ban is $\upsilon(1, \infty; \phi)$. After the ban, these riders must choose between paying the fixed cost and becoming pure card users, which gives the utility flow of $\upsilon(1, \infty; \phi) - \psi$, or ceasing to use Uber, which corresponds to the net utility flow $\upsilon(\infty, \infty; \phi)$. This last choice is accounted for with the term $\mathcal{V}(\infty, 1; \theta)$.[27]

Following standard arguments from demand theory, the consumer surplus lost after a cash-payment ban can be computed as the area below the aggregate demand for cash-fare Uber rides. First, we define the aggregate demand in a city that initially allows cash payments, where the cash price unexpectedly increases to $p_a \geq 1$

$$A(p_a, 1) = \int 1_c\,(1, 1; \theta)\,\tilde{a}(p_a, 1; \phi)\,\mathrm{d}F(\theta) + \int (1 - 1_c\,(1, 1; \theta))\,a^*(p_a, 1; \theta)\,\mathrm{d}F(\theta) \tag{5}$$

Note that this definition of aggregate demand breaks the integral into two groups of riders, as in equation (4). The first group has already registered a card, according to the decision at the original prices $(p_a, p_c) = (1, 1)$, for which $1_c(1, 1; \theta) = 1$. The second are the remaining riders, which have not registered a card and, hence, they may consider to do it optimally.

Proposition 1. Assume that $G(\cdot\,|\,\phi)$ has continuous density, and that almost all riders θ have sufficiently large income I to consume the outside good. Then

$$\mathcal{CS}_{\text{ban}} = \int_1^\infty A(p_a, 1)\,\mathrm{d}p_a \tag{6}$$

The demand satisfying equation (6) is the *aggregate* demand. A proof of Proposition 1 is provided in Supplementary Material, Section A.1, relying on the envelope theorem for the intensive-margin and assuming a density g for the fixed cost to account for the extensive-margin of adoption.

27. More generally, we can define for any $p_a \geq 1$ the consumer surplus cost that follows an increase in the price of cash from 1 to $p_a \geq 1$ as $\mathcal{CS}(p_a, 1)$. The ban is represented as $\lim \mathcal{CS}(p_a) = \mathcal{CS}_{\text{ban}}$ as $p_a \to \infty$.

Other means of transportation. Proposition 1 states that the consumer surplus depends only on the prices/quantities of Uber rides. We assume that the prices of taxis and other substitutes (complements), remain constant after a ban on cash payments, an assumption supported by ample empirical evidence. These findings allow us to evaluate the consumer surplus for cash-fare Uber rides *without* measuring the impact on the quantities of other goods. In fact, the availability of this information does not impact or improve the consumer surplus calculations. This is a general result of demand theory.[28] In Supplementary Material, Section D, we use demand theory to show that when the prices of substitutes (or complements) are fixed, cross-price elasticities and quantities demanded of these goods do not affect the calculation of consumer surplus. Supplementary Material, Section K develops a closed-form example to illustrate these results from demand theory, and Supplementary Material, Section L summarizes the empirical evidence on the price response of other means of transportation following a ban or the introduction of cash. Nevertheless, if the prices of other modes of transportation were increase in response to a ban on Uber rides, our consumer surplus estimates, which turn out to be large, would be a lower bound.

4.4. *Identification and functional forms*

In theory, based on Proposition 1, we could trace out the demand curve for Uber rides by increasing the cash price permanently. Repeating this exercise until the cash price reaches the choke price, we could directly estimate the consumer surplus of cash-fare Uber trips. In practice, however, Uber's policies render this exercise impossible. We overcome this challenge with large-scale field experiments in which we trace out the demand curve by reducing, rather than increasing prices, as well as bringing to bear information from the reaction of riders to the ban in Puebla. In concert with the structural model described above, we extrapolate from this data to estimate the consumer surplus. We use a parametric version of the model because our experiments contain a limited amount of price points and rewards variation.

We first divide an Uber rider's consumption problem into two stages to clarify the features of the indirect utility function that are identified by each experiment. As a preliminary step, we define the utility function $U(\cdot; \phi, p_2, \ldots, p_n) : \mathbb{R}_+ \to \mathbb{R}$ to embed all the information of the utility function u in a simple set up, for fixed prices of the related goods $\{p_2, \ldots, p_n\}$:

$$U(X; \phi, p_2, \ldots, p_n) \equiv \max_{x_2, x_3, \ldots, x_n} u\,(X, x_2, \ldots, x_n; \phi) + I - \left[\sum_{i=2}^{n} p_i x_i\right]. \tag{7}$$

This problem establishes a utility function for composite Uber rides, in which X serves as the main argument by maximizing out the remaining related goods 2 to n, at prices $\{p_2, \ldots, p_n\}$.[29] Using U, we can define the following indirect utility function $V(\cdot; \phi) : \mathbb{R} \to \mathbb{R}$ in the problem for a rider choosing the number of composite rides X at price P:

$$V(P; \phi) = \max_{x \geq 0} \; U(x; \phi) + \left[I' - Px\right] \tag{8}$$

Note that we are using that preferences are quasi-linear. We let the optimal solution be $X(P)$, with the first-order condition $U'(X(P)) = P$ if $X(P) > 0$ and $U'(X(P)) \leq P$ otherwise. These results will aid the below discussion of the assumptions needed to compute $\mathcal{CS}_{\text{ban}}$.

28. See Hausman (1981), who shows that one can recover an expenditure function whose derivative provides the appropriate compensated demand curve for the good whose price has changed, which enables the exact calculation of compensating variation and equivalent variation, which coincide under quasi-linearity.

29. As before, we omit the dependence of prices $\{p_2, \ldots, p_n\}$.

Cash-card choice utility *H*. For a given rider type ϕ, given that H is homogeneous of degree one, we can identify H if we observe the ratio of the choices $\tilde{a}(p_a, p_c; \phi)/\tilde{c}(p_a, p_c; \phi)$ as we vary p_a/p_c exogenously. Equivalently, we can identify H by tracing the share of trips paid in cash $p_a\tilde{a}/(p_a\tilde{a} + p_c\tilde{c})$ as a function of p_a/p_c (see Experiment 1 below). For $H(\cdot; \phi)$, we use a CES function described by two parameters

$$H(a, c) = \left[\alpha^{\frac{1}{\eta}} c^{\frac{\eta-1}{\eta}} + (1-\alpha)^{\frac{1}{\eta}} a^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}},$$

where η is the elasticity of substitution and the observable rider-specific parameter α represents the share of payments made with a card for mixed users. To be precise, if $p_a = p_c = p$ for any p, the optimal demand gives $p_c c/(p_c c + p_a a) = \alpha$ and $p_a a/(p_c c + p_a a) = 1 - \alpha$. The parameters (α, η) are contained in the type ϕ.[30]

Uber ride utility *U*. The definition of U in equation (7) and equation (8) make clear that U is identified by observing how $\tilde{c}(p, p; \phi)$ and $\tilde{a}(p, p; \phi)$ change as the price of an Uber ride with either payment method $p = p_a = p_c$ changes, since $p = \mathbb{P}(p, p; \phi)$. Moreover, for pure cash riders, we can also identify U by varying the price of trips paid in cash p_a, which gives $\mathbb{P}(p_a, \infty; \phi) = p_a \mathbb{P}(1, \infty; \phi)$ (see Experiments 1 and 2 below).[31] Importantly, we use the functional form of U, and its associated demand for rides X, to extrapolate the shape of the indirect utility function V estimated using experimental variation in prices. We let

$$U(x; \phi) = -k \exp\left(-\left(x + \bar{x}\right)/k\right)$$

such that U is described by $k >$ and $\bar{x} > 0$. The demand that solves equation (8) is

$$X(P; \phi) = -k \log P + k \log \bar{P} \tag{9}$$

so k and $\bar{P}$ are indexed by ϕ. This demand has constant semi-elasticity $k \geq 0$. Note that the price elasticity of this demand function is:

$$\epsilon(P) \equiv -\frac{P}{X(P)} \frac{\partial X(P)}{\partial P} = \frac{1}{\log\left(\bar{P}/P\right)}. \tag{10}$$

The consumer surplus of a rider with this utility function at initial price P_0 is[32]

$$C(P_0; \phi) = \int_{P_0}^{\bar{P}} X(p; \phi)\, \mathrm{d}p \quad \text{and} \quad \frac{C(P_0; \phi)}{P_0 X(P_0; \phi)} = \epsilon(P_0)\left[\exp\left(\frac{1}{\epsilon(P_0)}\right) - 1\right] - 1. \tag{11}$$

This semi-log demand curve has a finite choke price $\bar{P}$ (*i.e.* $X(\bar{P}; \phi) = 0$) given by $\bar{P} = e^{-\bar{x}/k}$ and is convex. These two features are *consistent* with our experimental and survey data. The ratio of the choke price to the current price for the demand function with constant semi-elasticity is $\bar{P}/P = \exp(1/\epsilon(P))$. For instance, at $\epsilon = 1.3$, the choke price is about 2.1 times greater than

30. As usual, the price of a composite Uber ride is $\mathbb{P}(p_a, p_c; \phi) = [\alpha p_c^{1-\eta} + (1-\alpha) p_a^{1-\eta}]^{1/(1-\eta)}$.

31. If we decrease p_a, we can also disregard pure cash riders' incentives towards registering a card. Also, if the constant $\mathbb{P}(1, \infty; \phi)$ is unknown, then we can identify U up to a constant; see Case 4 of Supplementary Material, Section A.5.

32. P_0 refers to the initial price at which the consumer surplus is to be evaluated, *e.g.* the elasticity of demand is not constant, and its value depends on the point at which it is evaluated. Since we normalize the length of each ride to make cash and card prices equal to 1, we use $P_0 = 1$.

the price at which we evaluate the elasticity.[33] The convexity of the demand curve implies a larger consumer surplus (relative to expenditures) compared with a linear demand curve, as the latter lacks local convexity and has lower choke prices. Considering a demand curve with constant elasticity would also not be reasonable due to the magnitude of choke prices, leading to a substantially larger consumer surplus.[34]

Distribution of fixed cost g. We assume that the indirect utility functions $v(p, \infty; \phi)$ and $v(p, p; \phi)$ are known and that pure cash riders are faced with different levels flow rewards d that can be obtained only if they register a card (see Experiment 3 below). Then, we can identify the distribution $\psi \sim g(\cdot \mid \phi)$ using the fraction of riders that have registered a card for different values of d.[35] The distributions of ψ and ϕ must also be consistent with the behaviour of pure cash users. Here, we list the relevant constraints:

(1) *The choice of pure cash users not to start using a payment card as long as cash payments are allowed.* The condition that ensures this is

$$\psi \geq v(1, 1; \phi) - v(1, \infty; \phi) \tag{12}$$

for all cash users and all values of ψ in the support of $G(\cdot \mid \phi)$. The right-hand side of this equation defines the lower bound of the support $G(\cdot \mid \phi)$, which we refer to as $\underline{\psi}(\phi)$.

(2) *The observed excess migration of pure cash users to pure card users after the ban in Puebla.* For the second condition, we use that fraction m_{ban} of pure cash users in Puebla migrated to card after the ban on cash, in excess to those that migrated before the ban

$$\psi \leq v(\infty, 1; \phi) - v(\infty, \infty; \phi) \text{ for fraction } m_{\text{ban}} \text{ and} \tag{13}$$

$$\psi \geq v(\infty, 1; \phi) - v(\infty, \infty; \phi) \text{ for fraction } 1 - m_{\text{ban}}. \tag{14}$$

The right-hand side of these inequalities defines a value of ψ such that for higher values pure cash riders prefer to stop using Uber. We refer to this value as $\psi_{\text{ban}}(\phi)$.

(3) *The change in trips for pure cash users that became pure card users after the ban in Puebla.* In Puebla, we keep tract of the number of trips for pure cash users that become pure card users after the ban. In the data, these users decreased their number of trips. Thus, for those values of ϕ, we must have

$$0 < \tilde{a}(\infty, 1; \phi) \leq \tilde{a}(1, \infty; \phi). \tag{15}$$

(4) *The experimental evidence on the excess migration for different reward levels.* In our experiment (Experiment 3 below), pure cash riders are offered a one time payment d_j, from which we measure the induced (excess) migration of fraction m_j of pure cash riders to become card/mixed riders by registering a card. We index each level incentives as well

33. In Supplementary Material, Sections A.5 and A.6, we derive expressions for the different demand curves for cash and card fares: $a(p_a, p_c; \phi)$, $\tilde{a}(p_a, p_c; \phi)$, $c(p_a, p_c; \phi)$, $\tilde{c}(p_a, p_c; \phi)$, the indirect utility function $v(p_a, p_c; \phi)$, and other comparisons between indirect utility functions used in the computation of the consumer surplus.

34. Specifically, the consumer surplus relative to expenditures with linear demand is $\frac{1}{2}\frac{1}{\epsilon(P_0)}$. Under a demand function with constant elasticity, the consumer surplus relative to expenditures is $\frac{1}{\epsilon - 1}$. This value can grow very large for elasticities closer to 1, as those we estimate below.

35. We assume that the density g of the distribution of fixed costs for registering a card ψ is the same for all pure cash users. This assumption is verified in Supplementary Material, Section G. Supplementary Material, Figure G4 shows that migration rates are independent of historical trips, a variable capturing heterogeneity among cash users.

as each fraction of the treatment group that migrate by j.

$$\psi \leq \upsilon(1,1;\phi) - \upsilon(1,\infty;\phi) + \rho d_j \text{ for fraction } m_j \text{ and} \tag{16}$$
$$\psi \geq \upsilon(1,1;\phi) - \upsilon(1,\infty;\phi) + \rho d_j \text{ for fraction } 1 - m_j \tag{17}$$

Although we do not know α for pure cash riders, given that they have not been faced with interior choices for card prices, there is a small interval of α's consistent with all these inequalities. In Supplementary Material, Section A.7, we find the remaining parameters of U and G and compute the consumer surplus lost in a ban on cash by pure cash users for each feasible value of α. Below, we aim to be conservative and report estimates using the value of α that is consistent with a *lower bound* of the net consumer surplus lost by pure cash users who switch to card payments after the ban. We also report estimates consistent with an *upper bound* of the consumer surplus lost by pure cash users assuming riders do not switch to card payments after a ban on cash. Our estimates show that both lower and upper bounds are very close.

Consumer surplus calculation. All the relevant functional forms for our analysis and expressions used for the consumer surplus calculation are listed in Table 1. The calculation of the consumer surplus is straightforward; here, we recap the steps. For pure cash users, we integrate the demand for Uber rides up to the choke price. After normalizing by total fares, the calculation only requires plugging in the empirical estimate $\epsilon(P_0)$ into equation (11) at the initial price $P_0 = 1$. For pure cash users who switch to cards after the ban, the consumer surplus has to be adjusted by subtracting the fixed cost of registering a card, $\psi \sim g(\cdot \mid \phi)$, estimated using the experimental and observational evidence described below. Lastly, for mixed users, the consumer surplus has to be adjusted for the share of payments made with a card α and the elasticity of substitution η. In this case, the calculation for the consumer surplus, normalized by total fares, requires plugging in α, which is observed, as well as our estimates for $\epsilon(P_0)$ and η into Supplementary Material, equation (29): $\epsilon(P_0)[-\frac{1}{\epsilon(P_0)} - 1 - \alpha^{\frac{1}{1-\eta}}(\log(\alpha^{\frac{1}{1-\eta}}) - \frac{1}{\epsilon(P_0)} - 1)]$, representing the difference between the indirect utility of mixed users and pure card users relative to total fares of mixed users.

Discrete choice model. Supplementary Material, Section B describes a discrete choice model of Uber ridership and the choice of means of payment, which gives rise to the same demand system described above.[36] In each subperiod of the discrete choice model, agents choose whether to take an Uber trip or to use an alternative mode of transportation, and, conditional on choosing to ride with Uber, they decide the means of payment they will use (*i.e.* cash or card). As standard, in each subperiod agents draw a set of random variables that indicate the values of these choices. Using particular distributional assumptions of these random variables, we obtain our functional forms. In particular, the discrete choice model has the property that each indirect utility function, before the realization of the shock, is the same as those described in Sections 4.1 and 4.2. Furthermore, the discrete choice model has a random demand whose expected values are the same as those demand functions described in Section 4.4. In Supplementary Material, Section C, we also discuss how to estimate the discrete choice model using sequential GMM.

5. EXPERIMENTS

This section describes the three field experiments that let us identify the parameters in our model and estimate the consumer surplus lost after cash payment is banned. The experiments took place

36. Our discrete choice model builds on the work by Anderson *et al.* (1987) and Dubé *et al.* (2022).

TABLE 1
Functional forms and main equations

	Model	Identification
Preferences	$u(H(a,c;\phi)), x_2, \ldots, x_n; \phi) + x_{n+1}$	Test using experimental data (Supplementary Material, Section A.2)
Extensive	$\mathcal{V}(p_a, p_c; \theta) \equiv \max\{v(p_a, p_c; \phi) - \psi, v(p_a, \infty; \phi)\}$	ψ: Experiment 3 (Section 5.3) and Puebla
Utility	$U(x;\phi) = -k\exp(-(x+\bar{x})/k)$	$\epsilon(P_0)$: Experiments 1 and 2 (Section 5), Panama (Section 5.2), Survey (Section 5.2)
Intensive	$H(a,c) = \left[\alpha^{\frac{1}{\eta}} c^{\frac{\eta-1}{\eta}} + (1-\alpha)^{\frac{1}{\eta}} a^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}}$	η: Experiment 2 (Section 5.2), α: Data
	Consumer surplus	**Derivation**
Pure cash	$\epsilon(P_0)\left[\exp\left(\frac{1}{\epsilon(P_0)}\right) - 1\right] - 1$	Indirect utility comparisons (Supplementary Material, Section A.4) Details in equation (11)
Pure cash (adj.)	$\int_{\widehat{\underline{\psi}}}^{\max\{\widehat{\underline{\psi}}, \widehat{\psi}_{\text{ban}}\}} \left[\psi - \widehat{\underline{\psi}}\right] \widehat{g}(\psi) d\psi$	Net consumer surplus lost (Supplementary Material, Section A.7) Details in Supplementary Material, equation (39)
Mixed	$\epsilon(P_0)\left[-\frac{1}{\epsilon(P_0)} - 1 - \alpha^{\frac{1}{1-\eta}}\left(\log\left(\alpha^{\frac{1}{1-\eta}}\right) - \frac{1}{\epsilon(P_0)} - 1\right)\right]$	Indirect utility comparisons (Supplementary Material, Section A.4) Details in Supplementary Material, equation (29)

Notes: The table shows the functional forms assumed in the model. It also shows the experiments that identify each of the parameters required to estimate the consumer surplus and the key equations used for the calculation of the consumer surplus.

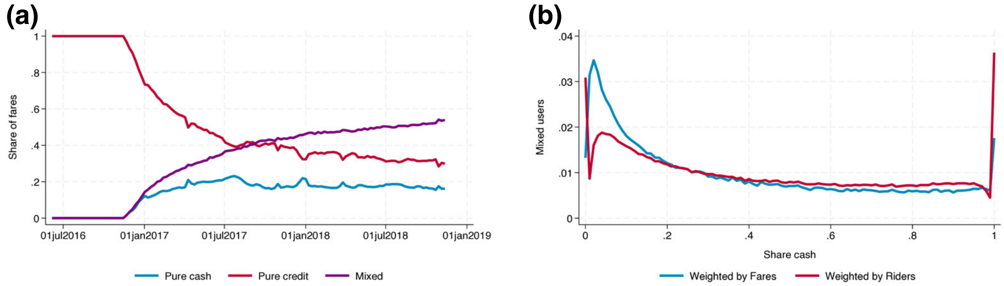


FIGURE 2

State of Mexico: share of fares by type of user (a) share of fares by user type and (b) distribution mixed users

Notes: Panel (a) shows the share of total fares paid by different types of users in the State of Mexico. The red line shows the share of fares paid by pure card users, those that have never paid for an Uber ride with cash. The blue line shows the share of fares for pure cash users, those who have not registered a card in the application. The purple line shows the share of fares of mixed users, those who have at least one trip paid in cash and at least one paid using a card. The cash payment option was introduced in November 2016. Panel (b) shows the distribution of mixed users as a function of the share of fares paid in cash. The sample includes users with at least 4 weeks of tenure that had used both methods of payment and that took at least five trips after becoming mixed users. The blue line shows the distribution of mixed users weighted by fares and the red line shows the distribution weighted by riders.

in the State of Mexico between August and September of 2018. The sample includes users who signed up in the State of Mexico and whose most frequent city for Uber trips is within the State of Mexico. All users have a verified mobile number and are not subject to other experiments simultaneously. The users in our sample took at least two trips in 2018 and took at least one trip since 1st April 2018.

In Experiment 1, we vary the prices of cash and/or card for mixed users to estimate the elasticity of substitution between cash and card payments η, as well as the price elasticity of demand $\epsilon(P)$. In Experiment 2, we vary the price p_a for pure cash users to estimate the price elasticity of demand $\epsilon(P)$. Lastly, in Experiment 3, we present pure cash users with different incentives towards registering a payment card to estimate the distribution of the fixed cost g.

Supplementary Material, Table E1 shows descriptive statistics for the users in our sample, including averages of variables such as fares, trips, fares paid in cash, trips paid in cash, share of fares paid in cash, and tenure. Mixed users pay higher fares per week than pure cash users (\$5.29 USD versus \$1.43 USD) and travel more (1.11 trips per week versus 0.36).

5.1. *Experiment 1: consumer surplus - mixed users*

This experiment focuses on mixed users, who have paid for at least one trip using both cash and card in their rider history before the experiment began.[37,38] Figure 2a shows the share of fares paid by mixed users over time in the State of Mexico. Mixed users account for approximately half of the fares paid in the State of Mexico. Panel (b) shows the distribution of mixed users over their share of fares paid in cash.

We have six treatment groups, each composed of approximately 11,000 riders and a control group of 90,000 riders. The treatment and control groups were balanced in the following observables: average weekly historical trips, average weekly historical fares, log tenure (in weeks), and average weekly historical fares paid in cash. Riders in the treatment groups were presented with

37. Users in the sample must have a card on file that is not banned by Uber.
38. Examples of the emails sent communicating the promotions can be found in Supplementary Material, Section Q.

the following promotional offers: (i) 10% off if the trip is paid with cash, (ii) 10% off if the trip is paid with card, (iii) 10% off regardless of the payment method, (iv) 20% off if the trip is paid in cash, (v) 20% off if the trip is paid with card, and (vi) 20% off regardless of the payment method. Note that this design includes both decreases and *increases* in relative prices. The discounts were applied to all trips taken by the riders in each treatment group during the entire week. At the beginning of the week, riders received an introductory email describing the promotion. At the same time, the promotion appeared on their phone's main screen once they opened the application. Two reminder emails were sent (in the middle of the week and 2 days before the promotion expired).[39]

We concentrate on the share of card payments, $s_c \equiv p_c c/(p_c c + p_a a)$, among mixed riders with positive trips during the week of the experiment in treatments facing different relative prices p_a/p_c. For a given α, this is a function of relative prices faced by riders. We choose s_c, as opposed to the ratio $\frac{c}{a}$, because it is well-defined even if a rider makes only one trip during the week of the experiment. Indeed, in a discrete choice model, where taking a trip or not is a probabilistic event, the comparison of the average value of s_c between the treatment and control gives the equation to estimate η. In Supplementary Material, Section B, we derive such a discrete-time model, and in Supplementary Material, Section C, we derive the moment conditions to estimate η.

To simplify the presentation and use regression analysis, we linearize the optimal choice of the share of card payments s_c for a CES function H, as a function of the relative prices p_a/p_c, the share parameter α, and the elasticity of substitution η.[40] The first- and second-order approximations around $p_c/p_a = 1$ are

$$s_c = \alpha - (\eta - 1)\alpha(1-\alpha)\ln\left(\frac{p_c}{p_a}\right) \quad \text{and} \tag{18}$$

$$s_c = \alpha - (\eta - 1)\alpha(1-\alpha)\ln\left(\frac{p_c}{p_a}\right) + \frac{1}{2}(1-\eta)^2(1-\alpha)\alpha\,[1-2\alpha]\left(\ln\left(\frac{p_c}{p_a}\right)\right)^2. \tag{19}$$

In our empirical implementation, we begin by using the first-order approximation. We observe the share of payments made with a card, $\hat{\alpha}^i$, for $p_c = p_a$ in our data (*i.e.* equation (18) becomes linear for $p_a = p_c$). However, $\hat{\alpha}^i$ could be measure with measurement error since it is estimated using riders' historical data, which depends on the number of trips riders have taken. We mitigate this concern by dividing each side of equation (18) by our estimate of $\alpha(1-\alpha)$ so that $\tilde{s}_c \equiv \frac{s_c}{\alpha(1-\alpha)}$

$$\tilde{s}_c = 1/(1-\alpha) - (\eta - 1)\log(p_c/p_a). \tag{20}$$

As a result, we run the following regression at the rider i level for data generated during the week of the experiment:

$$\tilde{s}_c^i = \varphi_0 + \varphi_1 \log(p_c^i/p_a^i) + \nu^i, \tag{21}$$

where φ_0 is a constant, our estimate of η is obtained from φ_1 (*i.e.* $\hat{\eta} = 1 - \hat{\varphi}_1$). The error term ν^i contains potential sampling errors; through the lens of our discrete choice model, it represents

39. In Supplementary Material, Section N.1, we show that the first-order approximation is highly accurate, and the second-order approximation is nearly exact. Our normalization on the length of each ride implies that the cash and card prices faced by the control group are $p_a = p_c = 1$.

40. Details can be found in Supplementary Material, equation (48) in Section C, which shows the moment equation for the share of card payments used in the GMM estimation.

the differences between the model-derived expected values and the realized values in the experiment, as in GMM.[41] Prices are rider-specific given that riders are assigned to either treatment or control groups. This regression has the advantage of moving the measurement error on α to the left-hand side variable, thereby mitigating the attenuation bias that such measurement error may cause. We refer to this specification as the transformed-share case. Table 2 shows our estimates of η, the elasticity of substitution between Uber rides paid in cash and Uber rides paid in card.[42] In these specifications, columns (1)–(3), we find an elasticity of substitution of approximately 3. In column (4), we include the constant $1/(1-\alpha)$ specified in equation (20) as a regressor using each mixed rider's historical trips to estimate $\hat{\alpha}^i$ and find similar results. In column (5), we go back to equation (18) and use our estimates for $\hat{\alpha}^i$ to construct $\Gamma^i \equiv \hat{\alpha}^i(1-\hat{\alpha}^i)\ln(p_c^i/p_a^i)$. Then, we estimate

$$s_c^i = \varphi_0 + \varphi_1 \Gamma^i + \nu^i.$$

In column (6), we follow a similar procedure using the second-order approximation to construct the second-order term $(\Gamma^i)^2 \equiv \frac{1}{2}(1-\hat{\alpha}^i)\hat{\alpha}^i[1-2\hat{\alpha}^i](\ln(p_c^i/p_a^i))^2$ and estimate

$$s_c^i = \varphi_0 + \varphi_1 \Gamma^i + \varphi_2 \left(\Gamma^i\right)^2 + \nu^i.$$

In column (7), we instrument α, to reduce potential bias introduced by measurement error. First, we compute the predicted share of fares paid in card using all the control variables. Then, we estimate equation (18), as in column (5), using the predicted share. Our preferred estimates are in columns (5) and (7). While the point estimates vary across the different specifications displayed in Table 2, we find $\eta \approx 3$ or smaller throughout.[43]

For robustness, we tested specifications with and without controls (historical fares and tenure in Uber), specifications that split price increases and price decreases, and specifications that use different thresholds to define the set of mixed users (those with more than 5% and less than 95% of their fares paid in cash, etc.). These robustness checks can be found in Supplementary Material, Section N.2. We find that the estimates for η are similar for *price increases* and price decreases (Supplementary Material, Table N32). In Supplementary Material, Table N33, we further confirm that similar estimates are obtained using the second-order term of the CES function's second-order approximation, φ_2 (*i.e.* $\hat{\eta} = 1-(\hat{\varphi}_2)^{\frac{1}{2}}$). Importantly, we find that the estimates for η are independent of observables such as share of rides paid with cash, total fares, total fares in cash, and riders' tenure (Supplementary Material, Figure G1), all of which are highly correlated with riders' income. These additional results provide additional portability to our estimates for this parameter and our structural specification.

An alternative estimate for the elasticity of substitution can be obtained by aggregating the decision about the share of card-payment trips across riders. For this purpose, we write the second-order approximation for this choice s_c as a function of the prices faced by a single rider and as a function of her share parameter α and of the common elasticity of substitution η.[44]

41. Our baseline specification includes mixed users with more than 1% of their fares paid in cash and less than 99% of their fares paid in cash. We present robustness checks with mixed users with more than 5% of their fares paid in cash and less than 95% of their fares paid in cash.

42. Throughout, we estimate the standard errors using the Delta Method. Our results are unchanged if instead we use robust standard errors.

43. In Supplementary Material, Section N.1, we show that for the range of the parameter of interest the first-order approximation is very accurate and the second-order approximation is nearly exact.

44. Alvarez and Argente (2022) propose a mechanism for the simultaneous use of cash and cards and for the imperfect substitutability across payment methods. They show that mixed users are more likely to pay with cash whenever they have it available; in particular, they are more likely to pay with cash after they get paid.

TABLE 2
Elasticity of substitution: mixed users (miles)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Elasticity	3.180***	2.914***	2.655***	2.997***	2.619***	2.618***	2.223***
	(0.364)	(0.340)	(0.177)	(0.212)	(0.102)	(0.102)	(0.074)
Obs.	53,966	53,966	46,328	53,966	53,966	53,966	78,265
Controls	No	Yes	Yes	Yes	Yes	Yes	No
Type	1 pct	1 pct	5 pct	1 pct	1 pct	1 pct	1 pct
Specification	Transf.	Transf.	Transf.	Transf. Cons	CES First	CES Second	CES First IV

Notes: The table reports estimates of the elasticity of substitution between cash and card payments for mixed users. The estimates are computed using experimental data collected in the State of Mexico. The dependent variable is the relative miles between card and cash payments for each user the week of the experiment and the independent variable is the relative price for cash and card trips. Column (1) reports the results after using the transformed-share specification in equation (20) and including mixed users with more than 1% of their fares paid in cash and less than 99% of their fares paid in cash. Column (2) reports the same specification including controls. The controls included for each user are the historical trips, trips squared, fares, fares squared, cash fares, cash fares squared, log of tenure, cash trips, and cash trips squared. Column (3) includes users with more than 5% of their fares paid in cash and less than 95% of their fares paid in cash. Column (4) includes the constant specified in equation (20) as a regressor. Column (5) estimates the elasticity using the CES first-order approximation in equation (18). Column (6) estimates the elasticity using the CES second-order approximation in equation (19). Column (7) reports the results of the elasticity of substitution estimated in two steps. First, we compute the predicted share of fares paid in card (*i.e.* $\hat{\alpha}$) using all the control variables. Then, we estimate equation (18) using the predicted share. The standard errors are computed using the Delta Method. The ***, **, and *, represent statistical significance at 1%, 5%, and 10% levels, respectively.

We interpret equation (19) as the expected value of the share of card trips. We let μ be the distribution of α across the experiment's population. Riders enter into this population if they satisfy the conditions to be in the experiment—such as being active mixed riders—and they do so with weights proportional to the probability of taking a trip within a week. Control and treatment groups differ only in the randomly allocated prices p_c/p_a, so the expected value of $\bar{s}_c(p_c/p_a)$ is given by

$$\bar{s}_c\left(\frac{p_c}{p_a}\right) = m_1 - (\eta - 1)m_2 \ln\left(\frac{p_c}{p_a}\right) + m_3\,(1-\eta)^2\left(\ln\left(\frac{p_c}{p_a}\right)\right)^2$$
$$m_1 = \int \alpha\,\mu(\mathrm{d}\alpha),\, m_2 = \int \alpha(1-\alpha)\mu(d\alpha),\ \text{and}\ m_3 = \frac{1}{2}\int (1-\alpha)\alpha\,[1-2\alpha]\,\mu(\mathrm{d}\alpha). \tag{22}$$

We estimate μ by using the distribution of the share of card payments prior to the experiment for 54,470 riders with positive trips during the experiment. The estimated values for the three moments are $\hat{m}_1 = 0.6187$, $\hat{m}_2 = 0.1349$, and $\hat{m}_3 = -0.0081$, with very small standard errors.

In Figure 3, we plot the actual average share across riders for each of the four treatment groups (10% and 20% cash discount and 10% and 20% card discounts) and for the control group, including its 95% confidence interval. We also plot three versions of the theoretical prediction equation (22), using the estimated moments $(\hat{m}_1, \hat{m}_2, \hat{m}_3)$. Each line corresponds to a different value of the elasticity of substitution, namely $\eta = 2.5$, $\eta = 3$, and $\eta = 3.5$, a range of values suggested by the regressions on Table 2. We note that given the small value of $\hat{m}_3$ the relationship between $\bar{s}$ and $\log(p_c/p_a)$ is almost linear, *i.e.* the first-order approximation for the expected share is very accurate. Second, the dots, which correspond to the average card share for control and treatment groups for each price, are arranged in a nearly linear segment. Third, the value of

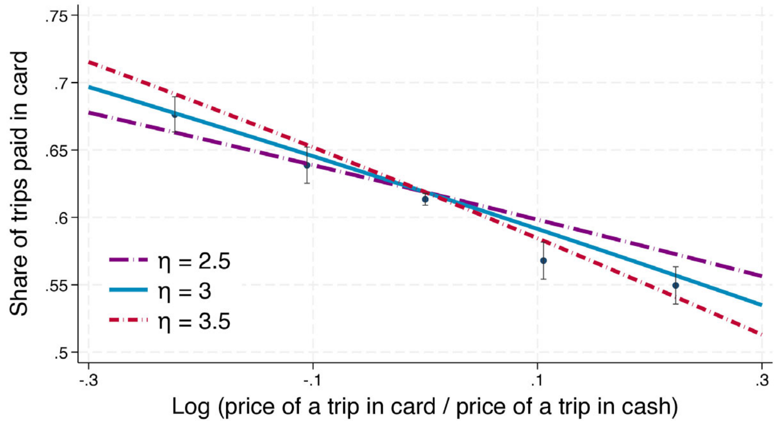


FIGURE 3

Experiment I and elasticity of substitution η

Notes: The dots are the average card share for control and treatment groups with the corresponding relative price. The vertical lines are 95% standard error bands. The solid and dotted lines are the theoretical prediction for the expected card share displayed in equation (22) using the estimated values of $\hat{m}_1$, $\hat{m}_2$, and $\hat{m}_3$. The lines differ in the value of the parameter η.

$\eta = 3$ gives a very good fit, providing further validation to our previous estimates.[45] Lastly, in Supplementary Material, Section C, we estimate the discrete choice version of our model and provide an alternative estimate for the elasticity of substitution, η. We again obtain estimates around 3 or smaller using sequential GMM following Hansen (2007).

We also estimate the composite Uber price elasticity ϵ for mixed users under our functional assumption of constant semi-elasticity, using the treatments where the cash and card prices $P = p_a = p_c$ are the same. These estimates are essentially regressions of the miles during the week of the experiment on the log of the price and a constant, following equation (9).[46] We find that the elasticity ϵ, evaluated at current prices, is approximately 1.1 or smaller, which corresponds to the first two columns of Table 3, labelled AA.

We also include the results of two other experiments conducted independently by Uber, labelled as Mandin and Ubernomics. We use these experiments to provide external validity to our estimates of the elasticity of demand for cash and mixed users. These experiments were not explicitly designed to give estimates of the elasticity and curvature of the demand function, but their results allow estimates of these parameters. We are able to select riders and construct control variables to make the samples comparable using historical data. These confirmatory exercises return elasticities similar those found in our experiments.[47] The Mandin experiment

45. The normalization on the length of each ride implies that using miles as dependent variable is equivalent to using total fares.

46. The Ubernomics experiment took place in the Greater Mexico City from 15th May to 22nd May 2017, only a few months after the introduction of cash in the State of Mexico. The treatment groups received discounts of 10% and 20% off in all rides taken the week of the experiment. The sample includes 4,869 pure cash users and 4,306 mixed users. The Mandin experiment took place in all areas of the Greater Mexico City in June 2018 and lasted 4 weeks. The treatment groups received discounts of 10%, 20%, and 30% off for all rides taken during the weeks of the experiment. The sample includes 5,668 pure cash users and 20,914 mixed users. More details on both experiments can be found in Supplementary Material, Section R.

47. The average of the ratio of consumer surplus to total Uber expenditures, using $\eta = 3$, $\epsilon = 1.1$, and the distribution of the α, weighted by fares, is 0.2463. This figure describes mixed riders who have taken more than five trips and have more than 4 weeks of tenure.

TABLE 3
Elasticity of demand: mixed users (miles)

	(1) AA	(2) AA	(3) AA	(4) Mandin	(5) Ubernomics
Elasticity	1.082***	1.030***	1.096***	1.278***	1.452***
	(0.103)	(0.086)	(0.093)	(0.075)	(0.296)
Observations	109,365	109,365	98,773	11,660	4,306
Controls	No	Yes	Yes	Yes	Yes
Type	1 pct	1 pct	5 pct	1 pct	1 pct

Notes: The table reports the elasticity of demand for mixed users estimated using Supplementary Material, equation (27) using miles as the dependent variable. Column (1) reports the estimates without using controls. Column (2) estimates the elasticity using controls. The controls included for each user are the historical trips, trips squared, fares, fares squared, cash fares, cash fares squared, log of tenure, share of fares paid in cash, cash trips, and cash trips squared. Column (3) reports the results using the users included in the Mandin experiment. Column (4) reports the results using the users included in the Ubernomics experiment. The standard errors are computed using the Delta Method. The ***, **, and *, represent statistical significance at 1%, 5%, and 10% levels.

is particularly relevant, since it had price variation that lasted 4 weeks. Thus, Uber riders were less likely to miss the promotion (avg. Uber rider in the State of Mexico takes 2.6 trips per month) and the experimental variation better approximates a permanent change in prices. Supplementary Material, Section N.2 contains several other robustness exercises that complement our estimated elasticities, including estimates of the semi-elasticity of demand, the elasticity of demand of number trips, the elasticity of demand for users that have taken at least five trips, the elasticity of demand in logs, the Poisson regression specification, and the Poisson pseudo maximum-likelihood specification.

We next use the estimated values of η and ϵ to calculate the consumer surplus enjoyed by mixed users. Using these elasticities, the observed distribution of the share of cash trips, and the observed distribution of total fares, we implement Supplementary Material, equation (29). We aggregate riders weighting them by their fares paid in cash (*i.e.* Figure 2) so that the aggregate consumer surplus is the total surplus of the cash option over total expenditures. Figure 4 displays the consumer surplus as the share of a user's expenditure on Uber, where the horizontal axis shows the share of cash fares. Each line corresponds to different parameter values for ϵ and η. We estimate the consumer surplus lost after a ban on cash payments to be 25% what mixed users spend on Uber rides.[48] Recall that mixed users account for about 50% all expenditures on Uber rides in the State of Mexico. Since the average mixed user pays for 37% of rides with cash, their consumer surplus decrease by 67% of their expenditures on trips paid in cash.[49] For $\eta = 5$, the consumer surplus lost for mixed users is 42.6% of cash fares.[50] Panel (b) shows that for $\eta \geq 5$, the consumer surplus lost is very similar for different values of ϵ.

48. To be precise, using the cash share for mixed users of 0.3685, we get $0.6682 = 0.2463/0.3685$.

49. An alternative research design would have been to only use the changes in cash prices to estimate the consumer surplus lost for mixed users (*e.g.* using the same four discounts as we use in Experiment 2). This alternative design has the advantage of yielding a more-direct measure of the curvature of mixed users' demand for cash trips and does not require an estimate of η. Supplementary Material, Figure E1 in Section E.1 shows that our implied functional form captures the shape of the mixed user's demand for cash trips in the data, which validates both the model and our parameter estimates. We opt for modelling payment choices for mixed users instead of following this alternative approach because it allows us to *increase relative prices* and estimate the elasticity of substitution η, which we believe is a parameter of interest that can be used for several counterfactuals.

50. Examples can be found in Supplementary Material, Section Q.

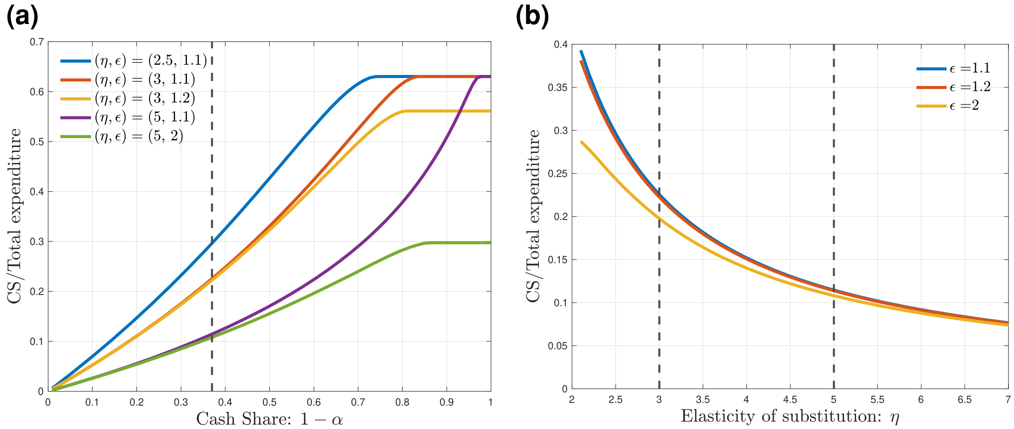


FIGURE 4
Consumer surplus: mixed users

Notes: Panel (a) shows the model estimates of the consumer surplus (as a multiple of initial total fares) as a function of the share of a user's cash trips. The graph plots the estimates for different combinations of the elasticity of demand ϵ and the elasticity of substitution between cash and card payments η. The consumer surplus estimates are for mixed users, those who have paid for at least one trip using a payment card and for at least one trip in cash. Panel (b) shows the model estimates of the consumer surplus (as a multiple of initial total fares) as a function of the elasticity of substitution, η, for three different values of the elasticity of demand, ϵ. The expression used for the consumer surplus calculations is: $\epsilon(P_0)\left[-1/\epsilon(P_0)-1-\alpha^{\frac{1}{1-\eta}}\left(\log\left(\alpha^{\frac{1}{1-\eta}}\right)-1/\epsilon(P_0)-1\right)\right]$.

5.2. *Experiment 2: consumer surplus - pure cash users*

This experiment was targeted to pure cash users in order to understand their card adoption patterns. We focus on users that have not registered a card with Uber. We have four treatment groups each composed of approximately 20,000 riders and a control group of 56,000 riders. The treatment and control groups were balanced in the following observables: average of weekly historical trips, average of weekly historical fares, and log tenure (in weeks). We have four treatment groups each getting 10%, 15%, 20%, and 25% off of all the trips taken during the week of the experiment. Note that, since the treatments cover several price points, the experiment is designed to provide information about the local convexity of the demand curve, which we use to inform our structural assumptions on the constant semi-elasticity demand function. At the beginning of the week the riders received an introductory email describing the promotion. At the same time, the promotion showed up in the main screen of their phone once they opened the application. Two reminder emails were sent (in the middle of the week and 2 days before the promotions expired).[51]

Using the miles travelled during the week of the experiment as the dependent variable, we estimate the price elasticity of demand, ϵ, to be approximately 1.38, when evaluated at current prices. Our baseline case is the semi-log demand corresponding to our functional form specification. Table 4 displays the estimates under columns AA, as well as estimates using the same

51. Other specifications and further robustness exercises can be found in Supplementary Material, Section N.2 including estimates of the semi-elasticity of demand, the elasticity of demand of number trips, the elasticity of demand for users that have taken at least five trips, the elasticity of demand in logs, the Poisson regression specification, and the Poisson pseudo maximum-likelihood specification.

TABLE 4
Elasticity of demand: pure cash users (miles)

	(1) AA	(2) AA	(3) Mandin	(4) Ubernomics
Elasticity	1.375***	1.383***	1.113***	0.813**
	(0.101)	(0.078)	(0.165)	(0.414)
Observations	138,725	138,725	4,279	3,569
Controls	No	Yes	Yes	Yes

Notes: The table reports the elasticity of demand of pure cash users estimated from Supplementary Material, equation (27) using miles as dependent variable. Column (1) reports the estimates without using controls. Column (2) estimates the elasticity using controls. The controls included for each users are the historical trips, trips squared, fares, fares squared, and log of tenure. Column (3) reports the results using the users included in the Mandin experiment. Column (4) reports the results using the users included in the Ubernomics experiment. The standard errors are computed using the Delta Method. The ***, **, and *, represent statistical significance at 1%, 5%, and 10% levels, respectively.

specification for the Ubernomics and Mandin experiments. This estimate is robust to using controls such as the average of weekly historical trips, average of weekly historical trips squared, average of weekly historical fares, and log tenure (in weeks).[52]

Using the estimated price elasticity and our demand specification we next estimate the consumer surplus for pure cash users. Figure 5 displays our estimates for different elasticity estimates. Using 1.38, we estimate the consumer surplus to be approximately 47% of the total fares per year. The consumer surplus lost displayed in Figure 5 is, however, an *upper bound* estimate given that, some users might decide to begin using a card rather than leaving Uber completely after a large price increase. In fact, when the cash option was banned in Puebla, only 70% of the users left the platform. To adjust the consumer surplus of these riders, we use both the experience in Puebla, as well as a third experiment to estimate the fixed cost of adopting a card payments. Section 5.3 provides more details.

The right axis of the figure displays the corresponding choke price implied by our functional form as a multiple of the current price. The choke prices corresponding to our preferred estimate for price elasticity are about double the current prices. We choose a demand function with constant semi-elasticity because it is consistent with the local convexity we found in the experimental data and because it implies a finite choke price. This implication is relevant because consumer surplus estimates are sensitive to the shape of the demand curve at very high prices, which are rarely explored in field experiments. For example, Table 5 reports estimates for pure cash users for different demand specifications using our estimated elasticity of demand. The table shows that our baseline specification predicts a consumer surplus 30% higher than the linear specification but at least 5.5 times smaller than the log-log specification. To go further in confirming our results, in the next subsection, we use a natural experiment in the country of Panama and a survey instrument to validate the functional-form assumptions for the demand curve at very high prices.

5.2.1. Panama: large price increase. We use data from a natural experiment in Panama, where the government abruptly limited the supply of drivers, to validate our functional-form assumptions and obtain an additional estimate of demand elasticity. This data is informative

52. Uber initially operated in three provinces: Panama City, Panama West, and Colon. A recent Supreme Court decision has allowed Uber to extend its services to more provinces. Panama City is the most active province in terms of rides.

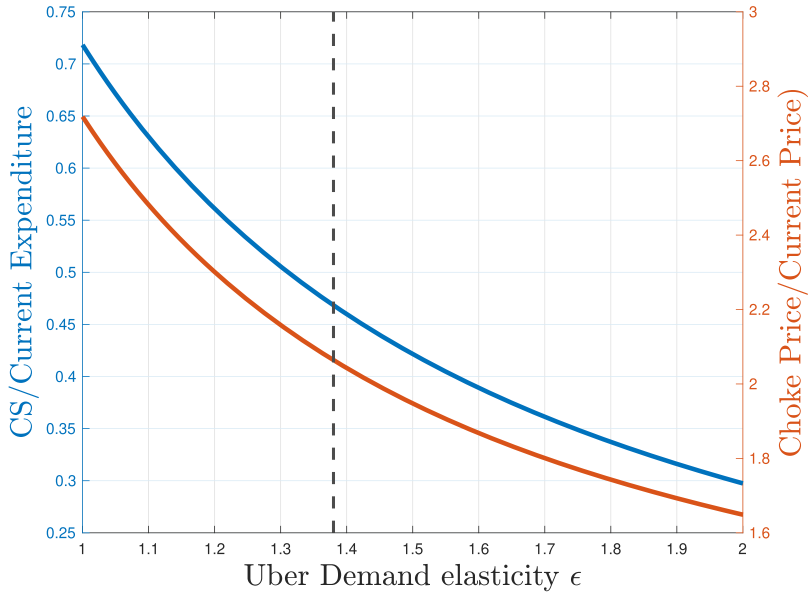


FIGURE 5
Consumer surplus and choke price: cash users

Notes: The figure shows the model estimates of the consumer surplus (as a multiple of initial total fares) as a function of the elasticity of demand ϵ. The graphs also shows the model estimates of the choke point, the price at which the demand for Uber trips is zero as a function of ϵ. The estimates are for pure cash users, those that never registered a card in the application. The expression used for the consumer surp.us calculations is: $\epsilon(P_0)[\exp(1/\epsilon(P_0))-1]-1$.

TABLE 5
Consumer surplus cash users: functional forms

	CS relative to expenditures	$\epsilon=1.2$	$\epsilon=1.38$ (baseline)	$\epsilon=2$
Linear	$\frac{1}{2}\frac{1}{\epsilon(P_0)}$	0.42	0.36	0.25
Semi-log (baseline)	$\epsilon(P_0)\left[\exp\left(\frac{1}{\epsilon(P_0)}\right)-1\right]-1$	0.56	0.47	0.30
Log-log (constant elasticity)	$\frac{1}{\epsilon-1}$	5	2.6	1

Notes: The table displays estimates of the consumer surplus relative to expenditures for different functional forms at initial price P_0=1. The table shows the equations used for the calculation and the estimates for $\epsilon=1.2$, $\epsilon=1.38$, and $\epsilon=2$. Our baseline estimates are under the semi-log specification and $\epsilon=1.38$.

about the demand curve at higher prices, as Uber ride costs nearly doubled following the regulation.

Uber launched in Panama in February 2014.[53] In August 2016, due to the low penetration of credit card use, cash payments were introduced nationwide. Within a year, over half of all trips were paid with cash. Panama's government imposed restrictions on Uber in October 2017.[54] The decree included a ban on cash payments for Uber rides, mandated a special license ("E1" type)

53. Restrictions were implemented after a nationwide ground transportation strike organized by the two Panamanian transportation unions.

54. The decree also introduced a fleet cap of two cars and geographical limitations, allowing Uber to operate in only four out of Panama's ten provinces.

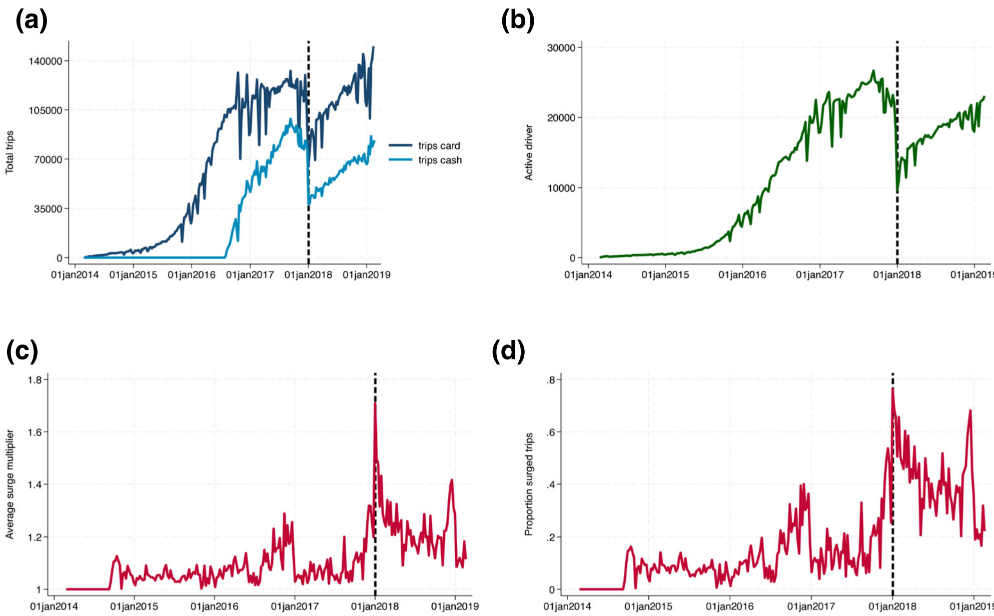


FIGURE 6

Panama: trips, fares, and drivers (a) trips. (b) Active drivers. (c) Avg. surge multiplier and (d) share surged trips

Notes: The figure shows the evolution of trips, active drivers, the average surge multiplier and the share of surged-price trips in Panama. The frequency of the data is weekly. The black dotted line marks the date that the restrictions went into effect.

for drivers, costing around \$200 USD and obtainable only by nationals over 21, following a 36-hour seminar.[55] Driver restrictions took effect on 2 January 2018.[56] As a result, 83% of Uber drivers were disconnected from the application due to lacking the E1 license. The unexpected reduction in driver supply led to a surge in surge-priced trips, increasing from an average of 16% in 2017 to 45% in 2018. As shown in Figure 6, the share of cash-fare trips notably decreased from over 50% in 2017 to less than 35% in 2018. This decline in cash payments exceeded that of card payments, consistent with our findings that the demand for Uber trips paid in cash is more elastic than for card payments.

Interpreting this natural experiment as an exogenous reduction in driver supply, we use data on total trips and average surge multipliers to trace the Uber demand function in Panama. In Figure 7, we show the number of trips plotted against prices for each of the 52 weeks in 2018 following the driver supply restriction. The blue line shows the fit of the semi-log demand function implied by our chosen functional forms. We estimate the elasticity of demand to be approximately 0.95 for all trips in Panama City. For cash-fare rides, the demand elasticity increases to about 1. The share of cash-fare rides was approximately 0.4 before the driver restriction but decreased afterward, consistent with the higher elasticity of cash trips. These trends are consistent with our data from the State of Mexico.[57] The graph shows that even at high prices not

55. Uber negotiated an extension for the cash ban deadline until May 2019, later renewed until October 2019. Cash payments were temporarily banned in February 2020.

56. Details on these estimates are provided in Supplementary Material, Section O.

57. The surveys were sent through email to all riders in Experiments 1 and 2 on 9th July 2019 and were open to responses until 16th July 2019. A total of 433,356 users received a survey, 287,233 participated in Experiment 1 (mixed and pure card users) and 146,123 participated in Experiment 2 (pure cash users). The response rate was 1.46%.

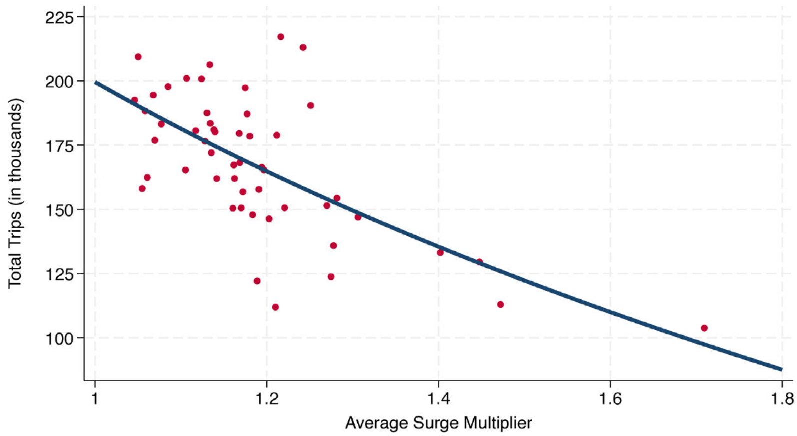


FIGURE 7
Panama: total trips and prices (2018)

Notes: The figure plots the total weekly trips and the average weekly surge multiplier for Panama. Each dot represents a week in 2018, the weeks after the decree went into effect reducing the supply of drivers in the country. The surge multiplier is seasonally adjusted. The line is a semi-log function.

explored in our experiments, the demand curve aligns well with observed patterns of total trips and prices. Our assumption of exponential utility accurately fits the observed patterns in Panama when prices nearly double.

5.2.2. Survey instrument: choke Prices. We also used a survey instrument to gain further insight into how price changes might affect consumer preferences and what the choke price might be. The survey was sent to the riders in our Experiments 11 months after the experiments concluded. Six different surveys were randomly given to users, with three questions each. We received more than 6,000 responses, an average of 1,056 responses per survey.[58] This format allowed us to minimize response times and, at the same time, allowed us to obtain a range of responses for a given question. For example, all surveys included the following question: “If you were to receive a 20% discount for 1 week, how would you change the number of trips you take....” Some users were given the options to respond (a) no change, (b) increase less than 10%, (c) increase more than 10%. A second set of users were given options to respond (a) no change, (b) increase less than 20%, (c) increase more than 20%. And a third set of users were given options to respond (a) no change, (b) increase less than 30%, (c) increase more than 30%.

Each survey also included two other symmetric questions, one related to a *permanent* large price decrease (*e.g.* “If the price of trips is permanently reduced by half, how would you change your trips...”) and another related to a permanent large price increase (*e.g.* “If the price of trips is permanently doubled, how would you change your trips...”). Half of the surveys sent asked users to respond to permanently doubling prices or permanently reducing prices by half, while the other half of surveys asked users to respond to prices permanently tripled or cut by a third. In response to the question about a permanent price increase, the users could respond with the following options: (i) no change, (ii) decrease substantially, (iii) stop travelling. We compare elasticities from the survey with those from our experimental design to validate the survey instrument and ensure the reported elasticities reflect revealed preference elasticities.

58. To minimize measurement error in the historical average of weekly fares, we trim top and bottom 1%.

TABLE 6
Distribution of choke prices

Choke Price	(1) Mean	(2) Std. Dev.	(3) 10th	(4) 25th	(5) Median	(6) 75th	(7) 90th
Mixed users	6.0	20.7	1.18	1.35	1.82	3.28	8.19
Pure cash users	4.4	10.9	1.47	1.62	1.99	3.06	6.42

Notes: The table shows moments of the distribution of choke prices implied by framework described in Section 4 for both mixed users and pure cash users. To approximate $X(P)$, we use each user's historical average of weekly fares. To minimize the measurement error, we trim the top and bottom 1%. The semi-elasticity k is that estimated for each group of users presented in Supplementary Material, Tables N1 and N9.

The last survey question on permanent price increases provides information on the distribution of choke prices, which we compare to the distribution implied by our structural framework.

To analyse users' responses, we follow three steps. First, we reweight the survey respondents' covariate distribution to match the entire experiment population based on their trip history and tenure. We accomplish this step with entropy balancing, a multivariate reweighting method described in Hainmueller (2012). Second, we use responses to the first question to validate the survey instrument, confirming that reported elasticities align with those from our experiments. Lastly, we use responses to the third question to compare reported choke prices with our model. The rest of this section focuses on this final step, with more details on the previous steps in Supplementary Material, Section F.

In our structural framework, for mixed users in the control group (facing prices equal to 1 in our model), the implied choke price is defined as

$$\bar{P} = \exp\left(\frac{X(P)}{-k}\right), \tag{23}$$

where $X(P)$ is the number of miles a rider travels in a week and k is the semi-elasticity we estimated from experimental data. Since the survey responses provide us with a distribution of choke prices, we implement equation (23) using the data to obtain the distribution of choke prices implied by our structural assumptions. This requires taking a stance on the riders' heterogeneity. In this case, we use each user's history of average weekly fares to approximate $X(P)$ and the semi-elasticity estimated in our experiments.[59] Table 6 presents the distribution of choke prices for mixed users.

The median choke price for mixed users implied by our model is 1.82. There is considerable heterogeneity in the choke prices; the ratio between the 75th and the 25th percentiles is 2.42. Given our structural assumptions, if we doubled prices, 56% of users would leave the platform and, if we tripled prices, approximately 73% of users would stop using Uber. These figures are remarkably close to the survey responses. Approximately, 56% of respondents said they would stop travelling if prices doubled and 67% responded that they would stop travelling if prices were tripled. Next, we study pure cash users. Their choke price is defined as

$$\bar{P} = \exp\left(\frac{\tilde{a}(p_a, \infty)}{k(1-\alpha)^{\frac{1}{1-\eta}}} + \log\left((1-\alpha)^{\frac{1}{1-\eta}}\right)\right), \tag{24}$$

59. To address concerns about potential contamination of the estimated elasticities and/or migration rates by the advertisement effect of receiving promotional emails from Uber, Supplementary Material, Tables N34, N26, and N37 present results using only the variation across treatment arms. These results closely align with those reported in our baseline specifications.

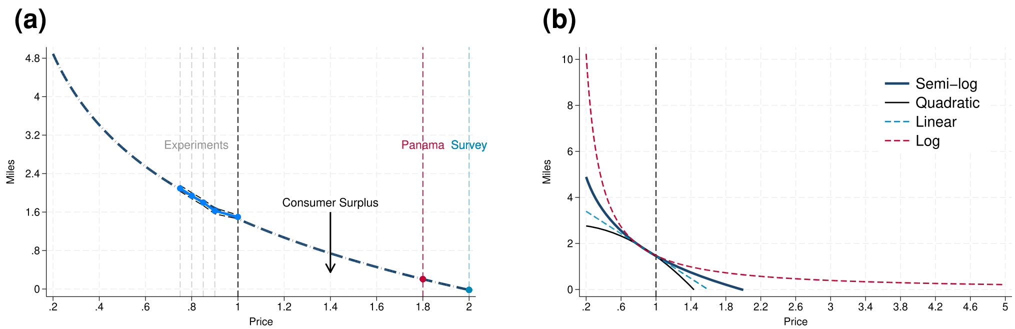


FIGURE 8
Demand curve: cash payments Uber rides

Notes: Panel (a) displays the semi-log function. The dashed lines represent the estimated points obtained from experimental, observational, or survey data as well as the equilibrium prices. The demand curve derived from experiments is estimated using the miles travelled during the week of the experiment as the dependent variable. To calculate each point, we first demean the following observables: trips, trips squared, fares, fares squared, and tenure. We then estimate a regression of miles on all the treatment arms in logs, without including a constant. Panel (b) displays other demand functions calibrated using the same procedure. It includes a semi-log, quadratic, linear, and log-log function.

where $k(1-\alpha)^{\frac{1}{1-\eta}}$ is the semi-elasticity of demand for pure cash users, as estimated directly from our regressions. The median choke price in this case is 1.99 and the ratio between the 75th and the 25th percentiles is 1.88. Our model implies that if we doubled prices, 51% of users would leave the platform, aligning with 54% in survey responses. If prices tripled, our framework implies that 75% would leave the platform, closely matching the 69% from surveys. Given that self-reports are informative about revealed-preference elasticities, these findings about the choke prices provide additional validation to our structural assumptions.

Figure 8a shows the demand curve for Uber rides paid in cash. The dashed lines indicate points of the demand curve that are estimated either by experiments, observational data (Panama), or survey data. The figure shows we cover almost all the relevant range of variation in prices, from below equilibrium prices to the choke price. The chosen semi-log demand form is consistent with the observed local convexity between miles and prices, as estimated from experimental evidence, and with a finite choke price, as observed in survey data. Figure 8b shows different demand curves for different functional forms. The figure illustrates that the location of the choke price significantly affects consumer surplus. In Supplementary Material, Section S.3, we compare the reported choke prices to those implied by linear, semi-log, and log-log demand functions. We show that under a linear demand specification, too many users stop using Uber relative to the estimates from the survey. Under a semi-log demand function, as we discussed before, the estimates on the fraction of users who would stop riding Uber after a price increase are remarkably close to the survey evidence. The CES functional form (or log-log after taking logarithms) can be rejected by the survey data since this functional form does not have a finite choke price. In other words, even for very high prices, riders continue to use Uber. The survey data shows that there is a sizable share of riders who would stop using Uber if prices were to double or triple, contradicting this demand specification.

5.3. *Experiment 3: net consumer surplus - pure cash users*

The estimates of the consumer surplus for pure cash users reported in Section 5.2 under our structural assumptions still need to be adjusted for the fact that, in the event of a ban on cash, riders could decide to pay the fixed cost of adopting a card and return to the application. Experiment 3 is designed to estimate the fixed cost of adopting a payment-card. The experiment is

TABLE 7
Extensive-margin: adoption of a payment-card

	(1) 1 week	(2) 1 week	(3) 1–6 week	(4) 1–3 week	(5) 4–6 week
Treatment 1–1 week	0.0241***				
	(0.004)				
Treatment 2–1 week	0.0269***				
	(0.004)				
Treatment 3–1 week	0.0366***				
	(0.004)				
Treatment 1–6 week		0.0166***	0.0333***	0.0283***	0.0112***
		(0.004)	(0.004)	(0.004)	(0.003)
Treatment 2–6 week		0.0217***	0.0394***	0.0382***	0.0088***
		(0.004)	(0.004)	(0.004)	(0.003)
Treatment 3–6 week		0.0390***	0.0468***	0.0485***	0.0088***
		(0.004)	(0.004)	(0.004)	(0.003)
Constant	0.0272***	0.0272***	0.0711***	0.0445***	0.0372***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Observations	20,609	20,677	46,996	36,184	46,996
R^2	0.005	0.005	0.005	0.006	0.001

Notes: The table reports the percent of users that adopted a payment-card for each of the treatment groups in experiment three relative to the control group. Migration is an indicator function that equals one if the user registered a card conditional on taking trip the weeks of the experiment. The variables "Treatment" report the migration rates relative to the control group of the three treatment groups in Experiment: 3, 6, and 9 times their average weekly fares if the users register a card in the application. Column (1) reports the rates of card adoption for the experiment that lasted 1 week. Column (2) reports the rates of card adoption during the first week for the experiment that lasted 6 weeks. Column (3) reports the rates of card adoption for the experiment that lasted 6 weeks. Column (4) reports the rates of card adoption during the first 3 weeks of the experiment. Column (5) reports the rates of adoption in the last 3 weeks of the experiment. The ***, **, and *, represent statistical significance at 1%, 5%, and 10% levels.

targeted to pure cash users in order to understand their card adoption patterns. We focus on users that have not registered a card with Uber.

We offered rewards if a user registered a card in the application, without imposing restrictions on the payment method used for subsequent rides. This was the first time Uber Mexico implemented an experiment of these characteristics. The treatment groups received rewards of 100, 200, or 300 pesos (5.2, 10.5, and 15.7 USD), which are approximately 3, 6, and 9 times the average weekly fares (or approximately 1, 2, and 3 average rides). The experiment is designed to obtain information about different points in the distribution of fixed costs. Given that pure cash users might or might not have a card already, the experiment had two treatments for each reward with two different time horizons. The first lasted only 1 week, targeting users that might already have a card, but have not registered it in the application. The second lasted 6 weeks in order to allow enough time for users to obtain a new card. These users received email reminders about the promotion every week. Overall, the experiment included six treatment groups with three incentive levels lasting 1 and 6 weeks, each made up of approximately 20,000 riders and a control group of 40,000 riders.

Table 7 shows the percent of pure cash users that adopted a payment-card (registered a card in the application) in each of the treatment groups conditional on having taken a trip during the weeks of the experiment. Column (1) and (2) show the adoption rates during the first week, for the 1- and 6-week experiments. The columns show that similar amounts of users register a card in the first week, regardless of the time horizon. In both cases, users in the treatment groups responded significantly to the incentives provided, relative to the control group. We observe more migration to card payments when larger incentives are offered. For instance, for a reward of slightly more than 15.2 USD the migration rate increases by 3.9%, which is statistically significantly larger than the migration rate with a reward of 5.2 USD, which is 1.6% –see column (2) of the table.[60]

Column (3) shows the overall migration rate over the span of 6 weeks and Column (4) and (5) examine the migration rates during weeks 1–3 and weeks 4–6, respectively. The columns show that significantly more users migrate during the first 3 weeks of the experiment than do in the latter 3 weeks. This indicates that, although our incentives were sufficiently enticing to encourage migration of marginal users, they were not enough to substantially incentivize users that did not own a card. Importantly, Supplementary Material, Table N35 shows that users in our treatment groups were more likely to use cards more than 6 months after our experiment ended. The table shows that, conditional on travelling 6–8 months after our experiments and having taken a trip during the weeks of our experiments, the probability of paying with a card is larger for users in our treatment groups.[61]

We next use a variety of observations to estimate the consumer surplus lost in a ban, taking into account the effect of those pure cash riders that choose to pay the fixed cost and become pure card users after the ban. To do so, we combine theoretical aspects with experimental evidence. On the theoretical side we use the specifications of preferences (described in Section 4.4), with their implications for demand (derived in Supplementary Material, Section A.5), the corresponding indirect utility functions derived in Supplementary Material, Section A.6, and the conditions that fixed cost and indirect utility have to satisfy for the optimal adoption of cards, as described in equations (13) and (16). On the experimental side, we use the parameters estimated in Experiment 2 for the demand of trips for pure cash users, the elasticity of substitution between cash and card payments estimated in Experiment 1 for mixed users, the migration rates under each of the incentive levels described in Section 5.3 from Experiment 3. With this information, we jointly estimate the counterfactual share parameter α for pure cash users, the parameters for the utility function U for composite rides for pure cash users (k and $\bar{P}$), and the distribution of the fixed cost G. In our choice of α, we strive to be conservative by making choices that give a *lower bound* to the net consumer surplus lost. All details can be found in Supplementary Material, Section A.7.

Approximately 70% of the pure cash riders stop using Uber after the ban of cash according to the evidence from Puebla, which is a very similar city to the State of Mexico in the context of the cities served by Uber across Mexico. From Table 4 our estimated elasticities at pre-ban prices are approximately 1.38 for this group, so their consumer surplus loss is almost 47% of their yearly expenditure in Uber. For the remaining 30% of riders, the losses are smaller.[62]

For the remaining 30% of riders, those who pay the fixed cost and return to the application, the losses are smaller. Pure cash users who transition to exclusively using cards significantly reduced their number of trips after the ban. Using this information, along with the excess migration rates estimated in Experiment 3, we calculate a lower bound for the net consumer surplus for pure cash users who register a card after the ban of about 44% of their yearly expenditure on Uber. The net consumer surplus lost varies significantly across the distribution of riders. The consumer surplus lost is higher for pure cash users who travel more due to the convexity of the

60. Supplementary Material, Table N36 shows unconditional migration rates—users registering a card in the application regardless of whether they took trips during the weeks of the experiment. The table shows that the overall unconditional migration, over the 6 weeks that the experiment lasted, are similar to those presented in Table 7.

61. In Supplementary Material, Section P, we compare Puebla and the State of Mexico and we correct our estimates to take into account observable differences between Puebla and the State of Mexico, which may lower this estimate up to 29%; Puebla's residents have in average about one more year of education, and have higher financial inclusion. In the spirit of obtaining a lower bound on the consumer surplus lost, we retain the 30% figure.

62. Supplementary Material, Section A.7 presents the detailed calculations for this lower bound. It also shows the net consumer surplus lost computed cell-by-cell, where the cells are percentiles of the distribution of the historical number of trips.

net consumer surplus lost and the skewness of the distribution of historical trips.[63] Aggregating both groups, riders who register a card and riders who do not register a card after a ban on cash, the ban on cash results in an average loss of consumer surplus for pure cash riders, amounting to about 46% of their total Uber expenditures.

6. CONSUMER SURPLUS ESTIMATES

6.1. *Taking stock*

The consumer surplus lost after a ban on cash payments has a *lower bound* of at least 50% of total expenditures on cash-fare Uber rides, and an *upper bound* of 57%. We proceed by summarizing how we computed these estimates. For mixed users, who account for about 50% of Uber fares, we estimate a loss in consumer surplus of about 25% of what they spend on Uber rides. For pure cash users, who account for 20% of all fares collected by Uber and tend to be poorer, we estimate a loss in consumer surplus of at least 46%. Adding up the loss of consumer surplus from mixed users and pure cash users, the consumer surplus lost is about 30% of what the two groups spend on Uber rides. Considering that mixed users pay for about 37% of their Uber rides with cash, we obtain our 50% headline figure for the lower bound of the consumer surplus lost in a ban on cash.[64] An upper bound of this estimate can be found if we do not account for pure cash riders registering a card in the app after a ban on cash. The upper bound estimates are very close to the lower bound estimates, at about 57% of total expenditure on cash-fare Uber rides. The magnitudes of our estimates reflect (i) the intensity with which cash is used in the application by *both* mixed users and pure cash users, (ii) the convexity of the demand curve for Uber rides, (iii) the high costs of registering cards, and (iv) the fact that users view cash and cards as far from perfect substitutes.

6.2. *Short-run versus long-run*

Throughout this paper, our objective has been to obtain credible consumer surplus estimates by considering price variations over several periods. First, we validate our estimates of the elasticity of demand, denoted as ϵ, using data from a 4-week price experiment that more closely approximates permanent changes in prices. We also rely on the natural experiment in Panama to gather information on large price increases lasting more than a year. Second, we estimate the distribution of fixed costs through a 6-week experimental design and data from an actual ban on cash payments in Puebla (*i.e.* ψ_{ban}). Lastly, our survey instrument is designed to specifically elicit responses to *permanent* price increases, and the choke prices implied by our structural framework align with the responses obtained from the survey.

However, the short-term nature of the price variations used to estimate the elasticity of substitution (η) could introduce a bias, potentially resulting in a lower elasticity estimate than the long-run elasticity, thus leading to an overestimation of the consumer surplus loss.

To ameliorate this concern, we draw upon the quasi-natural experiments documented in Alvarez and Argente (2022), which provide information on the long-run elasticity of substitution across payment methods, *given an estimate of the elasticity of demand and the combination of CES and semi-log assumptions*. Thus, the estimates in this paper for the elasticity of demand

63. The calculation for the consumer surplus lost is the average of the consumer surplus of pure cash users and mixed users weighted by their share of total cash expenditures: $0.46 \times \frac{0.20}{0.2+0.5\times 0.37} + 0.67 \times \frac{0.5\times 0.37}{0.2+0.5\times 0.37} > 0.5$.

64. More specifically, to estimate η from observational data, we rely on Supplementary Material, equation (28) to formulate an expression for the change in total trips before and after the ban on cash: $\Delta T = \alpha^{\frac{1}{1-\eta}} \left(1 - \frac{\epsilon}{1-\eta} \ln \alpha\right)$. Since the change in trips is observed in the data, we can use our estimates for ϵ and data for α to infer η.

ϵ allow us to recover η from observational data in Puebla, following an actual ban on cash payments.[65] We believe this estimate for η complements the estimates using experimental data presented in this paper. While the estimate in this paper comes from short-run variation, it is cleanly identified using exogenous changes in relative prices. On the other hand, the estimate that can be obtained using data in Alvarez and Argente (2022) is indirect, as it relies on assumed functional forms, given that prices change from zero to infinity. It may also be contaminated by confounding factors occurring after the ban on cash in Puebla, such as changes in crime. However, it offers a long-run estimate of the elasticity of substitution, which is relevant for the evaluation of a permanent ban on cash. Our short-run estimates (*i.e.* $\eta \approx 3$) are close but smaller than the long-run estimates (*i.e.* $\eta \approx 5$). This ordering is reassuring since it is consistent with the Samuelson/Le Chatelier principle. Using the *long-run* elasticity, the consumer surplus loss for mixed users is 11% of what mixed users spend on Uber. Considering both pure cash and mixed users, we obtain a long-run estimate of 38% of total cash expenditures. Figures 4 and 5 show the sensitivity of the consumer surplus estimates to a wider range of demand elasticities.

6.3. *Distributional consequences*

To determine the *distributional* consequences of a cash payment ban, we estimate the consumer surplus at the municipality level, the finest level of aggregation with available per capita income data. Using geolocalized trip information from the State of Mexico in August 2018, we assign each user to the municipality where most of their trips originated.[66] After classifying riders into pure cash users and mixed users, we calculate the consumer surplus for both groups; pure cash users spend approximately \$80 per year on Uber rides, while mixed riders spend about \$200 per year.[67] The consumer surplus lost at the municipality level is the average consumer surplus of pure cash users and mixed users, weighted by their share of total expenditures on Uber in the municipality.

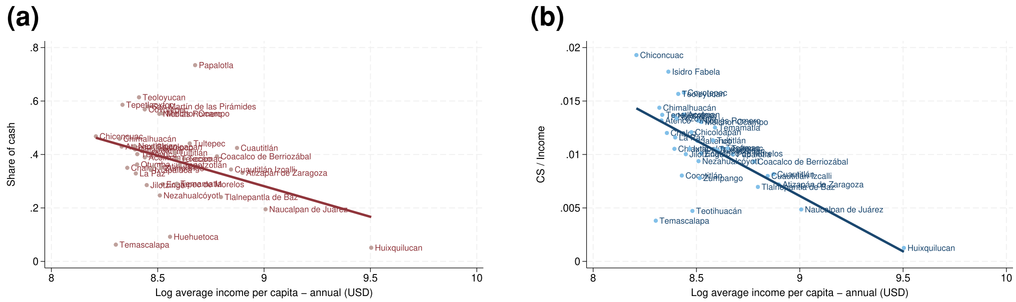
Figure 9b displays the consumer surplus loss as a share of the average annual income of Uber riders at the municipality-level. A rider who uses cash either sometimes or exclusively suffers an average loss in consumer surplus of approximately 0.8% of her annual income.[68] The figure also shows that cash ban losses fall mostly on households who reside in low-income municipalities. These households rely more heavily on the cash option. Figure 9a indeed shows that the share of cash fares declines with income per capita.

65. Supplementary Material, Section I shows that income per capita is correlated with financial access, commuting times, transportation modes, public infrastructure, and the use of cash in Uber.

66. Supplementary Material, Table E1 reports the weekly expenditures by pure cash users and mixed users. In order to accurately classify riders across user types, we consider riders with at least four trips in August 2018. We also use the share of payments made with card at the municipality-level to reduce potential measurement error.

67. The average annual income of Uber users in the State of Mexico is approximately 6,400 USD. This estimate is calculated by averaging the per capita income across municipalities weighted by the total number of Uber users in each municipality. We multiply this number by 1.24 to account for the fact that individuals with access to a smartphone earn higher income. The income data comes from Intercensal Survey of 2015 and the data on smartphone usage comes from the 2015 National Survey of Financial Inclusion (ENIF).

68. Supplementary Material, Figure I1 shows a similar pattern for the share of pure cash users at the municipality-level.



Share of Cash Fares Consumer Surplus

FIGURE 9

Share of cash fares and consumer surplus by income (a) share of cash fares and (b) consumer surplus

Notes: Panel (a) shows the share of cash fares and the average income per capita per year in USD. Panel (b) shows the consumer surplus from paying Uber in cash in each municipality of the State of Mexico as a fraction of the average income per capita per year. The consumer surplus lost at the municipality-level is the average of the consumer surplus of pure cash users and mixed users weighted by their share of total expenditures. We multiply income times 1.24 to account for the fact that individuals with access to a smartphones earn higher income. The income data comes from individuals that report labour income surveyed in the Intercensal Survey of 2015. The data on smartphone usage comes from the 2015 National Survey of Financial Inclusion (ENIF). The data of Uber rides are from August of 2018 in the State of Mexico.

7. CONCLUSION

Policies restricting means of payment have recently received great interest, and their possibility has been debated both by policymakers and academics. There are very few attempts to quantify the welfare effects of such policies, mainly because opportunities for accurate estimations of the relevant elasticities for this calculation for a given good or service are rare. In this paper, we combine a theoretical model with three large field experiments in Mexico to estimate the consumer surplus of using cash as a payment method in Uber. The total consumer surplus lost after a ban on cash payments is large, equivalent to 40–50% of total expenditure on cash-fare Uber rides. Given that the majority of trips paid in cash originate in low-income municipalities, these losses fall mostly on the least-advantaged households, who rely heavily on the cash payment option.

We have several other findings of interest for the literature on money demand and for the analysis of policies attempting to encourage or discourage payment methods. We found a statistically significant but small elasticity of the adoption/registration of cards when riders are given incentives. A reward of 15 USD increases the adoption rate by less than 4%, which is largely explained by the registration of existing cards. Nevertheless, users who registered a card after receiving a reward were more likely to use it to pay for rides in the future.

We also provide a well-estimated elasticity of substitution across payment methods using experimental data, an important input to models that incorporate a choice between means of payment. The low substitutability across payment methods implies that the optimal response of shifting away from cash payments (*e.g.* during the COVID-19 pandemic) is not without cost, even if people have access to other means of payment. This elasticity of substitution can be used to parameterize models designed to analyse counterfactuals in which means of payment are subject to a tax or a subsidy. For example, Alvarez *et al*. (2022) use our estimates to quantify the private costs of heavily taxing the use of cash to pay for all goods in Mexico and found that the private losses that follow a 40% tax on cash are approximately 6% of GDP. The extension of our analysis of Uber trips to the analysis of different goods and services, as well as our estimated elasticities, are important areas for future research.

Acknowledgments. We want to thank Andy Abel, Daron Acemoglu, Manual Amador, Marios Angeletos, George Alessandria, Andy Atkeson, Gadi Barlevy, Ben Bernanke, Mark Bils, Anmol Bhandari, Stephane Bonhomme, Sara Castellanos, Doireann Fitzgerald, Greg Kaplan, Narayana Kocherlakota, John List, Ellen Mc Grattan, Juanpa Nicolini, Francesco Lippi, Enrique Seira, Gaby Silva-Bavio, Rob Shimer, Harald Uhlig, Diana Van Patten, Venky Venkateswaran, and Ivan Werning for their comments and suggestions. We are grateful to Gabriel Chodorow-Reich for his very helpful discussion. We also want to thank Libby Mishkin and other members of the San Francisco Uber policy group, and especially the Uber Mexico team for general support and assistance, programming, implementing the field experiments, and for countless queries of data and background information, in particular to Federico Ranero, Daniel Salgado, and Hector Argente. We thank Basil Halperin for his many contributions during the initial phase of the project. We also thank Dan Ehrlich, Rafael Jimenez, and Francisca Sara-Zaror for excellent research assistance. None of the two authors are employees or consultants for Uber and have not received payments of any kind from Uber. David Argente is the brother of Hector Argente, who was the Research and Analytics Manager for Uber Latin America. This study is registered in the AEA RCT Registry and the digital object identifier (DOI) is: "10.1257/rct.6797-1.0." First draft: March 2019.

Supplementary Data

Supplementary data are available at *Review of Economic Studies* online.

Data Availability Statement

The data underlying this article were provided by Uber under a Data Usage Agreement and cannot be shared publicly. The code underlying this research is available on Zenodo at https://doi.org/10.5281/zenodo.13826268.

REFERENCES

ALLEN, R. and REHBECK, J. (2018), "Assessing Misspecification and Aggregation for Structured Preferences. Manuscript" (Working Paper).

ALVAREZ, F. and ARGENTE, D. (2022), "On the Effects of the Availability of Means of Payments: The Case of Uber", *The Quarterly Journal of Economics*, **137**, 1737–1789.

ALVAREZ, F., ARGENTE, D., JIMENEZ, R., *et al.* (2022), "Cash: A Blessing or a Curse?" *Journal of Monetary Economics*, **125**, 85–128.

ALVAREZ, F. and LIPPI, F. (2017), "Cash Burns: An Inventory Model with a Cash-Credit Choice", *Journal of Monetary Economics*, **90**, 99–112.

AMROMIN, E., JANKOWSKI, C. and PORTER, R. D. (2006), "Inducing More Efficient Payment on the Illinois Tollway" Chicago Fed Letter.

ANDERSON, S. P., DE PALMA, A. and THISSE, J. F. (1987), "The ces is a Discrete Choice Model?" *Economics Letters*, **24**, 139–140.

ANGRIST, J. D., CALDWELL, S. and HALL, J. V. (2021), "Uber Versus Taxi: A Driver's Eye View", *American Economic Journal: Applied Economics*, **13**, 272–308.

BORZEKOWSKI, R., KISER, E. K. and SHAISTA, A. (2008), "Consumers' Use of Debit Cards: Patterns, Preferences, and Price Response", *Journal of Money, Credit, and Banking*, **40**, 149–172.

BRIGLEVICS, T. and SCHUH, S. (2020), "This is 'what's in your wallet' ... and Here's How You Use it".

BUERA, F. J., KABOSKI, J. P. and SHIN, Y. (2021a), "The Macroeconomics of Microfinance", *The Review of Economic Studies*, **88**, 126–161.

BUERA, F. J., KABOSKI, J. P. and TOWNSEND, R. M. (2021b), "From Micro to Macro Development".

CHING, A. T. and HAYASHI, F. (2010), "Payment Card Rewards Programs and Consumer Payment Choice", *Journal of Banking & Finance*, **34**, 1773–1787.

CHODOROW-REICH, G., GOPINATH, G., MISHRA, P., *et al.* (2020), "Cash and the Economy: Evidence from India's Demonetization", *The Quarterly Journal of Economics*, **135**, 57–103.

COHEN, P., HAHN, R., HALL, J., *et al.* (2016), "Using Big Data to Estimate Consumer Surplus: The Case of Uber" (Working Paper No. 22627, National Bureau of Economic Research).

DEVIATOV, A. and WALLACE, N. (2014), "Optimal Inflation in a Model of inside Money", *Review of Economic Dynamics*, **17**, 287–293.

DUBÉ, J. P. H., JOO, J. and KIM, K. (2022), "Discrete-Choice Models and Representative Consumer Theory" (NBER Working Papers No. 30130, National Bureau of Economic Research, Inc).

FREEMAN, S. and KYDLAND, F. E. (2000), "Monetary Aggregates and Output", *The American Economic Review*, **90**, 1125–1135.

HAINMUELLER, J. (2012), "Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies", *Political Analysis*, **20**, 25–46.

HAINMUELLER, J., HANGARTNER, D. and YAMAMOTO, T. (2015), "Validating Vignette and Conjoint Survey Experiments Against Real-World Behavior", *Proceedings of the National Academy of Sciences*, **112**, 2395–2400.

HALL, J. V., HORTON, J. J. and KNOEPFLE, D. T. (2023), "Ride-Sharing Markets Re-Equilibrate" (Technical Report, National Bureau of Economic Research).

HANSEN, L. (2007), "Generalized Method of Moments Estimation", in Durlauf, S. and Blume, L. (eds) *The New Palgrave Dictionary of Economics*.
HAUSMAN, J. A. (1981), "Exact Consumer's Surplus and Deadweight Loss", *The American Economic Review*, **71**, 662–676.
HUMPHREY, D. B., KIM, M. and VALE, B. (2001), "Realizing the Gains from Electronic Payments: Costs, Pricing, and Payment Choice", *Journal of Money, Credit, and Banking*, **33**, 216–234.
KABOSKI, J. P. and TOWNSEND, R. M. (2011), "A Structural Evaluation of a Large-Scale Quasi-Experimental Microfinance Initiative", *Econometrica: Journal of the Econometric Society*, **79**, 1357–1406.
KARLAN, D., OSMAN, A. and ZINMAN, J. (2016), "Follow the Money not the Cash: Comparing Methods for Identifying Consumption and Investment Responses to a Liquidity Shock", *Journal of Development Economics*, **121**, 11–23.
KIYOTAKI, N. and WRIGHT, R. (1989), "On Money as a Medium of Exchange", *Journal of Political Economy*, **97**, 927–954.
KLEE, E. (2008), "How People Pay: Evidence from Grocery Store Data", *Journal of Monetary Economics*, **55**, 526–541.
KOCHERLAKOTA, N. (1998), "Money is Memory", *Journal of Economic Theory*, **81**, 232–251.
KOULAYEV, S., RYSMAN, M., SCHUH, S., *et al.* (2016), "Explaining Adoption and use of Payment Instruments by us Consumers", *The Rand Journal of Economics*, **47**, 293–325.
LACKER, J. M. and SCHREFT, S. L. (1996), "Money and Credit as Means of Payment", *Journal of Monetary Economics*, **38**, 3–23.
LAGOS, R. and WRIGHT, R. (2005), "A Unified Framework for Monetary Theory and Policy Analysis", *Journal of Political Economy*, **113**, 463–484.
LUCAS, R. E. (1987), "Money and Interest in a Cash-In-advance Economy", *Econometrica: Journal of the Econometric Society*, **55**, 491–513.
LUCAS, R. E. and NICOLINI, J. P. (2015), "On the Stability of Money Demand", *Journal of Monetary Economics*, **73**, 48–65.
MÉNDEZ, E. and VAN PATTEN, D. (2021), "Voting on a Trade Agreement: Firm Networks and Attitudes Towards Openness".
PARKER, J. A. and SOULELES, N. S. (2019), "Reported Effects versus Revealed-Preference Estimates: Evidence from the Propensity to Spend tax Rebates", *American Economic Review: Insights*, **1**, 273–90.
PRESCOTT, E. C. (1987), "A multiple means-of-payment model", in: Barnett, W. A. and Singleton, K. J. (eds) *New Approaches to Monetary Economics: Proceedings of the Second International Symposium in Economic Theory and Econometrics. International Symposia in Economic Theory and Econometrics* (Cambridge University Press) 42–51.
ROGOFF, K. S. (2017), *The Curse of Cash: How Large-Denomination Bills Aid Crime and Tax Evasion and Constrain Monetary Policy* (Princeton University Press).
STOKEY, N. (2019), "Means of Payment" (Working Paper. University of Chicago).
TOWNSEND, R. M. (2020), *Distributed Ledgers: Design and Regulation of Financial Infrastructure and Payment Systems* (MIT Press).
VIVES, X. (1987), "Small Income Effects: A Marshallian Theory of Consumer Surplus and Downward Sloping Demand", *The Review of Economic Studies*, **54**, 87–103.
WANG, L., WRIGHT, R. and LIU, L. Q. (2017), *Money and Credit: Theory and Applications* (International Monetary Fund).
WHITESELL, W. C. (1989), "The Demand for Currency versus Debitable Accounts: Note", *Journal of Money, Credit, and Banking*, **21**, 246–251.
YANG, B. and CHING, A. T. (2014), "Dynamics of Consumer Adoption of Financial Innovation: The Case of atm Cards", *Management Science*, **60**, 903–922.