Essays in Development, Environment and Health

Prabhat Barnwal

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2015

ABSTRACT

PhD

Prabhat Barnwal

This dissertation examines economics and public policy related topics in the context of developing countries using empirical methods.

In the first chapter, I study whether increased enforcement reduces leakage in public programs. In the context of fuel subsidies in India, this study attempts to understand how the formal sector and black market respond to a policy that reduces leakage. I explore these questions using a recent fuel subsidy transfer policy in India.

Second chapter is on health and wealth trade off near mineral mining operations in developing countries. Using extensive data on mining, health outcomes and assets from 44 developing countries, this study quantifies the wealth gain and adverse health impact of mineral mining.

In the third chapter, I estimate demand for a water quality diagnostic product – arsenic testing, when it is offered at a price. I further look into various aspects related to selection, learning and households behavioral response to the information. This study is based on a field experiment in Bihar, India.

Chapter 1

Curbing Leakage in Public Programs with Biometric Identification Systems: Evidence from India's Fuel Subsidies

Curbing Leakage in Public Programs with Biometric Identification Systems: Evidence from India's Fuel Subsidies

Prabhat Barnwal*

Abstract

In many developing countries, pervasive corruption and evasion often undermine the provision of public programs. I focus on India where a large universal program provides USD 8 billion in fuel subsidies for domestic cooking. The subsidy given to households, combined with taxes on commercial users, gives rise to a black market, where fictitious "ghost" beneficiaries are used to divert the subsidy from the domestic to the commercial sector. This paper studies the impact of a policy introduced in 2013 that makes biometrics-based Unique Identification (UID) mandatory for beneficiaries of the program in order to improve the state's capacity to purge ghost beneficiaries. The analysis is based on unique data that combine the administrative records from 23 million fuel purchase transactions and distributor-level fuel sales, with a novel survey data set which allows me to infer black market prices. My empirical strategy exploits two quasi-experiments: (i) the phase-wise policy roll-out across districts and (ii) its unexpected termination. I document four main findings: (1) the UID-based transfer policy reduces fuel purchases in the domestic fuel sector by more than 11%, suggesting a reduction in subsidy diversion; (2) after the policy is terminated, fuel purchases in the domestic sector revert to a similar level that existed before the policy was introduced; (3) a positive supply shock induced by the policy termination reduces black market prices between 13% to 19%, and; (4) in response to the lower prices, commercial firms reduce their fuel purchases in the formal sector and re-enter the black market for fuel. In sum, this paper illustrates that investment in enforcement capacity can significantly strengthen the state's ability to target program beneficiaries.

JEL Codes: D73, E26, O30, H26, Q48

Keywords: *State capacity, Corruption, Subsidy diversion, Black market, Fuel subsidy, Biometric authentication, UID, India*

---

*School of International and Public Affairs, Columbia University. *email:* pb2442@columbia.edu. I am grateful to the Ministry of Petroleum and Natural Gas (India) and Hindustan Petroleum Corporation Limited for their continued support. I especially thank my advisors Wojciech Kopczuk, Cristian Pop-Eleches and Eric Verhoogen. For helpful discussions, I thank Douglas Almond, Emily Breza, Ritam Chaurey, Ram Fishman, Raymond Fisman, Francois Gerard, Jessica Goldberg, Michael Greenstone, Jonas Hjort, Supreet Kaur, Dilip Mookherjee, Karthik Muralidharan, Suresh Naidu, Arvind Panagariya, Nicholas Ryan, Pronab Sen, Rathin Roy, Wolfram Schlenker, Sandip Sukhtankar, Anna Tompsett, Jan von der Goltz and the seminar participants at Columbia University, NEUDC 2014 – Boston University, IGC Growth Week – LSE and IGC conferences in India. I am specially grateful to Shyam Peri for the outstanding help with the data. I gratefully acknowledge the financial support from the International Growth Center (IGC) and Center for International Business Education and Research (CIBER) of Columbia University. All errors are mine.

1 Introduction

States with low fiscal capacity have inefficient tax and transfer systems, which directly affect the availability of public funds for economic development (Acemoglu, 2005; Besley and Persson, 2013). In developing countries, pervasive corruption and evasion often characterize inefficiency in public programs, whereas the presence of an underground economy further undermines the state's ability to collect taxes and provide welfare. Notwithstanding its central importance, there is little empirical research on investment in state capacity to enforce tax and transfer policies (Besley et al., 2013). This paper illustrates how investing in a technology-driven enforcement system can improve a state's ability to administer transfer programs.

In public programs, information asymmetries between the government and the local officials overseeing subsidy disbursement – where the former has little means of authenticating the identity of beneficiaries, and the latter have perverse incentives to misreport – manifest in *subsidy diversion* or *leakage* under the rubric of "ghost" beneficiaries. This term represents an existing or non-existing person, under whose name benefits are illegally drawn. Ghost beneficiaries represent a general problem observed in many developing countries.[1] For example, Fig. 1 illustrates that the number of households that drew fuel subsidy benefits in 2010-11 is about 50% higher than the estimates shown by the 2011 census. Policymakers often recognize the significant fiscal burden of subsidy diversion.[2] Yet caught between the incentives and constraints that are shaped by a weak institutional environment, states have little means to improve their targeting. This paper shows that investing in the enforcement capacity can significantly strengthen a state's ability to target program beneficiaries.

In this paper, I provide econometric evidence regarding the impact of biometrics-based enforcement on subsidy diversion. India's *Aadhaar* program seeks to cover all the 1.2 billion residents under the ambitious biometrics-based Unique Identification (UID) system. A unique identity number helps in the authentication of beneficiaries and facilitates a centralized payments infrastructure through which benefits can be directly transferred. Thus the UID-based transfer system makes it prohibitively expensive for the ghost beneficiaries to carry out identity fraud, and it shifts the

[1] A few examples include the 1.5 million ghost voters in Ghana (Smith, 2002); ghost farmers in Malawi (Eggen, 2012); ghost soldiers in Uganda (Tangri and Mwenda, 2008); the 40% figure for ghost workers in Zimbabwe (Parliament of Zimbabwe, 2012); and officials gaming the traditional ID system in China (Kaiman, 2013).

[2] For example, India's then finance minister stated in 2012 that he was *"losing sleep over subsidy leakage, not [the] subsidy itself."*, which is in line with Milka Casanegra de Jantscher's famous observation on developing countries: *"Tax administration is tax policy"*.

task of monitoring and disbursing transfers away from local officials. This paper investigates how the UID-based transfer policy enables the state to reduce subsidy diversion in the USD 8 billion cooking fuel subsidy program, which has 165 million listed beneficiary households (2013–14). Domestic cooking fuel is universally subsidized, while commercial fuel purchases are taxed.[3] Diverted domestic fuel changes hands in the black market, eventually being consumed for commercial purposes. Using novel administrative and survey datasets, I track the simultaneous responses to the UID-based enforcement in the domestic, commercial and black market sectors.

In my analysis, I first lay down a simple framework to study the impact of the UID-based transfer policy upon the cooking fuel markets. The de-duplication[4] aspect of biometrics-based enforcement, if effective, would enhance the state's ability to eliminate transfers to ghost beneficiaries in the subsidized domestic fuel sector (i.e., fuel for household domestic cooking use). This would, in turn, restrict the supply of subsidized fuel in the black market, increasing the black market equilibrium prices. Subsequently, higher black market prices would induce firms in the commercial sector to exit the black market and increase fuel purchases through formal channels. Overall, increasing enforcement with the UID-based transfer policy in the domestic sector should affect the fuel volume purchased in domestic and commercial sectors, and the black market prices.

Next, I test these predictions empirically using a difference-in-differences framework. My identification strategy exploits two quasi-experiments: (1) the time variation in the roll-out of the UID-based transfer policy and (2) its unexpected termination. The UID-based transfer system was coupled with the Liquefied Petroleum Gas (LPG) subsidy delivery program in June 2013, and gradually had been introduced in about 300 districts in six phases by January 2014. However, the program was unexpectedly terminated during the run-up to the general elections in early 2014.

My analysis is based on two unusual and attractive datasets. First, I tap into a rich transaction-level administrative database on cooking fuel purchases. LPG is the most preferred cooking fuel in urban areas. More than 3 million LPG refills are purchased every day, resulting in about one billion transactions a year. I use a representative sample of 23.2 million transactions that are carried out by about 4 million households in 509 districts. In addition, I use LPG distributor-level monthly data

[3]Even though this seems inefficient, policies like this exist in developed countries as well. For example, in the US, commercial use of diesel fuel is taxed, but residential use for heating purposes is not.

[4]De-duplication is the process to remove instances of multiple enrollments in public programs by the same individuals who use a fictitious or ineligible person's name to withdraw benefits.

from more than 3,000 distributors that covers fuel sales in the domestic and commercial sectors.

Second, I directly address a perennial problem faced by researchers studying the underground economy: detecting black-market transactions. Black-market activities are difficult to measure. Previous studies on black markets and tax evasion have generally employed a methodology to compare information from different sources (i.e. reporting gap) (Pissarides and Weber, 1989; Fisman and Wei, 2001, 2009).[5] My approach involves a direct audit survey with agents in the blackmarket, which, to my knowledge, is unique in this literature.[6] I organized a team to carry out the field work in 89 districts in 11 states specifically covering the period before and after policy termination. Using a systematic data-collection strategy in the black market, I collect prices from the demand side (i.e. small businesses) in local markets to construct a firm-level panel data. At the same time, I conduct innovative audit surveys of the LPG delivery men. Audit surveys help in disallowing any misreporting concerns. The approach used in my fieldwork leverages the fact that the fuel price information is readily revealed by the unsuspecting agents in the local markets, in part, because cooking fuel is such a ubiquitous commodity.

My main results suggest about a 11% to 14% reduction in fuel purchases in the domestic sector when the UID-based transfer policy is enforced. Exploiting the policy termination for identification, I show that domestic fuel sales in the treated districts converge to fuel sales in the control districts when the UID-based transfer policy is terminated. In the black market, the policy termination would bring a positive supply shock, and the analysis of the firm level panel data suggests about a 13% to 19% decrease in prices in treated districts. Black market supply side price data further confirms this. Because of the sudden drop in LPG black market prices, the firms would return to the black market. Using LPG distributor-month level panel data, I find the commercial sector reduced its formal fuel purchase by about 6% to 9% in response to the termination of the enforcement policy in the domestic sector. Taken together, the empirical evidence from these three LPG sectors suggest that the UID-based transfer policy was effective in reducing subsidy diversion.

I further address two important aspects regarding the enforcement and targeting in public programs. First, it is a valid concern that implementation issues in large scale technology-based enforcement programs may lead to an exclusion of genuine beneficiaries. In Section 5.3 and 5.4, I

[5]For recent papers on estimation of tax evasion and informal economy, see (Slemrod and Weber, 2012). Buehn and Schneider (2013) survey methods used to estimate the underground economy at macro level.

[6]See Merriman (2010) for an interesting application of indirect audit method to estimate cross-border tax evasion.

explore the purchase behavior of late complier and non-complier households. I find that the beneficiaries who drew more subsidies in the pre-enforcement period, are less likely to comply. Second, an increased in enforcement leading to an effective curb on corruption raises concerns about displacement in corruption (Yang, 2008), i.e., a reduction in corruption by the ghost beneficiaries may have to contend with a potential increase in subsidy diversion by genuine households. I find that increased fuel purchases by genuine households do not come close to outweighing the impact of increased enforcement.

This paper relates to three different strands in the literature. First, this paper aims to fill the gap between empirical and theoretical work on state capacity (Besley and Persson, 2009, 2010, 2013). In a weak institutional environment, there is a higher degree of uncertainty about the impact of large investments in the fiscal capacity to tax and transfer, such as in a UID-based transfer infrastructure.[7] My paper shows that there are tangible gains to strengthening the state's enforcement capacity. Using evidence from formal and informal sectors, I further show how the gains made during the enforcement can quickly be lost when the enforcement policy is reverted. In modern tax literature, strengthening enforcement with more information is discussed as the key to increasing tax compliance (Kopczuk and Slemrod, 2006; Gordon and Li, 2009; Kleven et al., 2009).[8][9] By highlighting the impact of an advanced information-based enforcement infrastructure, my paper illustrates that the prevalence of identity fraud may partly explain the high level of evasion in tax and transfer programs in developing countries.

Second, this paper contributes to the literature on targeting and corruption in developing countries (Reinikka and Svensson, 2001; Olken, 2006; Niehaus and Sukhtankar, 2013b,a).[10] I show that technology-based governance reforms can effectively reduce corruption. My results compliment the recent theoretical work on controlling corruption using a task-focused approach (Banerjee et al., 2012). Specifically, this paper shows that shifting the authentication and monitoring tasks away from the local officials can reduce corruption, while keeping the remaining Public Distribution System (PDS) intact. This paper directly contributes to a recent but growing literature on the

[7] E.g., more enforcement may not be optimal in the presence of an informal economy (Emran and Stiglitz, 2005).

[8] Recent papers have highlighted information-based approaches, as seen with the information trail (Gordon and Li, 2009; Kumler et al., 2013; Pomeranz, 2013), third party reporting (Alm et al., 2009; Kleven et al., 2011; Carrillo et al., 2014), and enforcement technology (Marion and Muehlegger, 2008; Casaburi and Troiano, 2013). See (Slemrod and Yitzhaki, 2002; Slemrod, 2007; Saez et al., 2012) for a review on the tax evasion literature.

[9] See Bird and Zolt (2008) for a review of technology's application in tax administration in developing countries.

[10] For a review of the empirical literature on corruption in developing countries, see Olken and Pande (2012).

effectiveness of payments and authentication infrastructure in developing countries.[11] Giné et al. (2012) show that fingerprinting borrowers leads to a higher payment rate in credit markets. In a context related to UID, Muralidharan et al. (2014) conducted a large scale policy experiment and show that the biometrics-based payments infrastructure increases efficiency in the transfers to authenticated beneficiaries. However, their setup does not allow the purging of ghost beneficiaries because the biometrics requirement was not mandatory. To the best of my knowledge, this paper is the first to study the impact of mandatory biometrics-based enforcement on corruption.

Thirdly, this paper highlights the inefficiency in fuel subsidy distribution in developing countries. Fuel subsidies are regressive, when they are cornered by the relatively better off population subgroups (McLure Jr, 2013; Arze del Granado et al., 2012). When subsidy diversion and fuel black markets are also taken into account, the standard economic guidance to eliminate fuel subsidies in developing countries is likely to have a different impact. Notwithstanding its importance, fuel subsidy diversion in developing countries remains relatively unexplored in the empirical literature, and my paper helps redress this.[12]

The rest of the paper is structured as follows. Section 2 discusses institutional context. Section 3 details a simple framework. Data and empirical strategy are discussed in Section 4. Results are reported in Section 5, and Section 6 concludes.

2 The Institutional Context and Policy Experiment

2.1 India's Biometrics-based Unique Identification Program

India is the first major country to roll out a national biometrics ID program in 2009. Similar to the Social Security Number system in the USA, registration under the Unique Identification (UID) program is not yet mandatory, but welfare transfers are being tied to it, thus assigning it a default status of national identification number.[13] Szreter (2007) argue that a national system of identity registration in 16^{th} century England contributed to the country's economic development by facilitating a functioning welfare system. Similarly, India's UID program is also frequently

[11] E.g., recent studies provide empirical evidence on the effectiveness of mobile money transfer programs in Africa (Mbiti and Weil, 2011; Jack and Suri, 2014; Aker et al., 2014).

[12] (Marion and Muehlegger, 2008; Kopczuk et al., 2013) discuss fuel tax evasion in the USA.

[13] UID is also known as *Aadhaar*. More details are available at `http://www.uidai.gov.in`. Gelb and Clark (2013b) provide a description of the UID program.

associated with potential efficiency gains in welfare programs and governance.[14] India aims to provide 1.2 billion residents a biometric ID – "nothing remotely similar in scale has ever been attempted" (Breckenridge, 2014). As of March 2014, around 600 million UIDs have been provided, which is in line with the target for the first five years. Every individual is allocated a unique 12-digit ID (Fig. A15).

Generally, identity fraud in public programs is feasible for two main reasons. First, welfare schemes accept documentary proof from multiple and isolated sources[15], which are not costly to counterfeit. This makes de-duplication and authentication very costly. Secondly, local level official can himself over-report (acting independently or in collusion with others) when the responsibility to authenticate and monitor lies with him.[16] Continuous de-duplication with audits is very expensive, particularly when the ghost beneficiaries immediately respond to new opportunities to divert subsidies.[17] De-duplication is central feature of UID program[18]. With the biometric data (i.e., fingerprints and an iris scan) and key demographic data collected during enrollment, an automated system carries out a 1:N match for each new record against the database of existing records. In January, for example, about a million UIDs were processed every day, which required more than 600 trillion checks in 24 hours.

UID further facilitates a centralized payments infrastructure (*Aadhaar Bridge*) by minimizing the possibility of fraudulent transfers.[19] Further, when households' bank accounts and their fuel subsidy accounts are linked with the UID, continuous verification in transactions is no longer needed. As UIDs become linked to various welfare programs, benefits are transferred directly to the bank accounts of the beneficiaries.[20] The UID-based transfer has been criticized for the high cost of UID

[14]On the other hand, there are privacy concerns associated with biometrics identification systems. The UK repealed its biometrics-based national identification card program (The Identity Cards Act 2006 (c 15)) in 2011. Developing countries may need to opt for second best institutions depending upon the constraints they face (Rodrik, 2008).

[15]For instance, an LPG subsidy beneficiary application provides the option to submit one of the 12 types of documents as proof of identity.

[16]In a strict sense, these two factors can be categorized as evasion and corruption respectively. In this paper, both are combined as "subsidy diversion", since they are often interlinked.

[17]A recent example is India's Food Security Act (FSA) 2013 which observed a quick flood of fictitious PDS applications. For instance, Bihar recently found more than half a million fake PDS cardholders with one year of the FSA 2013 (Singh, 2014).

[18]The UID program mentions in its mission statement – *"...that can be verified and authenticated in an online, cost-effective manner, which is robust enough to eliminate duplicate and fake identities"*.

[19]e.g., an individual can own multiple bank accounts and enroll as beneficiary in the same program multiple times. UID protects against multiplicity in bank transfers.

[20]The government of India has linked 28 welfare programs to the UID. While registering for a UID, the application asks for the individual's bank details and provides an option to open a new bank account. Bank accounts can also be linked separately.

and potentially low benefits on various grounds including exclusion errors (Khera, 2011).

2.2 LPG Sales, Subsidy, and the Black Market

India is one of the top five global LPG consumers. Government-controlled public sector sells LPG to households (for domestic cooking purpose) and to the commercial sector (i.e. businesses and the industrial and transport sectors). Household domestic cooking LPG sector is commonly called "domestic sector," whereas all remaining fuel users are classified under "commercial sector". [21] A total of 13,896 distributors deliver domestic and commercial LPG refills (Ministry of Petroleum and Natural Gas, India, 2014). Without considering subsidy diversion, the domestic sector consumes about 88% LPG fuel, with the remaining being consumed by the commercial, industrial and transport sectors. LPG is primarily an urban fuel with respectively 65.0% and 11.4% in urban and rural areas (Registrar General of India, 2011). Households are supplied with LPG in 14kg cylinders, whereas commercial customers are sold LPG in larger size cylinders.[22][23] An LPG refill transaction involves the exchange of an empty cylinder for an LPG-filled cylinder. Empty domestic-use cylinders are also rationed and are only supplied by the government appointed distributor [24], which restricts the LPG storage capacity of households.

The LPG Subsidy and Taxation

LPG is subsidized for domestic cooking use and is taxed when used for commercial purposes.[25] There is no free market, and *market price* is determined and revised by the government regularly according to the LPG price in international markets. The general LPG pricing structure is as follows. Households receive fuel by paying a subsidized fuel price of $p - s$, whereas firms purchase fuel at a tax-inclusive price of $p + t$. The regulated *market price* is p. Domestic subsidy and commercial tax are s and t, respectively. Fig. 3 (Panel (a)) illustrates this LPG pricing with January 2014 prices, when there was about 66% domestic cooking subsidy and 33% commercial taxes on LPG fuel. There is an annual cap on household refill purchase in the domestic sector, and beyond that

[21] Three public sector Oil Marketing Companies (OMCs) primarily supply LPG across the country. These OMCs are controlled by the Ministry of Petroleum and Natural Gas. Private LPG suppliers are allowed but their presence is limited to commercial fuel and the piped domestic supply in a few cities only. A heavy subsidy makes it difficult for the private companies to enter the domestic cooking LPG sector.

[22] LPG is also provided in 5kg cylinders in some selected cities.

[23] In other countries, LPG steel cylinder is also known as "bottle" or "canister".

[24] Of course, one can buy empty cylinder in the black market as well.

[25] Mostly excise tax and VAT. In a few states, domestic refills are also subjected to VAT, but overall there is substantial additional tax on commercial fuel.

beneficiary households have to pay market price. Annual cap is sufficiently relaxed and the current annual cap is not binding for more than 95% beneficiary households.

Enforcement of Market Segmentation

Since the same commodity is provided in two different sectors (domestic and commercial) at different prices, enforcement is needed to ensure market segmentation. Traditional enforcement of differential LPG pricing hinges on two main factors. First, the domestic LPG cylinder (i.e., LPG refills sold for domestic cooking use) is made to look different from the LPG cylinder meant for commercial consumption. Such visual differentiation in shape and size helps in reducing enforcement cost, similar to how red dye in untaxed diesel decreases enforcement cost in the USA (Marion and Muehlegger (2008)). However, this is not a foolproof strategy when monitoring is poor. Second, market segmentation is legally binding. LPG Regulation of Supply Distribution Order 2000 makes it illegal to sell, transport or re-refill LPG by unauthorized persons. Hoarding multiple LPG connections within a household is also illegal.[26] In practice, the audit and penalty policies are not very effective in curbing evasion, when enforcement requires greater coordination among agencies.[27] After-sales enforcement may also be susceptible to local-level corruption and side payments.

Subsidy Leakage and the Black Market

Because of the domestic-sector subsidy and the commercial-sector tax, differential pricing drives the incentive to trade LPG on the black market. Audits reveal that the number of fraudulently created LPG subsidy beneficiaries may run into the millions (Sastry, 2012). Further, the role of the delivery man (more commonly known as – the *gas hawker*) becomes pivotal in the black market. In general, the same delivery team is employed by the LPG distributorship to carry out regular doorstep deliveries to the households as well as to commercial customers. Thus, the delivery men hold permits to carry cylinders of both types at the same time. Even though it is difficult to track the original ghost cardholder, delivery men become the de facto supply side in the black market.

It should be noted that while severe legal provisions are applicable to the independent agents and cartels active in the black market, it is often difficult to punish the distributor or the delivery

[26] As per the LPG Order 2000, "No person other than a Government Oil Company, a parallel marketeer or a distributor shall be engaged in the business of selling LPG to the consumer." Similar restrictions are also present for transportation and storage. Further, using LPG cylinders meant for domestic sector in any other way is a non-bailable offence under the Essential Commodities Act, 1955.

[27] E.g.,Food Supplies, Weight and Measurement departments and the Police. Media often reports raids on cartels (PTI, 2014).

men even in cases where they have played an obvious role. In the case of critical irregularity by an LPG distributor, the stipulated penalty terms are less severe than the general penalty for operating in the black market (Ministry of Petroleum and Natural Gas, 2014).[28] LPG distributors have strong national- and state-level unions, and they regularly threaten to strike when reforms aimed at curtailing the subsidy diversion are announced.[29] LPG distributors also enjoy additional bargaining power because LPG fuel is an essential commodity, specially in urban areas.

2.3 UID Based Transfer Policy

The main policy intervention in this paper is the introduction of a biometric ID based transfer policy to provide fuel subsidies directly to the household. The Indian government introduced Direct Benefit Transfer for LPG (DBTL) in June 2013. After an initial transition period, the UID requirement was fully enforced and only compliant households received the subsidy. Irrespective of UID submission, non-subsidized LPG (i.e., at market price p) is available to all beneficiaries in the domestic sector.

One-time compliance requires LPG beneficiary households to provide their UID number either in person or through text message service, or the Internet. In addition, households need to link one bank account with the UID, if they have not already done so. Compliant households receive a subsidy amount in their bank account following each LPG refill purchase. In the transition period, households that had already complied receive transfers directly in their bank account, whereas non-compliant households kept availing LPG at the subsidized price. Once the policy was enforced in a district, all domestic fuel was sold at the market price p, and LPG subsidy was transferred only to the compliant households. Late complier households start receiving transfers after they comply.

A direct bank transfer can itself increase enforcement, irrespective of UID. This is possible if opening a bank account is subjected to stricter verification since the banks have different incentives than the LPG distributors. However, by design, bank accounts do not offer any de-duplication, and banks often have similar document requirements to those needed for opening an LPG beneficiary account, and in the absence of general overdraft privilege, banks may have little motivation to

[28] Distributors also enjoy a certain level of accounting autonomy. For example, *"Dealers find it terrible that sale of more than 50 'unaccounted' cylinders is regarded as a 'critical irregularity' [...They] want the limit to be increased to 300 cylinders."* (Jai, 2013).

[29] In 2013, delivery men threatened the government with strikes when the UID requirement was introduced (TNN, 2013). In 2012, LPG distributors went on strike when an annual cap on subsidized LPG refills per household was introduced (Kumar, 2012).

verify.[30] The real advantage the UID-bank account pairing offers, is that it does not require frequent verification for each subsidy transfer, and thus takes away the regular subsidy disbursement and monitoring responsibility from the LPG distributor.[31] Further, if there are barriers to banking access, it may affect the take up. However, according to 2011 census, bank account ownership in India dominates LPG penetration, in both urban and rural areas (Fig. A17).[32] For simplicity, I shall call the new UID-based direct bank subsidy disbursement policy as 'UID-based transfer'.

Phase-wise roll-out and selection of districts

UID-based transfer policy was implemented in phases. In Phase 1, the government of India introduced it in 20 districts (for timeline, see Fig. 2).[33] In Phase 2, the policy was introduced in 34 new districts in September 2013 and it was fully enforced in January 2014.[34] In total, the UID-based transfer policy was introduced in 291 districts from Phase 1 to Phase 6 by January 2014, out of which 42 districts (from nine states) had it fully enforced by January 2014. Selection of the districts in six phases was made on the basis of initial UID penetration in early 2013. The actual enforcement date was also subjected to a pre-determined cutoff to ensure wider UID take before making the compliance mandatory for subsidy transfer. Non-policy districts are not included in the 2013-2014 program for a reason not related to the LPG subsidy program.[35]

2.4 Policy Termination

On January 31, 2014, the Cabinet Committee on Political Affairs unexpectedly decided to terminate the UID-based LPG subsidy transfer policy. About 40 days after the surprising political announcement, the government terminated the policy and the household LPG subsidy transfer system was restored to the old system on March 10, 2014 (Nambiar (2014)). About 17 million

[30] India's Central Bank has expressed concerns about bank accounts with fake documents (PTI, 2014). There is evidence of *ghost bank accounts* in other developing countries as well (Mayah, 2012).

[31] Direct transfers may also reduce corruption in service provision (e.g., asking for bribe at the time of purchase).

[32] Due to greater emphasize on financial inclusion in recent years, banking access has significantly expanded since 2011. Moreover, LPG is primarily a fuel used by middle- to higher-income households, which makes it even less prone to concerns about banking access.

[33] District names are listed in Fig. A16.

[34] In Phase 1, two districts Mysore and Mandi had delayed implementation because of by-elections. In Phase 2, twelve districts in Kerala were not included in the mandatory implementation in January 2014.

[35] In January 2012, a cabinet committee divided biometric collection tasks between the Planning Commission (i.e., the UID program) and the Home Ministry (i.e., the National Population Register), which ultimately left a group of states with little UID coverage. Most of the 360 non-policy districts fall in these states – Bihar, Chhattisgarh, Jammu & Kashmir, Orissa, Tamil Nadu, Uttarakhand, Uttar Pradesh and northeastern states.

households already complying with the program were also moved back to the old regime.[36] While implementation related issues were cited for the decision in a closed door meeting, it does not fit with the Petroleum Minister's assertion that the policy was a success.[37] I look at the impact of delay in compliance in the results section. Having completed about six months of enforcement in Phase 1 districts, it seems the policy retrieval transpired after continued political lobbying in the run-up to the general elections. Around election time, special interests group may try harder to obstruct policy reforms in order to ensure the survival of the traditional rent-seeking structure.[38] This may get further reinforcement by politicians who have their own special interests in LPG distributorships.[39] The expected electoral costs of reduced rents may affect the timing of the adoption of enforcement technologies that improve transparency (Bussell, 2010).An incumbent government may also strive to reduce bias in the media, if teething problems with a new policy are selectively highlighted. Moreover, fuel subsidy reforms are generally susceptible to political business cycles in developing countries.[40]

3 Conceptual Framework

In this setup, LPG fuel is subsidized exclusively for household cooking usage (i.e., the domestic sector), whereas it is taxed in the commercial sector. Market segmentation creates an opportunity to make profit. Tax evasion and subsidy diversion are complimentary here – to evade the fuel tax, a firm has to purchase diverted subsidized fuel from the black market. An entrepreneurial agent can make a profit by obtaining fake or multiple beneficiary accounts. This ghost beneficiary then purchases subsidized fuel and sells it to a firm on the black market. I consider agents on both sides of the black market to be risk neutral and their decision to deal in the black market is a choice under certainty. p is the *market price* (i.e. the government regulated price based on LPG prices in the

[36]The policy termination also accompanied an increase in the annual refill cap from 9 to 11 in 2013–2014 in all the districts.

[37]The constitutional status of UID was also challenged in court (TNN, 2014), but this was not cited as the reason for termination.

[38]On policy manipulation by politicians seeking re-election (Nordhaus, 1975; Alesina, 1997). Some related contexts of corruption are discussed in (Cole, 2009; Kapur and Vaishnav, 2011; Burgess et al., 2012; Sukhtankar, 2012; Foremny and Riedel, 2014)

[39]A former minister in Karnataka state, himself an LPG distributor, explains – *"As a politician, I am telling you that 90% of the LPG dealers and black-marketeers in the state are either politicians, bureaucrats, or their kin."* This news story was about 2.4 million fake LPG beneficiaries who were uncovered in Karnataka state (Aji, 2012).

[40]E.g., in Indonesia – *"Parliamentary and presidential elections in 2014 mean that there is likely to be little appetite for further subsidy reforms in the first two thirds of 2014."* (International Institute of Sustainable Development, 2014).

international market), s is the subsidy in the domestic sector and t is tax in the commercial sector. Before the UID-based transfer policy, all domestic fuel is sold at subsidized price $p - s$[41], and the commercial fuel is available at $p+t$. Total LPG sales (D_{Total}) will consist of cooking fuel demand of genuine households ($D_{household}$), the ghost beneficiaries (D_{ghost}) and the commercial firms (D_{firm}). Ghost beneficiaries purchase fuel in the domestic sector and sell it to the firms.

$$D_{Total} = \underbrace{D_{household} + D_{ghost}}_{Domestic\ fuel\ sector} + \underbrace{D_{firm} - D_{ghost}}_{Commercial\ fuel\ Sector} \tag{1}$$

The black market supply and demand is conditional on the level of enforcement C. Say, C_{ghost} and C_{firm} denote the net expected cost to the supply and demand side respectively,[42]. The equilibrium black-market price P is determined by the supply (S_{bm}) and demand (D_{bm}) in the black market:

$$S_{bm}(P - C_{ghost}) = D_{bm}(P + C_{firm}) \tag{2}$$

Thus, the black-market price is increasing in C_{ghost}. The UID-based transfer policy makes it prohibitively expensive for the ghost beneficiaries to receive a fuel subsidy (i.e. C_{ghost} increases significantly), and so, the black-market price increases. Further, firms have an outside option available for purchasing the fuel legally at a tax-inclusive price ($p + t$):

$$D_{firm} : D_f[min(p + t, P + C_{firm})] \tag{3}$$

When $P \geq p+t-C_{firm}$, firms may decide to leave the black market and purchase fuel in the formal sector. If there are other fuel options available below $p + t$, firm may even switch to a new fuel.

However, in our setup, a ghost beneficiary can still continue to obtain non-subsidized fuel (i.e., at price p) without a UID. Before the UID-based transfer policy, the ceiling and floor in the black market is determined by the subsidy diversion opportunities available to the ghost, tax level and existing enforcement on either side of the black market, $P \in (p - s + C_{ghost}, p + t - C_{firm})$. When the UID-based transfer policy in enforced, the price floor shifts up because the UID requirement eliminates subsidies to the ghost beneficiaries: $P_{UID} \in (p + C_{ghost}, p + t - C_{firn})$ (Fig. 3).

[41]Annual cap on subsidized fuel is ignored here, which was not binding for more than 95% of the beneficiaries in 2013-14.

[42]Here firm level heterogeneity in C_{firm} is ignored. In a richer model, there may be different categories of firms depending upon their preference for buying from the black market – (1) firms buying all the fuel in black market only; (2) firms buying fuel partially from the black market, and the rest through the formal channels; and (3) firms purchasing fuel only through formal channels.

To sum, we have following predictions on the impact of the UID-based transfer policy:

Prediction 1. Domestic Sector: Decrease in fuel purchase

Prediction 2. Commercial Sector: Increase in fuel purchase

Prediction 3. Black Market: Increase in price

Similarly, an opposite impact is expected when the policy is terminated.

Finally, it may concern the policymakers if increasing enforcement on corruption encourages micro-level evasion by the households. A household may choose to divert unused fuel when the annual cap is not binding to its expected total fuel consumption. A household's decision to engage in the black market depends on the expected cost and benefit. Increased enforcement on the ghost beneficiaries leads to a higher black-market price, which in turn, may tilt the decision in favor of diversion. Since household level subsidy diversion is still illegal, this increase in micro-level evasion is similar in spirit to the displacement in corruption, perhaps with redistribution consequences. Fig. 4 illustrates the expected effects in the black market.

4 Data and Identification

4.1 Data

4.1.1 Beneficiary Level Data

I use a transaction-level data from 3.79 million fuel subsidy beneficiaries. This sub-sample consists of 10% randomly selected beneficiaries from the HPCL (Hindustan Petroleum Corporation Limited) database of about 40 million beneficiaries.[43] This leads to a sample size of 23.2 million transactions in 509 districts in 25 states (Table 1). This dataset covers a 12 month period from April 2013 to March 2014 (i.e. one financial year) and includes complete LPG refill history for all the beneficiary accounts in the sample. The median household purchases seven LPG refills in a year ($\simeq$ 100kg LPG per year). The transaction-level data is rich in information. It provides the number of refills, the LPG refill order date, the LPG refill delivery date and information on compliance with the UID-based transfer program. I consolidate the data into a household-month level panel.

[43]HPCL is one of the three OMCs under the Ministry of Petroleum and Natural Gas and owns about 25% LPG market share. It is uniformly present in about 80% of the country.

4.1.2 Distributor-level Data

Distributor-month level LPG sales data covers 3,341 distributors in 504 districts (Table 1). The data segregates monthly LPG sales by sector, i.e., domestic fuel (14kg refills) and commercial fuel (19kg refills). On average, an LPG distributor sales around 6670 domestic fuel refills and 460 commercial-fuel refills in a month. The sample used in this study is from April 2013 to April 2014 (total 13 months). Note that 19kg LPG refills sales does not represent all of the commercial sector LPG demand. Small businesses and industries can buy LPG either in 19kg, 35kg, 47.5kg refills or in bulk. LPG is also distributed as transport fuel. Thus, I have commercial fuel sales information only from a subset of commercial sector (i.e. all 19kg LPG refill sales). The non-domestic sector sales is about 12% of the total LPG consumption in India in 2013-14.

4.1.3 Survey in the Black Market for LPG Fuel

I conducted fieldwork between December 2013 and March 2014 to collect black-market price information. Using two different survey instruments, data was gathered from: (1) the supply side, i.e., LPG delivery men, and (2) the demand side, i.e., small businesses. The survey instrument was carefully designed with pilot surveys and an understanding of the LPG black market in urban areas. These districts are randomly selected from the different phases. There are 10 districts in Phases 1 and 2 (i.e., treated districts), 50 districts in Phases 3 through 6 (i.e., transition districts), and 29 districts in the non-policy group. Until the date of termination, the UID requirement was fully enforced only in Phase 1 and 2 districts. Non-policy districts were never introduced to the policy. As it rightly appears, Phase 3 through 6 and non-policy districts are over-sampled. This is because policy-termination was not at all anticipated before the beginning of the field work. Initially, the fieldwork had a forward looking design to compare the pre- and post-treatment prices in the districts where the policy had not yet enforced. So, initial power calculations were based on a relatively mild impact expected on the fuel prices in the presence of an anticipated effect in the black market. Phase 1 and 2 were included right after the political announcement of policy termination, but about a month before the actual termination. Further, to ensure against any concerns about lag in surveying the treatment districts, an additional four control districts were surveyed simultaneously with Phase 1 and 2 districts.[44] In this study, data is used from the two survey rounds covering the period before and after the policy-termination (Table 2). As an advantage to this paper, the unexpected termination of the UID-based transfer policy made it possible to observe the sudden and much higher impact on prices.

Demand Side: Survey of the Small Businesses

In each district, 15 small businesses were surveyed in each round. A total of 1452 small businesses from 89 districts were surveyed. For sampling, first an area roaster was created after listing the main market areas, and up to three local areas were randomly selected. Next, small businesses in these local markets were listed and sampled. Due attention was given to ensure the inclusion of small businesses with a similar production function, such as snack counters and restaurants.[45] For example, more than 60% of the sampled small-businesses sell Samosa (a popular North-Indian snack) and chai (hot tea). About 20% attrition was observed in the post-termination round and replacement firms were surveyed from the same local markets. With a number of questions on LPG prices asked in different ways, the survey instrument was carefully designed to collect LPG refill price information without confronting the business-owner about his dealings in the black market. The survey collects data on ongoing black market prices as well as on LPG refill history.

Supply side: audit survey with LPG delivery men

Up to seven LPG delivery men were surveyed in each district in an audit format.[46] The same local market areas were surveyed as in the small business survey and on the same day. The enumerator approached every alternate delivery man. A delivery man, as part of his regular job, carries empty and filled LPG cylinders of both types (domestic 14kg and commercial 19kg), so he can be recognized easily. The enumerator asked for a quote for a 14kg LPG refill to be delivered on the same day. Specifically, as per the script provided, the enumerator told the delivery men that she did not have an LPG beneficiary card and wants to buy an LPG refill today *in black*.[47] One round of bargaining

[44]The fieldwork in 14 additional districts (i.e., 10 districts in Phases 1 and 2, and 4 districts in non-policy group) began within five days of the regular survey. The difference in the mean black-market price collected from the control districts surveyed in these two batches in the same wave is statistically insignificant. In post-policy termination round, all the districts in the sample are surveyed simultaneously.

[45]Metal-cutting shops, other small industrial firms and commercial taxi drivers may be other potential buyers on the black market.

[46]Enumerators could not find enough number of delivery men in all the areas.

[47]Inquiry to purchase an LPG cylinder "in black" or "without a beneficiary card" is readily interpreted as the potential black market transaction. In the pre- and post-termination round, the percentage of delivery men who refused to entertain a black market transaction inquiry was 18.2% and 4.2% respectively.

was also embedded in the script to elicit the true offer price for a black market LPG refill.

4.2 Empirical strategy

Using a difference-in-differences method, I exploit the phase-wise roll-out of the UID-based transfer policy and its unexpected termination in order to test the predictions developed in Section 3. The phase-wise roll out-provides a treatment variation in time and districts. Selection of districts in earlier phases is on the basis of initial UID penetration in these districts, which is assumed to be uncorrelated with subsidized fuel diversion. I also use multiple control groups to test for the robustness checks. Next, I exploit unexpected decision to terminate the policy in early 2014 to estimate causal effect of UID-based transfer policy termination. Post-termination, LPG subsidy transfer system was restored as it was before introduction of UID-based transfer. The assumption that the selection of districts into initial phases is not correlated to subsidy diversion, equally applies here. Equation 4 shows the main empirical specification.

$$Y_{idm}=\alpha + \beta.Post_m * UIDrequired_d + \gamma.Post_m + \delta.UIDrequired_d + \mu_i + \pi_m + \epsilon_{idm} \tag{4}$$

where i in district d in given month m. $Post_m$ is defined as the dummy for treatment period taking value 1 after UID-based transfer is enforced and 0 otherwise, $UIDrequired_d$ is dummy for the districts reflecting treatment status: districts in treated group has $UIDrequired_d = 1$ and 0 otherwise. μ_i is the dummy for each household controlling for household specific time invariant factors such as family size, education, wealth, eating preferences etc. ; π_m denotes month dummies controlling for month specific effects. ϵ_{idm} is household specific error term. With household-month level data, Y_{idm} is total number of LPG refills purchased by the household. The difference-in-differences coefficient β on $Post{*}UIDrequired_{dm}$ (i.e. an interaction of $UIDrequired_d$ and $Post_m$), provides average treatment effect of biometrics-based enforcement on domestic sector fuel purchase. I use OLS to estimate coefficients. Standard errors in all the regressions are clustered at the district level. In the preferred specification, Phase 1 districts are in the treated group, and are compared with the upcoming phases (where UID requirement was introduced but not yet enforced) and non-policy districts (where UID requirement was not yet introduced). This provides a longer treatment period of six months that reassures against any short term time-substitution in fuel purchases and delay in compliance. Impact of policy termination is estimated in a similar way by replacing $Post_m$

with $Post - termination_m$.

In order to show dynamic response in domestic sector fuel purchase, I estimate month-wise treatment effect in treated districts with the following specification.

$$Y_{idm}{=}\alpha + \beta.UIDrequired_d * \theta_m + \gamma.UIDrequired_d + \theta_m + \mu_i + \epsilon_{idm} \tag{5}$$

where β provides month-wise marginal effect on recorded domestic sector LPG refills. I present the results with month-wise estimated coefficient plot. These plots also enable comparison of pre-treatment and post policy termination trends in pre-treatment and post-termination period. Household fixed effects (μ_i) are included in above specification. Y_{idm} is total domestic LPG refills quantity purchased by the households in a given month and district. Further, I employ similar specification with distributor-month level panel. The outcome variable in *log(total LPG refills sales)* and I look at domestic and commercial fuel sales separately. Distributor fixed effects are included to control for distributor-level time invariant factors. Using the survey data, I estimate the impact of policy termination on black market fuel prices using similar specification (Equation 4). In the preferred specification, the outcome variable is *log(black-market price)*. With the supply side (delivery man survey) analysis, district level fixed effects are used, and with the demand side (firm-level survey) analysis, firm fixed effects are included.

5 Results

5.1 Impact of UID-based Transfer Policy Roll-out

5.1.1 Domestic Fuel Sector

Fig. 6 reports the estimated month-wise treatment effect of the UID-based transfer policy on fuel purchases in the domestic sector (Equation 5). Regarding the pre-treatment trends, this plot shows that beneficiary fuel purchase trends in treated and control districts are similar in the five months prior to the enforcement of the UID-based direct transfer policy. The estimates suggest the UID-based transfer policy enforcement caused a significant drop in the fuel purchase in the domestic sector. Table 3 shows the estimated average treatment effect using the same sub-sample as in Fig. 6. After controlling for household and month fixed effects, the estimated effect is an 11.2% to 13.8% decrease in domestic fuel purchase. The coefficient on the interaction term in Col (1) suggests a decrease of 0.0664 in monthly household LPG refills, which is about 11.8% of average

LPG refills per household per month.[48] The estimates are robust to using different control districts. Col (2) and (3) show estimates of the effect of policy separately with transition (Phases 3 through 6) and non-policy districts in the control group. In all the regressions, the treatment group includes districts from Phase 1.[49] Further, I confirm these estimates using the distributor-level panel dataset. In Fig. 7, the estimated coefficients suggest a similar impact of the UID-based transfer policy on fuel sales in the domestic sector. Difference-in-differences estimates show that the enforcement of UID-based transfer policy reduces fuel demand in domestic sector by 13% to 17% (Table 4). These estimates are slightly on the higher side, but are very close to the ones shown with the beneficiary-level panel above (Table 3). Overall, the results show that the UID-based transfer policy reduced the purchase of fuel in the domestic sector. This is in line with the first prediction and indicates that fuel purchase by the ghost beneficiaries decreased with the policy's enforcement. Note that in all of these regressions, our outcome variable includes all domestic LPG refills irrespective of the subsidy transfer. So, the fuel purchases by the households who did not comply or complied late, is taken into account. A large effect in the first month indicates a delay in compliance and timing to purchase the fuel, likely by the households that were late to comply. A lower UID penetration and lack of access to a bank account may further bias these estimates by completely excluding genuine beneficiaries.[50] I discuss the policy effect on late complier and non-complier households in Section 5.3 and 5.4.

5.1.2 Commercial Fuel Sector

A reduction in fuel sales in the domestic sector should accompany an increase in the commercial-fuel demand. However, the estimated effect of UID-based transfer on fuel sales in the commercial sector is not significant (Table 5). Although this is surprising, several potential explanations are possible. First, the program enforcement date was known for a long time and firms as well as agents active in the black market may have anticipated the effect well in advance. Fuel switching and up to certain extent, fuel stock-piling because of anticipated black market price changes can also be

[48] Mean of number of monthly refills per household is provided in the Table 3.

[49] Phase 2 districts are not included here because the UID-based transfer policy was enforced in Phase 2 districts during the treatment period for Phase 1. Results with Phase 2 districts are presented in (Table A13).

[50] Migrant households and fuel switching can also explain a part of the effect. It is not a direct concern if the non-complier or migrant household buys LPG in the black market, since all domestic refills are included in the outcome variable. A beneficiary can switch to kerosene, but that will require surrendering their LPG beneficiary card. Potential concerns may be about switching to coal or firewood, which cannot be completely ruled out.

not ruled out. Second, as already mentioned in the data section, the commercial LPG sales data do not cover all non-household LPG sales. The data used covers 19kg LPG refills only, whereas commercial supply is also provided in 35kg, 47.5kg, in bulk and also as transport fuel.

5.2 Policy Termination

5.2.1 Domestic Fuel Sector

Policy termination provides another opportunity to identify the impact of the UID-based transfer. The policy termination has already been shown with the coefficient plots (Fig. 6 and Fig. 7). With the distributor-level sales data, the covered time period includes one additional month (April 2014), providing a longer post-termination period. Note that Fig. 7 confirms that post-termination trends are similar in the treated and control districts. In addition, domestic sector fuel sales in treated districts revert to the level in the control districts, i.e., the difference between treated and control districts in terms of domestic fuel sales is insignificant after the UID requirement was relaxed.

Table 6 (Col (1) and (2)) reports about a 6 to 7.5% increase in the domestic fuel purchase due to the reversal in the UID-based transfer policy. The estimate in Col (3), where the control group consists of districts not yet planned for policy roll-out, is positive but not significant.[51] Similarly, with the distributor-level sales data, estimated coefficients suggest the symmetric impact of policy termination, when compared with the effect of policy enforcement (Table 7). About a 6% to 13% increase in distributor-level fuel sales in the domestic sector is observed when the UID-based transfer policy was reverted. Taken together with the estimated effect of UID-based transfer roll-out, we see the opposite impact on the fuel sales in domestic sector.

5.2.2 Commercial Fuel Sector

The termination of the UID-based transfer policy lowered fuel sales in the commercial sector by 6% to 9% (Table 8). Together with the results on fuel sales in the domestic sector, we see that the UID-based transfer policy termination, which was meant for the domestic sector only, affects fuel sales in the domestic as well as the commercial sector. While confirming prediction (2), this result indicates that the relaxation in enforcement in the domestic sector led to an increase in the diversion of subsidized fuel to the black market and fuel tax evasion in the commercial sector. Note

[51] This is likely because the post-termination dummy (termination occurred on March 10, 2014) does not fully align with the monthly structure of data, but I carry the same structure throughout this paper for consistency.

that, in terms of levels, the impact on fuel sales in the commercial sector is lower than the increased amount of domestic fuel sales, but the direction of impact is in the same direction as predicted.

5.2.3 Black Market Price

***Supply Side*: Delivery Men Audit Survey**

Using black market LPG refill data from the pre- and post-termination periods, Table 9 (Panel A) reports the estimated coefficients for difference-in-differences specification. The UID-based transfer policy termination reduced black-market price quotes from 13% to 16%. These estimates are also robust to more demanding time fixed effects as in Col(2). Since black market prices were higher during enforcement, it can be inferred that the impact on enforcement has had greater impact on pre-enforcement base prices. In other words, with respect to the pre-enforcement base price level, the impact of the UID-based transfer will be even higher than these estimates.

Demand Side: Small Businesses Survey

Similarly, Table 9 (Panel B) presents estimates using the data reported by small businesses. This provides us with slightly higher, but still closely aligned, estimates. Col (3) and (4) show that the policy termination in treated districts caused about a 20% decrease in black-market price. This jump in black-market price indicates a positive supply shock on the backdrop of the termination of the UID-based transfer policy. Note that the effect is slightly higher than the estimates with the supply side survey. A higher level of bargaining by long-time buyers after the sudden policy termination could be one potential explanation. Interestingly, the coefficients on the interaction and treatment terms sum up to zero, indicating that the policy termination led to a convergence in price levels in the treated and control districts. This is a significant result, considering that the treated and control districts had very similar trends in fuel sales in the pre-treatment period. Table A15 in the Appendix provides additional results with the transition districts in the control group. It confirms that the policy termination unambiguously brought black market fuel prices down.

In above analysis, firm and location fixed effects help in controlling for any consistent misreporting. By design, each round of survey was completed in a tight time window, usually within a week, to counter concerns about any time variation in prices. Next, the audit survey (supply side) and the small businesses survey (demand side) provide similar estimates, that help in ruling out any strategic misreporting by the respondents. The demand-side survey also included questions about LPG refill history that provided retrospective data on the date and price of the last five LPG refills. This helps in creating an unbalanced daily panel of the black-market prices, which I use to estimate the impact of the policy. Using a range of different treatment and control districts, Table A16 confirms that the estimated effect is robust to many different combinations of control and treatment districts. Unambiguously, the policy termination brought black-market prices down. However, effect size is smaller, when the treatment group consists of transition districts only.[52]

The refill history data provides a price time line that allows testing for any discontinuity in black market prices on the date of the policy termination. Fig. 14 shows the results with a 40 day window around the policy termination date. Note that the discontinuity in black market prices only shows up in the treated districts, and price levels in the transition and non-policy districts are stable around the policy termination date. This offers additional support for the main estimates.

5.3 Slow Compliance and Implementation Issues

While the UID-based transfer program has a significant effect on reducing subsidized fuel diversion into the black market, there are various factors that may affect the above estimates. First, implementation issues and delay in compliance may also affect the above estimates. This is a particular concern if some of the genuine beneficiaries, if they could not submit a UID or a bank account, are also eliminated. Compliance requires linking the UID to the LPG beneficiary account number and a bank account. Slow uptake can be attributed to UID and bank accounts coverage. However, by January 2014, UID penetration had covered 98.5% of the population in Phase 1 districts (Fig. A16). The 2011 census shows that bank account ownership dominates LPG adoption by households (Fig. A17). However, about 20% of households in the treated districts fulfilled the compliance requirement (i.e., submitting a UID and a bank account number) after Sep 2013 (Fig. A18). A household is more likely to fulfill the compliance requirement the next time an LPG refill is needed. Note that more than 90% of households do not need to purchase an LPG refill every month, since the median household consumes seven LPG refills per year. Further, Fig. 13 provides a reasonable assurance against the concerns of implementation issues pertaining to compliant households (such as, delay in bank transfers or UID reporting affecting purchase behavior of compliant beneficiaries in the treated districts) since purchase behavior of genuine households in treated districts is mostly

[52] This is likely because even before the UID-based transfer policy is enforced, black-market price starts increasing in the market.

parallel with the genuine households in Phase 6.

Second, the strategic timing of an LPG refill purchase can change pre- and post-treatment purchase. One possible channel is LPG storage. Households have limited, but some, storage capacity by design, since a 14kg LPG refill lasts almost two months for the median household.[53] Thus, the relatively large dip we observe in the first month of the UID-based transfer enforcement period (as in Fig. 6) may be partially attributed to an inter-temporal substitution in purchase decisions. This attenuates the estimates for the impact on fuel sales in the domestic sector. However, given a total period of six months, this should not significantly affect the magnitude of the average effect.

There are several pieces of supporting evidence to suggest that late compliance does not drive our main results. First, Fig. 6 and particularly, Fig. 9 show that the post-enforcement gap between the treated and control group remains mostly constant (except for the first and last month during the enforcement period), despite more than 20% of households complying from September 2013 to February 2014. In other words, the relatively constant LPG refill purchase gap between the treated and control districts should have otherwise gradually reduced if the delay in compliance played a significant role in bringing down the amount of LPG purchased in the domestic sector. Second, I directly test it by comparing the purchase behavior of late complier households in the treated districts (Phase 1) with the fuel purchase by households in the control districts (Phase 3 through 6). Fig. 10 shows that households that complied late did reduce LPG refill purchases right after the enforcement, but they more than compensated for it in subsequent months. Since the post-enforcement period covers all six months, there should be little effect on the estimates. Overall, this indicates that households that could not comply before UID-based transfer policy was enforced did not change their average fuel purchase behavior significantly during the enforcement period.

I compare the fuel purchases by early and late complier households in the treated districts. Fig. A20 shows that the late complier households gradually increased their LPG refill purchase. In fact, during the enforcement period, fuel purchase by late compliers reaches the fuel purchase level of early complier households, although in the pre-enforcement period average fuel purchase was lower in late complier sub-group. This suggests that, up to some extent, late complier households timed the decision to purchase fuel. Further, the decrease in fuel purchases contributed by late complier

[53] Having more than two LPG cylinders per household is not allowed. The government controls the supply of empty LPG cylinders as well.

households is small (i.e., the area between the treated and control plots in Fig. A20), when the proportion of late complier households is also take into account. At the most, it may explain a less than 4 percentage point decrease in LPG refill purchases, though additional robustness checks do not support this estimate.[54]

5.4 Non-Complier Households and Potential *Ghost Beneficiaries*

Less than 20% of beneficiaries in the treated districts did not comply by March 2014. With the descriptive data, Fig. 5, Panel (a) shows that beneficiaries who drew more refills in the pre-enforcement period are less likely to comply.[55] Fig. 5 further shows that the percentage compliance slows down at higher pre-treatment LPG refill counts. These non-compiler beneficiaries are also more likely to stop purchasing fuel. Fig. 11 reports that there was a huge drop in LPG refills purchased by non-compliant beneficiaries, when compared with complier households in Phase 3 through 6. On the other hand, we do not see such a drop in fuel purchases by late complier beneficiaries (households that complied after the program was enforced) in Fig. 10.

I further explore whether the count of beneficiaries purchasing zero refills responds to the introduction of the UID-based transfer policy. Table 11 reports that the enforcement of the UID-based transfer policy caused a 10% to 13% increase in the number of households that did not buy a single refill in a given month. This sample includes all the households, irrespective of their compliance status. Further investigating whether this effect is driven by low LPG refill purchases in the first enforcement month (or the last month before policy-termination), Fig. 9 reports that the treated districts observed more households that purchased no refills throughout during the enforcement period.

Exploring the take up rate in the treated districts, I plot monthly compliance conditional on pre-enforcement fuel purchase (Fig. A19). The plot shows that compliance rate responds to the enforcement, but overall compliance stays much lower in high fuel purchase sub-group. Note that the steepest compliance rate is shown by the households which did not purchase a single refill in pre-enforcement period. This likely represents new enrollments in the LPG subsidy program.[56] This

[54] As an additional robustness check, I compare late complier households in Phase 1 districts with complier households in Phase 6 districts (Fig. A21). Phase 6 provides an attractive control group for robustness check since it was introduced to the policy in the end.

[55] High-frequency beneficiaries are defined as the beneficiary who purchase five or more LPG refills, that is more than the prorated annual LPG refill cap at that time.

[56] Available data does not provide LPG subsidy beneficiary enrollment date.

can also be explained by timing to comply only when the household needs next LPG refill.

These empirical patterns highlight a couple of important points. First, we see different pre-treatment purchase behavior of the households that failed to comply with the UID requirement until the very end. Second, while late complier households do not show a significant reduction in LPG refills purchased during the UID-based transfer enforcement period, non-complier beneficiaries significantly reduced their fuel purchase in this period. The number of households who purchased zero-refills during the enforcement period also suggests a pattern that is expected from the ghost beneficiaries, who may opt to exit the black market when the UID-based transfer policy is enforced. Even though not all non-compliant beneficiary households can be categorized as ghost beneficiaries, it is likely that a good number of non-complier beneficiaries could not comply because they do not exist. However, using household level fuel purchase data, it is difficult to completely separate ghost beneficiaries from genuine non-compliant households.

5.5 Heterogeneous Effects

In order to see whether the UID-based transfer policy affected various subgroups in different ways, I use pre-treatment purchase behavior and the compliance-status of the beneficiaries to analyze heterogeneous treatment effects. This endogenous analysis is obviously limited in the sense that I do not have socio-economic data on the households. It is possible that a large number of these LPG beneficiary cards were held by ghost beneficiaries, who drew a higher number of LPG refills in the pre-enforcement period.[57]

I explore the impact of UID-based transfer on high and low LPG purchase frequency subgroups using a triple difference design. A prorated annual LPG cap cutoff (i.e., five or more than five refills) is used to divide households into two sub-groups as per their pre-treatment purchase behavior: (1) High-frequency beneficiary), and (2) Low-frequency beneficiary. Similarly, using the information on household compliance date, we have another division (1) Compliant beneficiaries, and (2) Non-compliant beneficiaries as per their status in the first month of enforcement. Results are shown in Fig. 8. The triple difference coefficient shows more than 30% impact on the high-frequency non-compliant households, which is more than twice as large as the average treatment effect estimated in

[57] A genuine beneficiary who prefers not to comply (or, cannot comply) would likely try to use the entire fuel quota before the UID-based transfer policy came into force, but households have limited capacity to store LPG because of the cap on empty cylinders.

Section 5.1.1. It suggests that high-frequency non-compliant households are more likely to decrease their LPG purchase, once the UID-based transfer policy was enforced (Table 10).

5.6 Household Level Diversion

It is of policy interest to see whether the increase in enforcement on ghost beneficiaries causes an increase in subsidy diversion by genuine households. High black-market price (when the UID-based transfer policy was enforced) and the non-binding annual LPG refill cap can provide lucrative incentives to trade on the black market. I empirically test whether the UID-based transfer policy induces additional households to divert domestic fuel to the black market.

Fig. 12 shows that there is indeed a steep jump in LPG purchases by early complier households, when UID-based transfer is enforced. An early complier household is the one who complied by the time the policy is enforced. This single difference plot reports that early complier households increased their fuel purchase substantially when UID-based transfer policy was enforced. In order to elicit causal estimates, I further use compliant households in yet-to-be treated districts (Phase 3 through 6) as control. Table 12 shows the results. Enforcing the UID-based transfer caused a 4% increase in LPG refill purchases by the genuine households in treated districts (Col (1)). While this is a positive effect, when compared to the overall impact of the UID-based transfer policy on the supply in the black market, the size of this effect is small.

As a robustness check, to allay concerns about any effect of the UID-based transfer policy introduction on the households in the control group, I test the same specification only with households in Phase 6 as control. Phase 6 districts entered the transition period in January 2014, thus the first four months in the Phase 1 treatment period (September to December 2013) are relatively less prone to any concern about changes in purchase behavior during the transition period. Fig. 13 provides little support in favor of an increase in the subsidy diversion by households, as reported in Col (2) in Table 12. There are two potential explanations. First, households may face a higher legal and logistic cost when they try to sell an LPG refill on the black market. Second, social norms may make it costly for a household to participate in the black market. Overall, household-level diversion does not seem to outweigh the reduction in household fuel sales after increased enforcement.

5.7 Summary

To summarize, the results show that the UID-based subsidy transfer policy reduced fuel purchases in the domestic sector. This confirms Prediction (1) and indicates a reduction in subsidy diversion through the ghost beneficiary accounts. Analysis of commercial fuel demand did not show a significant positive effect during the enforcement period, but the impact of unexpected policy termination confirms Prediction (2). Black-market price decreased when the UID-based transfer policy was terminated, which is in line with Prediction (3).

I also investigated several other strands. First, it is evident that the beneficiaries who bought more subsidized fuel before the enforcement are less likely to comply with the UID requirement. I showed that these beneficiaries are also more likely to change their purchase behavior after the increased enforcement. This is consistent with the story of subsidy diversion through the ghost beneficiaries that are unable to procure a UID. Second, I show that results are not driven by late compliance or under-provision to compliant households. Finally, I test whether the increased enforcement affected subsidy diversion by compliant households in treated districts and, at least in the short run, it does not seem to be high enough to outweigh the reduction in subsidy diversion by the ghost beneficiaries. Low take up by genuine beneficiaries remains a potential concern in this paper, which could not be sufficiently addressed with the available data.

The main estimate on reduction in fuel purchases (i.e., 11% to 14%) in the domestic sector is less than the *subsidy gap* shown in Fig. 1. However, it is not far from other governments audits. For example, Karnataka state carried out rigorous verification of the LPG subsidy beneficiary list in 2012 and found about 22% illegal LPG beneficiary cards (Sastry, 2012). Note that a certain degree of subsidy diversion could still continue under the UID-based transfer policy, since a lower but significant margin to make profit remained available. A part of the *subsidy gap* can also be explained with a pre-existing level of subsidy diversion by compliant households.

6 Policy Discussion and Conclusion

This paper underscores the importance of investing in state's enforcement capacity. When reducing subsidy diversion and curtailing the black markets using traditional enforcement methods (e.g., audit and penalty) is not feasible, it may be more effective to invest in enforcement infrastructure such as a UID-based transfer platform. More than eighty developing countries may soon employ biometrics-based identification in some way or another (Gelb and Clark, 2013a). Breckenridge (2014) mentions that the success or failure of India's UID system will determine whether other developing countries will subscribe to biometrics. Using the UID's application to a major subsidy program in India, this paper presents an array of empirical evidences from formal and informal sectors to illustrate that there are significant efficiency gains of increasing enforcement in developing countries. With a reduced subsidy burden, there are direct benefits in terms of more funds for other public programs and a reduction in taxes. Black-market activities are also a waste of resources, which could be put to more productive use. In addition, since reduced evasion changes the ultimate fuel price that firms face, an increase in fuel efficiency may also have its own environmental and macro-economic benefits. Moreover, there is an effect expected in terms of decreasing informality and cash in the underground economy. Gordon and Li (2009) discuss how, in developing countries, informal firms decide to stay in the cash economy by weighing the potential cost and benefits of using the financial sector. Fuel purchases in the black-markets play a similar role by allowing firms to avoid leaving an information trail of their fuel consumption, which in turn can be linked to the size of their businesses by tax inspectors.

A complete welfare analysis would require the analysis of general equilibrium effects, for which no attempt is made here. Increasing enforcement on fuel subsidies may increase the general price levels. Through increase in price levels, there could also be a heterogeneous impact on firms who can find a way to manipulate the system. Second, there are implications for redistribution if poor households benefit from diverting subsidized fuel. Again, heterogeneity in household's ability to make profits on the black market would matter.

Tax and transfer administration reforms are important for developing countries. Yet, the enormous investment required often makes it difficult for policymakers to commit to new enforcement systems. India's UID-based direct transfer policy rides on heavy investment in the Unique Identification system and its cost and expected benefits are widely debated. Since the UID program is also linked to many other welfare programs, it is difficult to gauge its net benefits in this paper. However, at the expected cost of USD 3 for each UID issued (Gelb and Clark, 2013b), it is likely that, in the long term, expected efficiency gains in tax and transfer programs will provide a good return on the government's investment in the UID system.

On a broader level, reforms to combat corruption and evasion need much stronger political support and are susceptible to political business cycles. India's unexpected termination of the UID-based fuel subsidy transfer program also illustrates that as long as political acceptance remains insufficient, technology alone cannot help. Casaburi and Troiano (2013) show that incumbent politicians may enjoy higher likelihood of re-election because of wider approval with an anti-tax evasion policy. However, such potential gains may not be visible to the politicians themselves ex ante.

References

Acemoglu, D. (2005). Politics and economics in weak and strong states. *Journal of Monetary Economics*, 52(7):1199–1226.

Aji, S. (2012). Stink of a Gas Scam: Over 24 lakh of the 79 lakh LPG Connections in Karnataka belong to 'Ghost' Consumers. India Today.

Aker, J. C., Boumnijel, R., McClelland, A., and Tierney, N. (2014). Payment Mechanisms and Anti-Poverty Programs: Evidence from a Mobile Money Cash Transfer Experiment in Niger. *Unpublished working paper.*

Alesina, A. (1997). *Political cycles and the macroeconomy.* MIT press.

Alm, J., Deskins, J., and McKee, M. (2009). Do individuals comply on income not reported by their employer? *Public Finance Review*, 37(2):120–141.

Arze del Granado, F. J., Coady, D., and Gillingham, R. (2012). The unequal benefits of fuel subsidies: A review of evidence for developing countries. *World Development*, 40(11):2234–2248.

Banerjee, A., Mullainathan, S., and Hanna, R. (2012). Corruption. *National Bureau of Economic Research*, (Working Paper 17968).

Besley, T., Ilzetzki, E., and Persson, T. (2013). Weak states and steady states: The dynamics of fiscal capacity. *American Economic Journal: Macroeconomics*, 5(4):205–235.

Besley, T. and Persson, T. (2009). The origins of state capacity: Property rights, taxation and policy. *The American Economic Review*, 99(4):1218–1244.

Besley, T. and Persson, T. (2010). State capacity, conflict, and development. *Econometrica*, 78(1):1–34.

Besley, T. J. and Persson, T. (2013). *Handbook of Public Economics: Taxation and Development*, volume 5. Prepared for A. Auerbach, R. Chetty, M. Feldstein and E. Saez.

Bird, R. M. and Zolt, E. M. (2008). Technology and Taxation in Developing Countries: from Hand to Mouse. *National Tax Journal*, pages 791–821.

Breckenridge, K. (2014). *Biometric State: the Global Politics of Identification and Surveillance in South Africa, 1850 to the Present.* Cambridge University Press.

Buehn, A. and Schneider, F. (2013). Estimating the size of the shadow economy: Methods, problems and open questions. Technical report.

Burgess, R., Hansen, M., Olken, B. A., Potapov, P., and Sieber, S. (2012). The political economy of deforestation in the tropics. *The Quarterly Journal of Economics*, 127(4):1707–1754.

Bussell, J. L. (2010). Why get technical? corruption and the politics of public service reform in the indian states. *Comparative Political Studies.*

Carrillo, P., Pomeranz, D., and Singhal, M. (2014). Dodging the taxman: Firm misreporting and limits to tax enforcement. *National Bureau of Economic Research*, (Working Paper 20624).

Casaburi, L. and Troiano, U. (2013). Ghost-house busters: The electoral response to a large anti tax evasion program.

Cole, S. (2009). Fixing market failures or fixing elections? Agricultural credit in India. *American Economic Journal: Applied Economics*, pages 219–250.

Eggen, O. (2012). Performing Good Governance: The Aesthetics of Bureaucratic Practice in Malawi. *Ethnos*, 77(1):1–23.

Emran, M. S. and Stiglitz, J. E. (2005). On selective indirect tax reform in developing countries. *Journal of Public Economics*, 89(4):599–623.

Fisman, R. and Wei, S.-J. (2001). Tax rates and tax evasion: Evidence from "missing imports" in China. Technical report, National bureau of economic research.

Fisman, R. and Wei, S.-J. (2009). The smuggling of art, and the art of smuggling: Uncovering the illicit trade in cultural property and antiques. *American Economic Journal: Applied Economics*, 1(3):82–96.

Foremny, D. and Riedel, N. (2014). Business taxes and the electoral cycle. *Journal of Public Economics*, 115:48–61.

Gelb, A. and Clark, J. (2013a). Identification for development: The biometrics revolution. *Center for Global Development, Washington DC, accessed April*, 16:2013.

Gelb, A. and Clark, J. (2013b). Performance lessons from India's universal identification program. *CGD Policy Paper*, 20.

Giné, X., Goldberg, J., and Yang, D. (2012). Credit market consequences of improved personal identification: Field experimental evidence from malawi. *American Economic Review*, 102(6):2923–2954.

Gordon, R. and Li, W. (2009). Tax structures in developing countries: Many puzzles and a possible explanation. *Journal of Public Economics*, 93(7):855–866.

International Institute of Sustainable Development (2014). Indonesia Energy Subsidy Review. `http://www.iisd.org/gsi/sites/default/files/ffs_indonesia_review_i1v1.pdf`.

Jack, W. and Suri, T. (2014). Risk sharing and transactions costs: Evidence from Kenya's mobile money revolution. *The American Economic Review*, 104(1):183–223.

Jai, S. (2013). Distributors want Licence to sell Cylinders in Black. The Economic Times.

Kaiman, J. (2013). Chinese fury as ID Fraud becomes Recurring Motif in Property Scandals.

Kapur, D. and Vaishnav, M. (2011). Quid Pro Quo: Builders, Politicians, and Election Finance in India. *Center for Global Development Working Paper*, 276.

Khera, R. (2011). The UID project and welfare schemes. *Economic & Political Weekly*, 46(9):38–44.

Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S., and Saez, E. (2011). Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica*, 79(3):651–692.

Kleven, H. J., Kreiner, C. T., and Saez, E. (2009). Why can modern governments tax so much? An agency model of firms as fiscal intermediaries. *National Bureau of Economic Research*, (Working Paper 15218).

Kopczuk, W., Marion, J., Muehlegger, E., and Slemrod, J. (2013). Do the Laws of Tax Incidence Hold? Point of Collection and the Pass-through of State Diesel Taxes. *National Bureau of Economic Research*, (Working Paper 19410).

Kopczuk, W. and Slemrod, J. (2006). Putting firms into optimal tax theory. *The American economic review*, pages 130–134.

Kumar, M. (2012). Cap on LPG Fuels Nationwide Rage; Distributors Call for Strike on Oct 1. International Business Times.

Kumler, T., Verhoogen, E., and Frías, J. A. (2013). Enlisting Employees in Improving Payroll-Tax Compliance: Evidence from Mexico. *National Bureau of Economic Research*, (Working Paper 19385).

Marion, J. and Muehlegger, E. (2008). Measuring Illegal Activity and the Effects of Regulatory Innovation: Tax Evasion and the Dyeing of Untaxed Diesel. *Journal of Political Economy*, 116(4):pp. 633–666.

Mayah, E. (2012). 73,000 Ghost Accounts used in Nigeria Pension Scam. Africa Review.

Mbiti, I. and Weil, D. N. (2011). Mobile banking: The impact of M-Pesa in Kenya. (Working Paper 17129).

McLure Jr, C. E. (2013). Reforming Subsidies for Fossil Fuel Consumption: Killing Several Birds with One Stone. Technical report.

Merriman, D. (2010). The micro-geography of tax avoidance: evidence from littered cigarette packs in Chicago. *American Economic Journal: Economic Policy*, 2(2):61–84.

Ministry of Petroleum and Natural Gas (2014). Marketing Discipline Guidelines.

Ministry of Petroleum and Natural Gas, India (2014),). Report on Marketing Activities. `http://petroleum.nic.in/docs/mktact.pdf`.

Muralidharan, K., Niehaus, P., and Sukhtankar, S. (2014). Building State Capacity: Evidence from Biometric Smartcards in India. *National Bureau of Economic Research*, (Working Paper 19999).

Nambiar, N. (2014). LPG Subsidy: Direct Benefit Transfer System Stands Scrapped from Mar 10. The Times of India.

Niehaus, P. and Sukhtankar, S. (2013a). Corruption dynamics: The golden goose effect. *American Economic Journal: Economic Policy*, 5(4):230–269.

Niehaus, P. and Sukhtankar, S. (2013b). The marginal rate of corruption in public programs: Evidence from India. *Journal of Public Economics*, 104:52–64.

Nordhaus, W. D. (1975). The political business cycle. *The Review of Economic Studies*, pages 169–190.

Olken, B. A. (2006). Corruption and the costs of redistribution: Micro evidence from Indonesia. *Journal of Public Economics*, 90(4):853–870.

Olken, B. A. and Pande, R. (2012). Corruption in Developing Countries. *Annual Review of Economics*, 4(1):479–509.

Parliament of Zimbabwe (2012). Proceeding on Feb 28, 2012.

Pissarides, C. A. and Weber, G. (1989). An expenditure-based estimate of Britain's black economy. *Journal of Public Economics*, 39(1):17–32.

Pomeranz, D. (2013). No Taxation Without Information: Deterrence and Self-Enforcement in the Value Added Tax. *National Bureau of Economic Research*, (Working Paper 19199).

PTI (2014). Jan Dhan Scheme: RBI Governor Raghuram Rajan Cautions Banks Against Running after Numbers. The Economic Times.

Registrar General of India (2011). Houselisting and Housing Census Data - 2011. `http://www.censusindia.gov.in/2011census/hlo/HLO_Tables.html`.

Reinikka, R. and Svensson, J. (2001). Explaining Leakage of Public Funds. *World Bank Policy Research Working Paper*, (2709).

Rodrik, D. (2008). Second-Best Institutions. *American Economic Review*, 98(2):100–104.

Saez, E., Slemrod, J., and Giertz, S. H. (2012). The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of Economic Literature*, 50(1):3–50.

Sastry, A. K. (2012). Illegal LPG Connections: Centre Cold to Karnataka Model. The Hindu.

Singh, S. (2014). 7.77 lakh Fake Ration Cards Uncovered in Bihar, Govt Orders Probe. The Indian Express.

Slemrod, J. (2007). Cheating ourselves: the economics of tax evasion. *The journal of economic perspectives*, pages 25–48.

Slemrod, J. and Weber, C. (2012). Evidence of the invisible: toward a credibility revolution in the empirical analysis of tax evasion and the informal economy. *International Tax and Public Finance*, 19(1):25–53.

Slemrod, J. and Yitzhaki, S. (2002). Tax avoidance, evasion, and administration. *Handbook of Public Economics*, 3:1423–1470.

Smith, D. A. (2002). Consolidating Democracy? The Structural Underpinnings of Ghana's 2000 Elections. *The Journal of Modern African Studies*, 40(04):621–650.

Sukhtankar, S. (2012). Sweetening the Deal? Political Connections and Sugar Mills in India. *American Economic Journal: Applied Economics*, 4(3):43–63.

Szreter, S. (2007). The right of registration: development, identity registration, and social security – a historical perspective. *World Development*, 35(1):67–86.

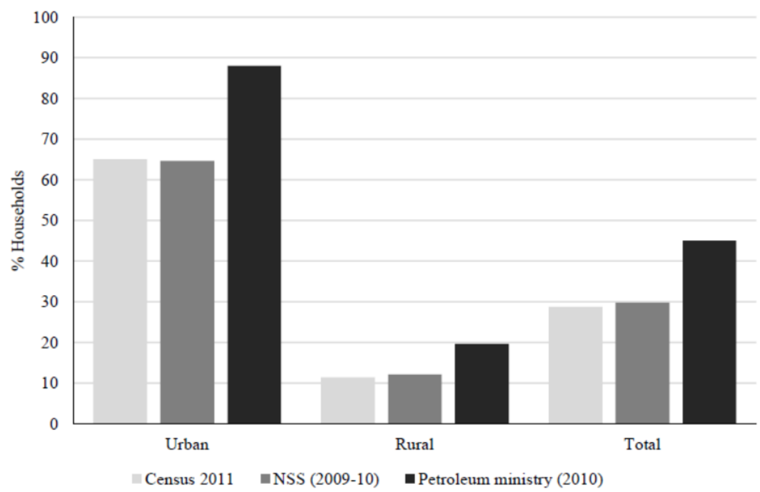
Tangri, R. and Mwenda, A. M. (2008). Elite Corruption and Politics in Uganda. *Commonwealth & Comparative Politics*, 46(2):177–194.

TNN (2013). 'Aadhaar-hit' Gas Delivery Boys begin Strike. Frontline.

TNN (2014). Aadhaar Link needed to Control Subsidy Leakage: Moily. The Times of India.

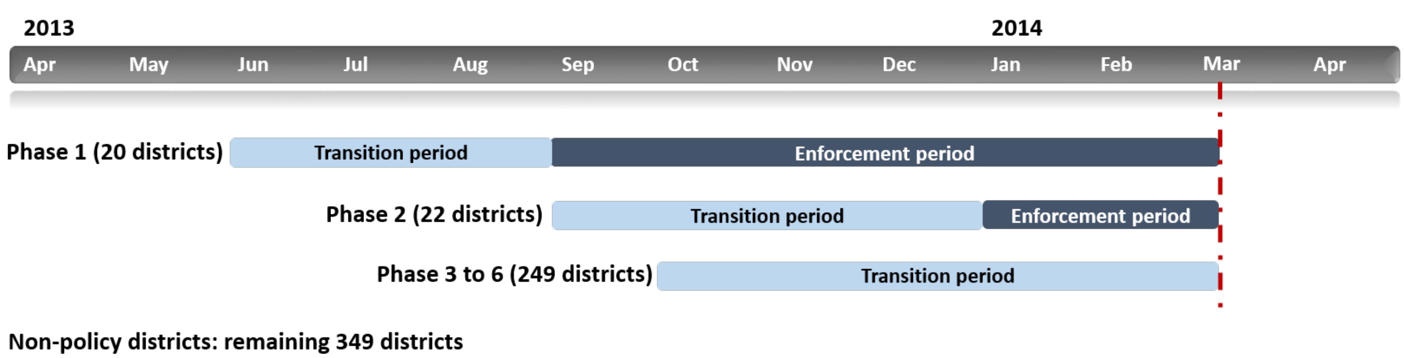
Yang, D. (2008). Can enforcement backfire? Crime displacement in the context of customs reform in the Philippines. *The Review of Economics and Statistics*, 90(1):1–14.

Figure 1: Subsidy gap: Cooking fuel subsidy in India



Note: Above plot illustrates the significant gap between the number of LPG user households in data collected from different sources. 2011 Census shows about 30% households use LPG for cooking purpose, which is also confirmed by the National Sample Survey (Round 66). On the other hand, the Ministry of Petroleum and Natural Gas reports about 45% LPG user households. All three sources consider almost similar base number of total households ($\sim$ 250 million) in 2011. Since LPG subsidy is universal for households, this gap directly corresponds to a *Subsidy gap*. While LPG and Kerosene are the preferred urban fuel; dirty fuels (such as firewood, kerosene, cow-dung cakes, coal) are used by households in rural areas. A household is allowed subsidies either on Kerosene or on LPG for cooking purpose, but not on both at the same time. LPG users predominantly belong to higher income groups.

Figure 2: Timeline: The UID-based transfer policy roll-out and termination

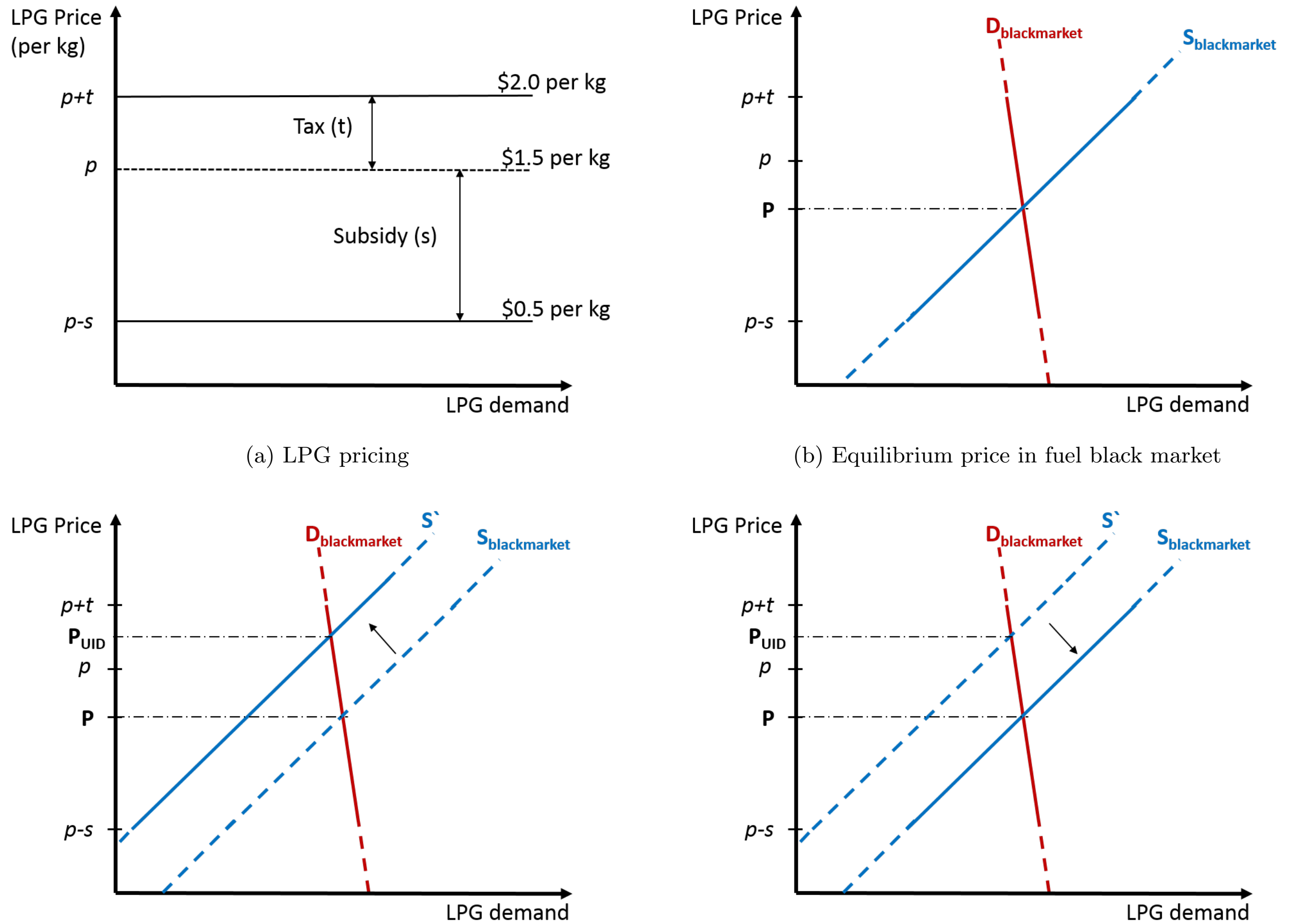


Note: **UID-based Transfer Roll-out:** Above timeline shows variation in the time of policy enforcement. Phase 1 districts were introduced with the policy in June 2013. In the transition period, a household who has complied (i.e. submitted UID and bank account details), received fuel subsidy directly in the bank account. A non-compliant household (as well as *ghost beneficiaries*) continue to purchase LPG refills at subsidy-inclusive price. Once transition period is over, the UID-based transfer is fully enforced in Phase 1 districts starting from 1 September 2013. All households avail LPG only at the non-subsidized price, whereas only compliant households get subsidy amount transferred to their bank accounts. This policy was gradually introduced in five next phases covering almost half of country. By January 2014, the policy was introduced in 291 districts, and 42 districts had it fully enforced.

Policy Termination: The UID-based transfer policy was terminated starting from March 10, 2014. A political announcement was made on January 31, 2014 in this regard, but the actual termination occurred after about 40 days. Policy termination effectively brought the subsidy delivery mechanism back to the old system.

Data Coverage: Administrative data on household-level fuel purchases covers whole 12 months period (April 2013 – March 2014). Distributor level LPG sales data for domestic and commercia fuel is month-wise and covers 13 months period (April 2013 – April 2014). Black-market survey data collection was done before and after the policy termination.

Figure 3: Impact of UID-based Transfer on Fuel Black market



(a) LPG pricing

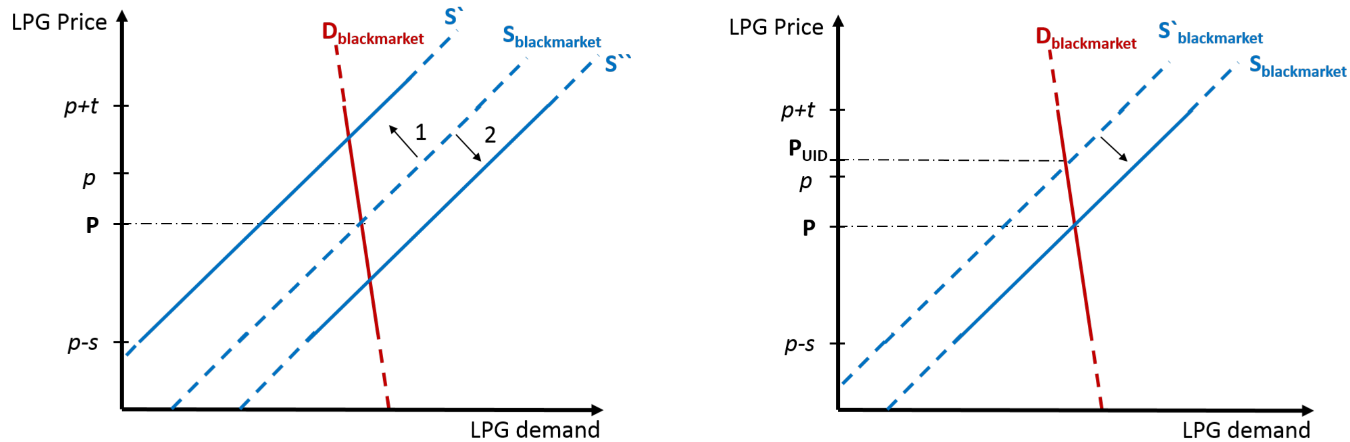
(b) Equilibrium price in fuel black market

(c) UID-based transfer: Negative supply shock

(d) UID-based transfer terminated: Positive supply shock

Note:
Panel (a) shows different price slabs as per a pricing structure regulated by the government. Households face subsidized fuel price $(p - s)$, whereas commercial firms purchase fuel at a tax inclusive price $(p + t)$. Households which exceed their annual cap on LPG refills, can buy LPG at non-subsidized prices (p), that is exogenously determined. January 2014 price levels are shown (approximate). Since supply is not constrained, $p - s$ and $p + t$ also represent formal sector supply curve for the domestic and commercial consumers.
Panel (b) presents the supply-demand and the equilibrium price in the fuel black market. The difference in prices provides arbitrage opportunity. Price floor and ceiling automatically become effective because of the market segmentation. Selling or purchasing LPG from non-authorized sources (i.e. non-PDS sources) is illegal, so these transactions take place in the black market.
Panel (c) shows impact of the UID-based transfer enforcement on the black market supply. LPG subsidy s is transferred directly to households after they submit the UID number. Thus, UID-based transfer facilitates a de-duplication process where *ghost beneficiaries* do not receive fuel subsidy anymore. However, any LPG beneficiary (that includes late and non-compliant households as well as *ghost beneficiaries* can purchase LPG at the non-subsidized price p. Thus enforcement with the UID-based transfer policy causes a supply shock in the black market.
Panel (d) explains the expected impact of policy termination. When the UID-based transfer policy for subsidy transfer is terminated, old system is restored (as in Panel (b)). A positive supply-shock is expected in fuel black market because now *ghost beneficiaries* are re-allowed to buy LPG at subsidized price $p - s$.

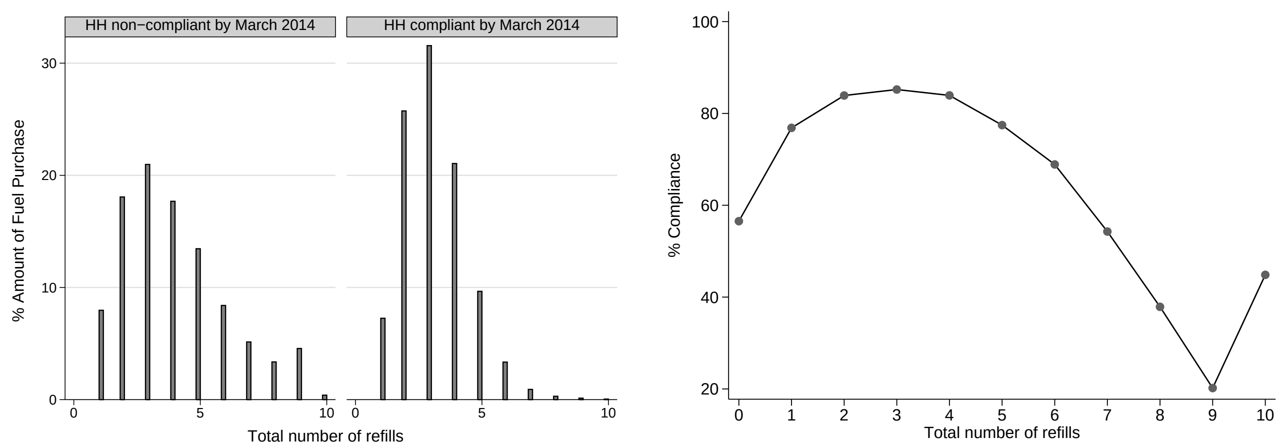
Figure 4: Increased enforcement and household-level diversion



(a) Enforcement on *ghost Beneficiaries* vs. Diversion by households (b) Impact of enforcement policy termination

Note:**Panel(a)** illustrates two opposite effects of the UID-based transfer policy. First, the elimination of *ghost beneficiaries* causes a negative supply shock pushing the black-market price up (1). Next, since the payoff in the black-market is higher now, households may increase the supply of subsidized-fuel from their personal quota. Thus figure represents the trade-off between two factors – reducing corruption Vs. increasing micro-level evasion, when enforcement is targeted only at the first factor. The net effect on supply and equilibrium price in the black market will depend on the magnitude of these two factors.
Panel(b) highlights the case when *ghost beneficiaries* are the dominant factor in the black market. The policy termination should move the black-market price down to the equilibrium level in control districts.

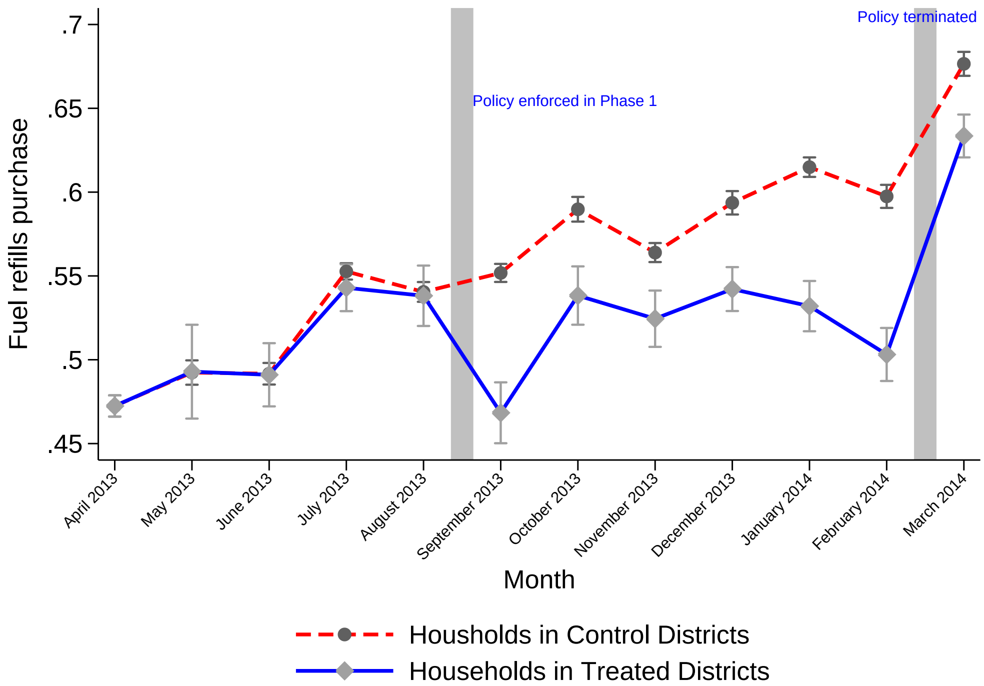
Figure 5: Pre-treatment fuel purchase and household's compliance status



(a) Fuel purchase by compliance (b) Compliance percentage by LPG Refills purchased

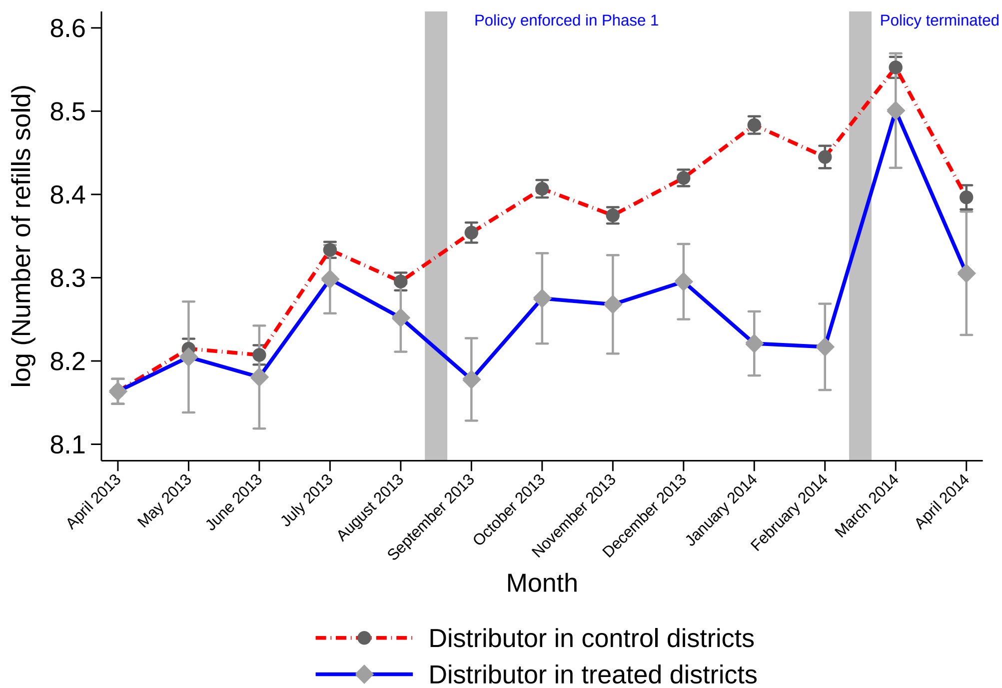
Note: The above plots show the distribution of pre-treatment LPG refill purchases and household compliance in the UID-based transfer policy. Y-axis in the Subplot (a) reports volume of fuel bought by households in pre-treatment period. Subplot (b) shows % compliant households (at distributor level) over total number of LPG refills purchased during pre-treatment period in Phase 1 districts. It is evident in the histogram that the beneficiaries who purchased beyond their prorated annual LPG refill cap (i.e. 5 or more), are less likely to comply. In 2013, annual LPG refill cap was 9, which is clearly reflected on the curve in lower panel. Compliant households are defined considering their compliance status by March 2014. Note that higher number of LPG refills are bunched together at 10 in both subplots.

Figure 6: Fuel purchase in the domestic sector (Beneficiary level data)



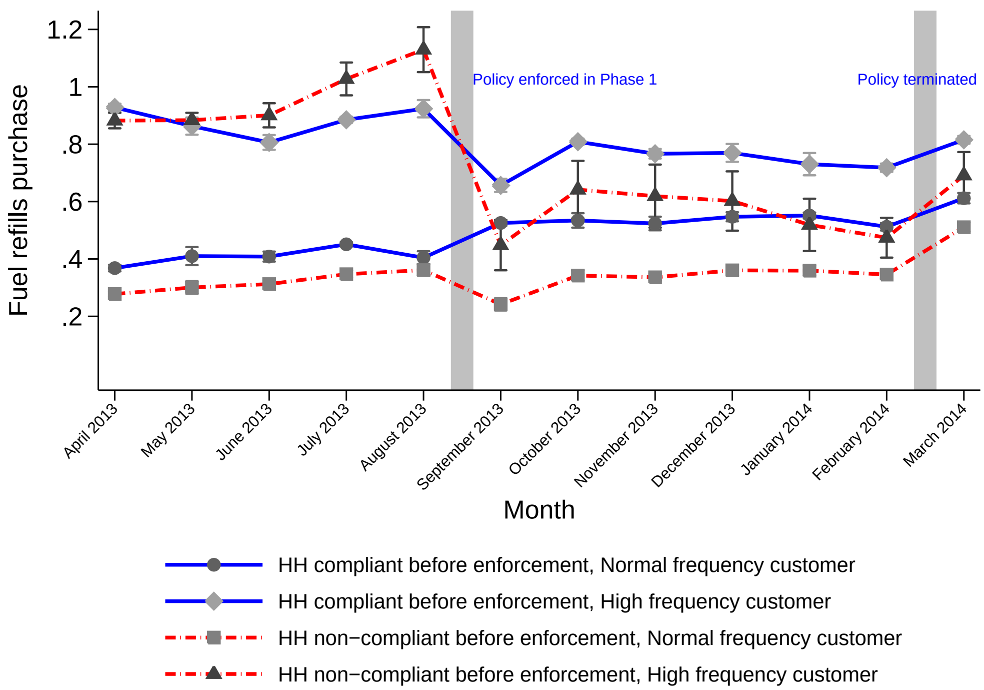
Note: The above plot shows the impact of UID-based transfer policy on household LPG refill purchases using a household-month panel dataset. This plot reports month-wise treatment effect (Equation 5) after controlling for household fixed effects. Dependent variable is the number of 14kg domestic LPG refills (subsidized as well as non-subsidized) by a household in a given month. Domestic fuel purchase by beneficiaries drops down right after the policy is enforced and it increases again when the policy is terminated. Confidence intervals are shown on this plot but they are extremely tight. LPG refills transaction data from 3,481,298 households is used covering Apr 2013 to Mar 2014. This is 10% randomly drawn sample from administrative records. There are total 16 Phase 1 districts in the treatment group and 473 districts in the control group (i.e. Phase 3 to Phase 6 and remaining non-policy districts). Policy was terminated on March 10, 2014. Phase 2 districts are excluded for clarity because the UID-based transfer policy was enforced in January 2014.

Figure 7: Fuel sales in the domestic sector (Distributor level data)



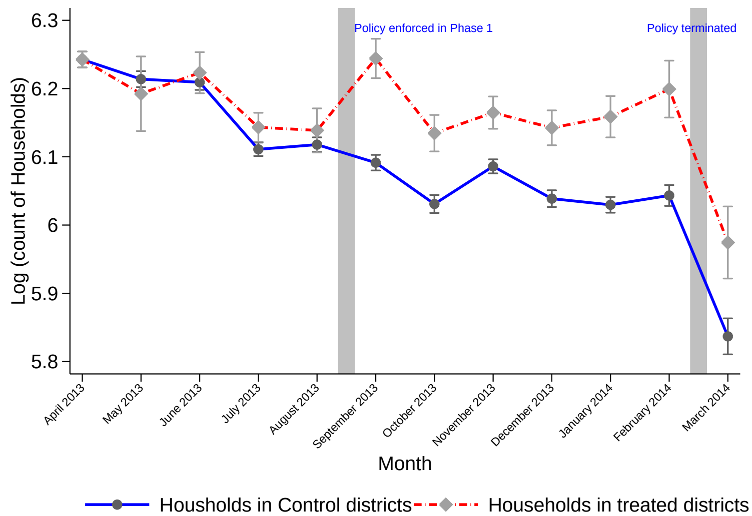
Note: The above plot shows the impact of UID-based transfer policy on fuel sales in domestic sector using a distributor-month panel dataset. It confirms that the domestic fuel sales decreased during the enforcement period and reverted back when the policy is terminated. This plot reports month-wise treatment effect (Equation 5 after controlling for household fixed effects. Outcome variable is 'log of total number of 14kg domestic LPG refills' (subsidized as well as non-subsidized). This sample covers data from 3013 distributors in 485 districts. Treatment group consists of 15 districts from Phase 1, whereas control groups has all the districts from Phase 3-6 and non-policy districts. Phase 2 districts are excluded. Policy was terminated on March 10, 2014.

Figure 8: Heterogeneous effect in treated districts



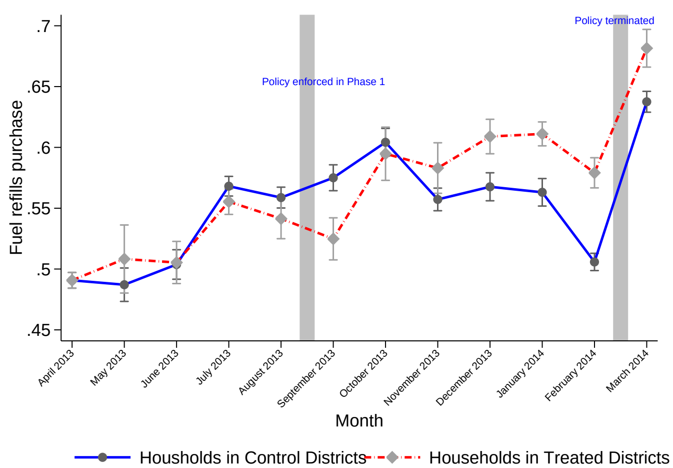
Note: This graph shows the impact of UID-based transfer policy estimated with a triple difference empirical model for beneficiaries in Phase 1 districts. Non-complier beneficiaries who purchased higher number of LPG refills in pre-treatment period, exhibit higher impact of the enforcement. High and normal frequency beneficiaries are defined as per their fuel purchase in the pre-treatment period. 'Normal frequency customer' denotes the household who has bought LPG refills as per its prorated annual cap (i.e. up to 4 LPG refills) and 'High frequency customer' is the household which bought five or more. We see that pre-treatment high frequency customers, who failed to comply by the first month of enforcement, significantly reduce their LPG refills purchase. Bunching i.e. more LPG purchases in anticipation before the UID-based transfer policy is enforced, is also observed. Triple difference estimates are presented in Table 10. District level fixed effects are used.

Figure 9: Number of beneficiaries who purchased no fuel



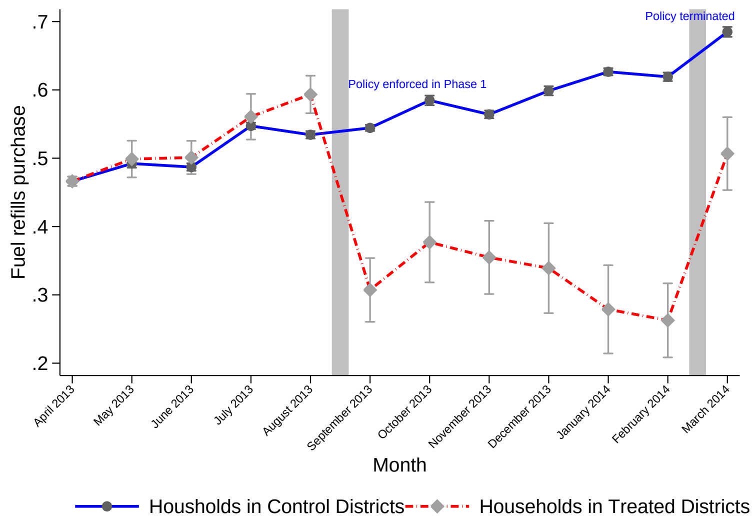
Note: This plot shows the impact of UID-based transfer on the number of zero-refill beneficiaries using a distributor-month level panel dataset. Zero-refill beneficiary is defined as the beneficiary who does not purchase any LPG refill in a given month. Outcome variable is log(number of zero-refill households). Treated group includes distributors in Phase 1 districts, whereas control group includes distributors in Phase 3 to 6. The panel is created by collapsing household-month panel. Distributor fixed effects are included. This plot suggests significant increase in the number of households who bought no refills during the UID-based transfer policy enforcement period.

Figure 10: Late complier households



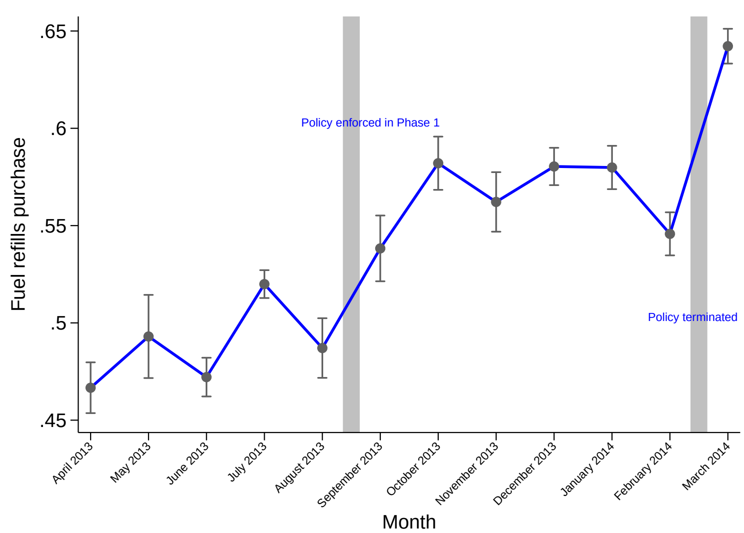
Note: This graph investigates the purchase behavior of late complier households in the treated districts with respect to complier households in upcoming phases. Estimated monthly coefficients show that the late complier households in Phase 1 districts increased their LPG purchases after first month in the enforcement period. This suggests delaying in the compliance may be governed by household's need for the next LPG refill. Also, relative size of coefficients, when averaged over the enforcement period, indicates that the late complier households did not change their purchase behavior much and so, late compliance does not seem to drive our main results. Note that complier households in districts in upcoming phases are included as control here. These households fulfill the compliance requirement, but the UID-based transfer policy was not yet enforced.

Figure 11: Non-compliant beneficiaries



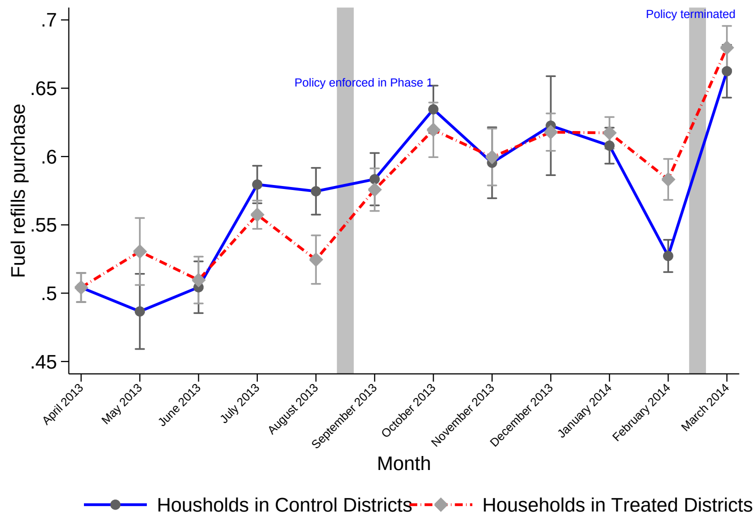
Note: Above plot shows the purchase behavior of non-compliant beneficiaries in treated districts when compared with compliant households in control districts. These beneficiaries did not comply with the requirements for UID-based transfer till March 2014. Using monthly coefficients from an empirical specification similar to (Equation 5), this plot reports monthwise treatment effect on the non-compliant beneficiaries. Household fixed effects are included. Above plot also shows evidence of bunching right before the enforcement of the UID-based transfer policy.

Figure 12: Early complier households: Increase in fuel purchase



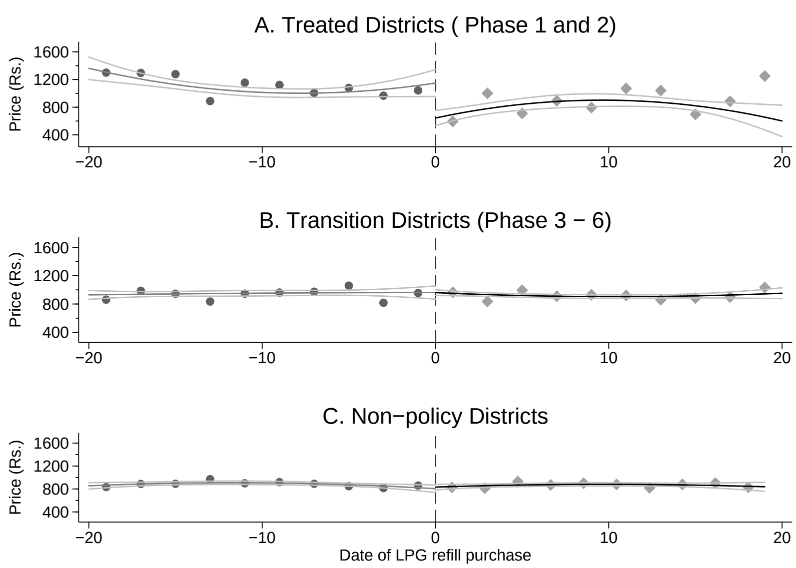
Note: The above plot shows the purchase behavior of early compliant households in Phase 1 districts. This plot reports monthwise coefficient on LPG refill purchase by the early complier households. Household fixed effects are included. Sub-sample includes households who fulfilled the compliant requirement before treatment (i.e. before Sep 2013). There is a steep increase in LPG refill purchase, when UID requirement is enforced in Sep 2013.

Figure 13: All compliant households: Similar purchase behavior across treated and control districts



Note: This plot shows the LPG refill purchase behavior of all compliant households in treated districts when compared with compliant households in Phase 6 districts. Since the policy was introduced in Phase 6 districts in January 2014, this control group is relatively free of any concerns about strategic refill timing by households during the transition period (specifically, between September to December). This graph provides a reasonable assurance against concerns about implementation issues pertaining to compliant households (such as, delay in bank transfers or UID reporting affecting purchase behaviors in treated districts) since purchase behavior of households is mostly comparable with households in control group. Next, this plot also shows that early complier households in treated districts do not cause any significant increase in fuel purchase. Note that Phase 6 compliant households are early complier households, since the policy was introduced (in January 2014) but was not yet enforced.

Figure 14: Policy Termination: Discontinuity in Black market Price



Note: This plot shows discontinuity in the reported black market LPG refill prices right after enforcement policy is terminated in treated districts (Panel A). A 20-days window is used on either side of the termination date. Transition districts (Panel B) and non-policy districts (Panel C) observe no change.

Table 1: Descriptive statistics – Administrative data

A. Household level LPG transactions data			
	Mean	Median	SD
Subsidized Refills	6.523	7	2.853
Total Refills	6.575	7	2.935
Monthly Refills	0.553	1	0.586
Households		3.79 million	
Distributors		3165	
Districts		509	
States		25	
Time period		12 months (Apr 2013 - Mar 2014)	
Transactions		23.17 million	

B. Distributor level LPG sales data			
	Mean	Median	SD
14kg refills (Domestic)	6,670	5,656	5,530
19kg refills (Commercial)	459.8	150	1,007
Distributors		3341	
Districts		504	
States		25	
Time period		13 months (Apr 2013 - Apr 2014)	
Monthly observations		43433	

Note: Details of beneficiary level LPG refills transaction data are shown in Panel A. Fuel for domestic cooking use is sold as 14kg LPG refills. This data is 10% randomly selected sample from Hindustan Petroleum Corporation Limited (HPCL) and it represents about 2.5% of total household LPG purchase in India in 2013-2014. More than 95% households received all subsidized refills (i.e. annual cap is not binding for almost whole population). Panel B shows LPG distributor level data. This data covers all the LPG distributors under HPCL. HPCL has about 25% of total LPG market share in India.

Table 2: Descriptive statistics – Survey data

Variable	N	Mean	SD	Min	Max
Demand side price	2369	1039.13	241.26	430	1600
Supply side price	1202	1062.49	233.41	550	1950
Firms		1452			
Delivery men		1202			
District		89			
State		11			

Note: The above table shows summary statistics of black market LPG refill prices collected from the supply and demand side in two rounds (before and after policy termination). Price values are in INR. In demand side survey, about 20% of attrition is observed in the post-termination round.

Table 3: Impact of UID-based transfer policy on domestic fuel sales (Beneficiary level data)

	(1)	(2)	(3)
Outcome variable: Household monthly LPG refill purchase			
Post	0.126***	0.108***	0.158***
	(0.00532)	(0.00568)	(0.00753)
UID-based transfer X Post	-0.0664***	-0.0621***	-0.0769***
	(0.00375)	(0.00401)	(0.00466)
Constant	0.484***	0.485***	0.475***
	(0.00319)	(0.00378)	(0.00396)
Observations	37,408,250	27,389,714	13,064,788
Household	3,400,750	2,489,974	1,187,708
Mean of outcome var	0.561	0.556	0.556
Control group	Ph 3-6 & Non-policy	Ph 3-6	Non-policy
Household FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes

Note: This regression estimates the impact of UID-based transfer policy on domestic fuel purchase using OLS. A household-month level panel is used. Estimates suggest about 11% to 14% reduction in fuel purchase in domestic cooking sector (i.e. coefficient on the interaction term as the percentage of mean value). Outcome variable is – number of LPG refills purchased by a beneficiary in a given month. Household and month fixed effects are included. Treated group includes all Phase 1 districts in the sample (16 districts). Phase 2 districts are not included. Col (1) combines all upcoming phases and non-policy districts together in the control group. Col(2) and Col (3) present estimates from the same specification, but with two different sub-groups as control. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses. Standard errors are clustered at the district level.

Table 4: Impact of UID-based transfer policy on domestic fuel sales (Distributor level data)

	(1)	(2)	(3)
Outcome variable: log(Domestic LPG refills sales)			
Post	0.285***	0.243***	0.341***
	(0.0126)	(0.0152)	(0.0171)
UID-based transfer X Post	-0.149***	-0.134***	-0.174***
	(0.0110)	(0.0118)	(0.0128)
Constant	8.178***	8.357***	7.975***
	(0.00716)	(0.00927)	(0.00953)
Observations	31,322	19,944	13,135
District	485	236	264
Distributor	3013	1909	1269
Control	Ph3-6 & Non-policy	Ph3-6	Non-policy
Distributor FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes

Note: This regression estimates the impact of UID-based transfer program on domestic-use fuel sales using OLS. A distributor-month level panel is used. Estimates suggest about 13% to 17% reduction in domestic-use LPG purchase. Outcome variable is – log(Total domestic-use LPG refills sold to households in a given month). Distributor and month fixed effects are included. Treated group includes all Phase 1 districts in the sample (16 districts). Phase 2 districts are not included. Col (1) combines all upcoming phases and non-policy districts together in the control group. Col(2) and Col (3) present estimates from same specification, but with two different control groups. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses. Standard errors are clustered at the district level.

Table 5: Impact of UID-based transfer policy on commercia fuel sales (Distributor level data)

	(1)	(2)	(3)
Outcome variable: log(Commercial LPG refills sales)			
Post	-0.0121	0.00294	-0.0468
	(0.0173)	(0.0206)	(0.0285)
UID-based transfer X Post	0.0113	-0.000159	0.0326
	(0.0212)	(0.0217)	(0.0234)
Constant	5.119***	5.447***	4.672***
	(0.0110)	(0.0132)	(0.0182)
Observations	24,288	16,303	9,475
District	482	235	262
Distributor	2678	1745	1082
Control	Ph3-6 & Non-policy	Ph3-6	Non-policy
Distributor FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes

Note: This regression estimates impact of the UID-based transfer program on commercial fuel sales. A distributor-month level panel is used. Outcome variable is – log(Total commercial-use 19kg LPG refills sold to businesses in a given month). Estimates are not significant. Note that the data does not include all commercial LPG sales, since LPG is distributed with higher size cylinders, in bulk and as auto-fuel. Distributor and month fixed effects are included. Treated group includes all Phase 1 districts in the sample (16 districts). Phase 2 districts are not included. Col (1) combines all upcoming phases and non-policy districts together in the control group. Col(2) and Col (3) present estimates from the same specification, but with two different sub-groups as control. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses. Standard errors are clustered at the district level.

Table 6: Impact of UID-based transfer policy termination on domestic fuel sales (Beneficiary level data)

	(1)	(2)	(3)
Outcome variable: Household monthly LPG refill purchase			
Post termination	0.101***	0.0849***	0.141***
	(0.00513)	(0.00557)	(0.00620)
UID-based transfer X Post termination	0.0343***	0.0444***	0.00963
	(0.00509)	(0.00519)	(0.00676)
Constant	0.560***	0.558***	0.538***
	(0.00314)	(0.00382)	(0.00279)
Observations	23,885,798	17,481,131	8,347,633
Mean of outcome var	0.605	0.597	0.603
Control group	Ph 3-6 & Non-policy	Ph 3-6	Non-policy
Month FE	Yes	Yes	Yes
Household FE	Yes	Yes	Yes

Note: This regression estimates the impact of UID-based transfer policy termination on domestic fuel purchase by the beneficiaries. Outcome variable is – the number of LPG refills purchased by a beneficiary in a given month. A household-month level panel is used. Comparing the interaction coefficient with the mean value, estimates suggest about 7% increase in fuel purchase in domestic sector (i.e. coefficient on the interaction term as a percentage of the mean value). Household and month fixed effects are included. Treated group includes all Phase 1 districts in the sample (16 districts). Phase 2 districts are not included. Col (1) combines all upcoming phases and non-policy districts together in the control group. Col(2) and Col (3) present estimates from same specification, but with two different sub-groups as control. Note that the sample includes only one month (March 2014) in the post-termination period. Whole March month is considered as post-treatment period for consistency, though the exact date of policy termination is 10 March. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses. Standard errors are clustered at the district level.

Table 7: Impact of UID-based transfer policy termination on domestic fuel sales

	(1)	(2)	(3)
Outcome variable: log(Domestic LPG refill sales)			
Post termination	0.0396***	-0.0108	0.115***
	(0.0112)	(0.0122)	(0.0161)
UID-based transfer X Post termination	0.101***	0.127***	0.0593**
	(0.0234)	(0.0235)	(0.0247)
Constant	8.303***	8.492***	8.069***
	(0.00678)	(0.00808)	(0.00812)
Observations	23,396	14,826	9,854
District	485	236	264
Distributor	3060	1932	1294
Control	Ph3-6 & Non-policy	Ph3-6	Non-policy
Distributor FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes

Note: This regression estimates the impact of the UID-based transfer policy termination on fuel sales in the domestic sector. A distributor-month level panel is used. Above estimates suggest about 6% to 13% increase in fuel purchase in domestic sector. Outcome variable is – log(Total domestic-use LPG refills sold to households in a given month). Distributor and month fixed effects are included. Treated group includes all Phase 1 districts in the sample (16 districts). Phase 2 districts are not included. Col (1) combines all upcoming phases and non-policy districts together in the control group. Col(2) and Col (3) present estimates from the same specification, but with two different sub-groups as control. Post-termination period includes two months (March and April 2014). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses. Standard errors are clustered at the district level.

Table 8: Impact of UID-based transfer policy termination on commercial fuel sales

	(1)	(2)	(3)
Outcome variable: log(Commercial LPG refill sales)			
Post termination	-0.0291*	-0.0429**	0.00494
	(0.0167)	(0.0206)	(0.0272)
UID-based transfer X Post termination	-0.0727**	-0.0637*	-0.0895**
	(0.0348)	(0.0353)	(0.0377)
Constant	5.124***	5.467***	4.636***
	(0.0106)	(0.0126)	(0.0186)
Observations	17,661	11,862	6,883
District	480	234	261
Distributor	2637	1727	1060
Control	Ph3-6 & Non-policy	Ph3-6	Non-policy
Distributor FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes

Note: This regression estimates the impact of the UID-based transfer program on commercial fuel sales. A distributor-month level panel is used. Estimates suggest about 6% to 9% reduction in commercial fuel purchase. Outcome variable is – log(Total commercial-use 19kg LPG refills sold to households in a given month). This does not include all commercial LPG sales. Distributor and month fixed effects are included. Treated group includes all Phase 1 districts in the sample (16 districts). Phase 2 districts are not included. Col (1) combines all upcoming phases and non-policy districts together in the control group. Col(2) and Col (3) present estimates from the same specification, but with two different sub-groups as control. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses. Standard errors are clustered at the district level.

Table 9: Impact of UID-based transfer policy termination on black-market price

	(1)	(2)	(3)	(4)
	A. Supply Side		B. Demand Side	
Outcome variable: log (price)				
UID-based transfer			0.188***	0.196***
			(0.0371)	(0.0582)
Post termination	-0.205***	-0.203***	-0.203***	-0.444***
	(0.0369)	(0.0606)	(0.0360)	(0.0715)
UID-based transfer X Post termination	-0.127***	-0.159***	-0.175***	-0.192***
	(0.0417)	(0.0406)	(0.0610)	(0.0736)
Constant	7.058***	7.023***	6.973***	7.134***
	(0.0137)	(0.0234)	(0.0363)	(0.00787)
Observations	504	504	1,000	1,000
Treatment	Ph1&2	Ph1&2	Ph1&2	Ph1&2
Control	Non-policy	Non-policy	Non-policy	Non-policy
Firm			602	602
District	38	38	38	38
District FE	Yes	Yes		
Firm FE			Yes	Yes
Date FE		Yes		Yes

Note: This table presents difference-in-difference estimate of the impact of enforcement policy termination on the black market fuel prices. The outcome variable is log(price). The sample consists of the treated districts (Phase 1 and 2) and non-policy districts in the control group. These 38 districts are from 8 states. Panel A (Col (1) and (2)) shows the impact of policy termination on the quoted price by the supply side (delivery man). Coefficients on the interaction term suggest 13% to 16% decrease in the black-market prices in treated districts, when the UID-based transfer policy is terminated. Note that the dummy on the treated group is included in the empirical specification, but the variation is absorbed by district fixed effects in Col(1) and (2). In Panel B, the outcome variable is the ongoing black-market price mentioned by small businesses. This panel shows about 18% to 19% decrease in the black-market price, when the policy is terminated. Districts, firms and interview date fixed effects are used where applicable. Additional robustness checks are provided in the Appendix. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses. Standard errors are clustered at the district level.

Table 10: Heterogenous effect on domestic fuel purchase

	(1)	(2)	(3)
Outcome variable: Household monthly LPG refill purchase			
Post	0.100***	0.0971***	0.140***
	(0.00771)	(0.00796)	(0.00926)
High frequency HH	0.548***		0.464***
	(0.00607)		(0.00348)
Post X High frequency HH	-0.345***		-0.249***
	(0.0273)		(0.0155)
HH not complied		-0.0233	-0.0789***
		(0.0147)	(0.00735)
Post X HH not complied		-0.131***	-0.0816***
		(0.0134)	(0.00792)
High frequency HH X HH not complied			0.154***
			(0.0119)
Post X High frequency HH X HH not complied			-0.154***
			(0.0345)
Constant	0.338***	0.461***	0.378***
	(0.00444)	(0.00694)	(0.00548)
Observations	3,095,114	3,095,114	3,095,114
Households	281,374	281,374	281,374
Month FE	Yes	Yes	Yes
District FE	Yes	Yes	Yes

Note: This table shows the heterogeneity in the impact of UID-based transfer policy on different subgroups in treated districts. Beneficiaries are divided based on their pre-treatment fuel purchases. Next, households compliant status is used to segregate the impact on compliant and non-compliant beneficiaries. With triple difference estimate, the coefficient on *"Post X High frequency HH X HH not complied"* shows that the UID-based transfer policy had a much stronger impact on the high frequency beneficiaries who failed to comply. District fixed effects are used. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. Standard errors are clustered at the district level.

Table 11: Impact of UID-based transfer policy on the beneficiaries that purchased no refills

	(1)	(2)	(3)
Outcome variable: log(count of households who purchased zero refills)			
Post	-0.194***	-0.161***	-0.249***
	(0.00834)	(0.00949)	(0.0136)
UID-based transfer X Post	0.107***	0.0993***	0.126***
	(0.00565)	(0.00644)	(0.00684)
Constant	6.827***	6.934***	6.704***
	(0.00454)	(0.00599)	(0.00489)
Observations	38,294,158	27,954,157	13,435,115
Control group	Ph 3-6 & non-policy	Ph 3-6	Non-policy
Mean of outcome var	1057	1159	990.3
Month FE	Yes	Yes	Yes
Distributor FE	Yes	Yes	Yes

Note: This table reports the impact of UID-based transfer policy enforcement on the count of beneficiaries who did not purchase fuel in a given month. This includes the subsidized as well as non-subsidized fuel refill purchase. A distributor-month level panel is created by collapsing the household-month data. Outcome variable is in log. The estimated coefficients suggest about 10% to 13% increase in the number of beneficiaries who did not purchase a single LPG refill. Distributor and month fixed effects are included. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses. Standard errors are clustered at the district level.

Table 12: Domestic fuel purchase by compliant households

	(1)	(2)
Outcome variable: Household monthly LPG refill purchase		
Post	0.0525***	0.0659***
	(0.00953)	(0.0118)
UID-based transfer X Post	0.0215***	-0.00195
	(0.00476)	(0.00590)
Constant	0.479***	0.488***
	(0.00330)	(0.00564)
Observations	11,803,583	3,899,918
Control group	Ph 3-6	Ph 6
Mean of outcome var	0.537	0.551
Month FE	Yes	Yes
Household FE	Yes	Yes

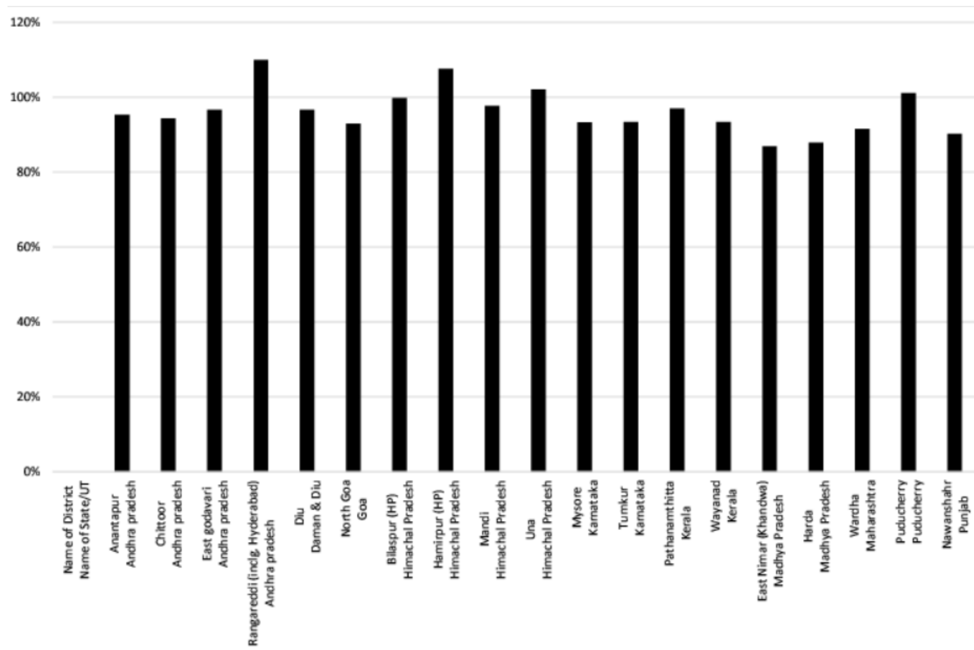
Note: This table reports the impact of the UID-based transfer policy enforcement on the increase in LPG refill purchase by compliant households. All the households who complied by March 2014 are included. Col (1) suggests about 4% increase in LPG refill purchases by compliant households in the treated districts. Col (2) uses the Phase 6 districts as control and does not show any significant effect. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses. Standard errors are clustered at the district level.

A APPENDIX

Figure A15: Sample Unique Identification (UID)

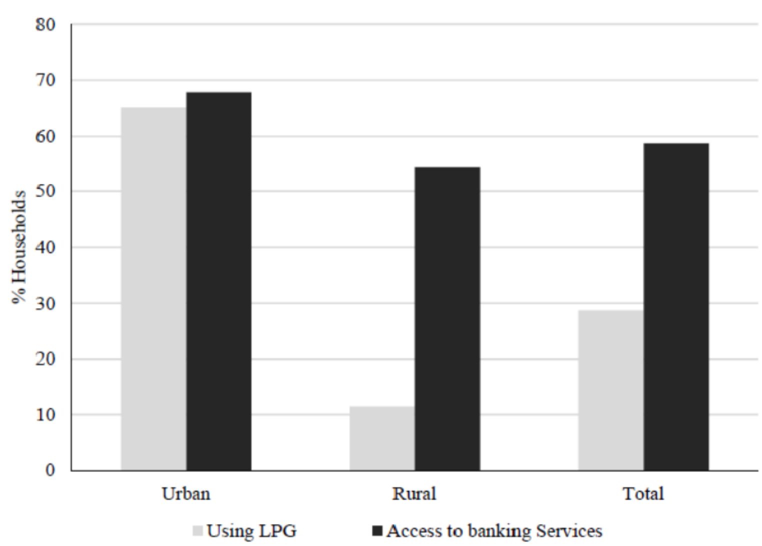
Note: UID number is communicated to the residents in above format. Also known as Aadhaar number, each 12 digit number is unique and is de-duplicated with already existing biometric database.(Source: `http://indane.co.in/images/aadhar-reg.jpg`

Figure A16: UID penetration in Phase 1 districts



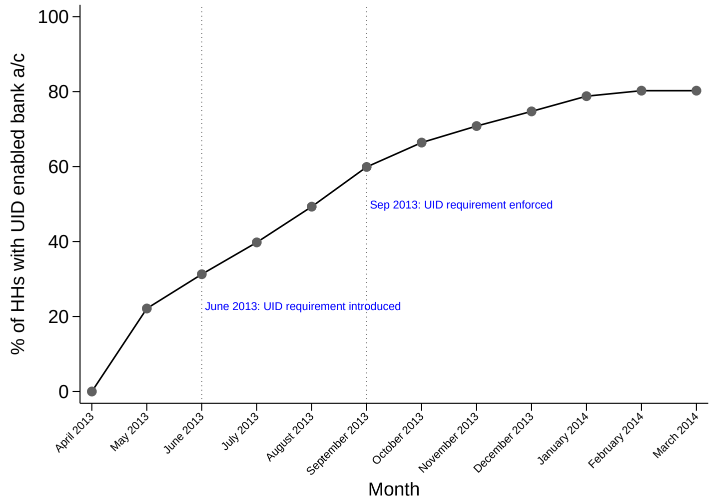
Note: This graphs shows UID coverage in Phase 1 districts in January 2014 (right before the policy termination), as per the data provided by the UID Authority of India. In these districts, 98.5% population (compared to the 2014 census) had received UID number. UID penetration in four districts exceeds their respective population, likely because of in-migration since 2011. Similarly, it is possible that some of the other districts could not reach 100% penetration because of an out-migration.

Figure A17: Access to banking services and LPG adoption



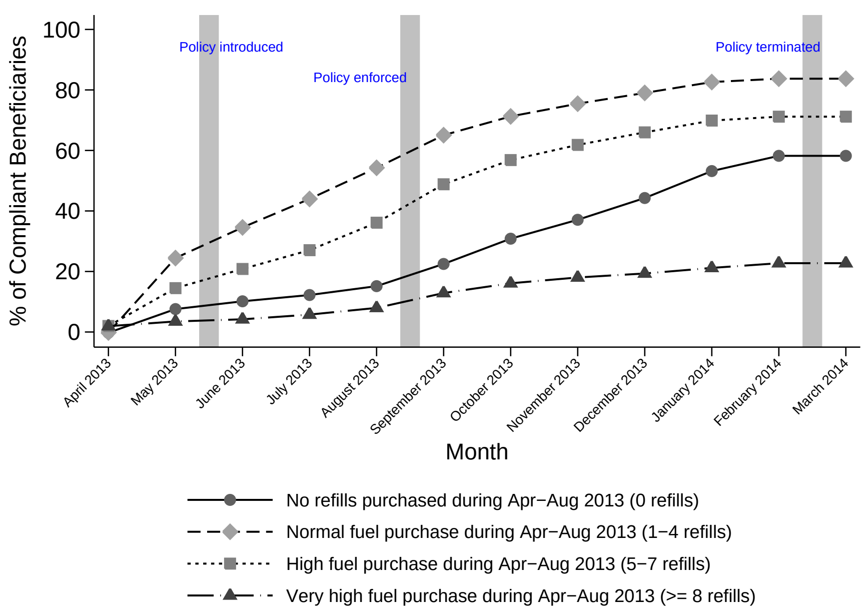
Note: 2011 Census collects information on "household having any type of bank account" (that does not include Self-help Groups, Agricultural Credit Societies etc.). Comparing data on LPG user households from the same source, the access to the banking services is more than the adoption of LPG for cooking purpose. This holds for urban as well as rural areas. Further, India has conducted dedicated campaign to increase the financial inclusion during 2011 to 2013.

Figure A18: Fuel subsidy beneficiary household compliance



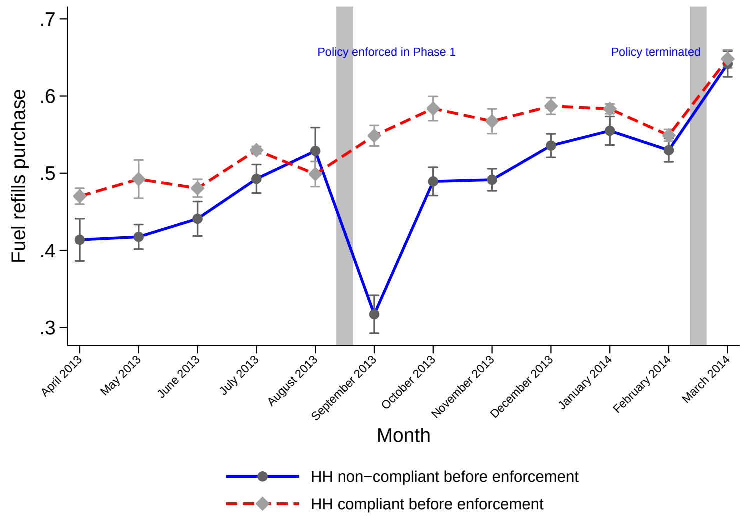
Note: This table shows household compliance in treated districts (Phase 1). Compliance requires households to submit their UID and bank account number. Right after the introduction of the UID-based transfer policy, compliance increased steeply and gradually the take up rate decreases. Households are not necessarily required to comply if they do not want subsidy transfer. When UID-based transfer policy was enforced, a non-compliant beneficiary could continue to avail domestic fuel, but not the subsidy. It is likely that the timing of next refill would affect household's decision to comply. Overall more than 80% compliance was achieved during the six month enforcement period.

Figure A19: Compliance and pre-enforcement fuel purchase



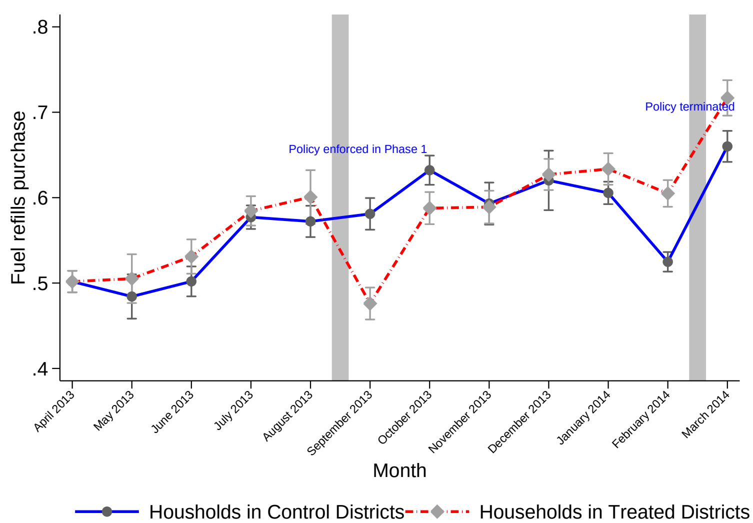
Note: The above plot shows how compliance levels vary with the number of LPG refills bought by a beneficiary households before the UID-based transfer program was enforced (Apr-Aug 2013). Sample consists of data from Phase 1 districts. A "normal household" purchases 4 LPG refills as per the pro-rated annual cap basis. Higher purchase frequency households consistently show low compliance. Note that households who did not purchase a single refill in the pre-enforcement period, show a much higher compliance rate after the policy is enforced.

Figure A20: Comparison of the early and late complier households in treated districts



Note: This coefficient plot shows comparison of late vs early compilers in the treated districts (all Phase 1 districts). It suggests late complier households decrease their LPG purchase in the first month of enforcement and then eventually catch up with the complier households. Note that the late complier households are about 20% of total number of households.

Figure A21: Comparison of the late complier households in treated districts with complier households in control districts



Note: This plot shows LPG refill purchase behavior of all late compliant households in treated districts when compared with compliant households in Phase 6 districts. Since the policy was introduced in Phase 6 districts in January 2014, this control group is relatively free of any concerns about strategic refill timing by households during the transition period (specifically, between September to December). This graphs shows that late complier in treated districts, on average, did not buy lower fuel than the complier households in control districts. First two months in the enforcement period observe a decrease in the fuel purchase, but households increase their fuel purchase in subsequent months. Note that Phase 6 compliant households are early complier households, since the policy was introduced (in January 2014) but was not yet enforced.

Table A13: Domestic fuel sales in Phase 2 (Household level data)

	(1)	(2)	(3)
Outcome variable: Household monthly LPG refill purchase			
Post	0.131***	0.114***	0.177***
	(0.00563)	(0.00626)	(0.00594)
UID-based transfer X Post	-0.134***	-0.123***	-0.162***
	(0.00660)	(0.00693)	(0.00705)
Constant	0.481***	0.481***	0.466***
	(0.00321)	(0.00382)	(0.00412)
Observations	37,579,113	27,560,577	13,235,651
Household	3,416,283	2,505,507	1,203,241
Mean of outcome var	0.559	0.554	0.550
Control group	Ph 3-6 & Non-policy	Ph 3-6	Non-policy
Household FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes

This table reports estimates of the impact of UID-based transfer program in Phase 2 districts. A household-month level panel is used. Outcome variable is – *number of LPG refills purchased in a month*. Estimates suggest about 22% to 29% reduction in domestic-use LPG purchase (i.e. coefficient on the interaction term as a percentage of mean value). Phase 2 districts had the UID-based transfer policy enforced for a relatively short period, so these estimates include households' timing behavior. Phase 1 districts are not included. Note that control groups are different in three columns and provide a robustness check. Household and month fixed effects are included. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses. Standard errors are clustered at the district level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses.

Table A14: Fuel purchase in domestic sector: Comparison of OLS and Poisson estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	Poisson	OLS	Poisson	OLS	Poisson
Outcome variable: Household monthly LPG refill purchase						
Post	0.125***	0.230***	0.110***	0.206***	0.152***	0.274***
	(0.00786)	(0.00797)	(0.00925)	(0.00948)	(0.00981)	(0.0135)
UID-based transfer X Post	-0.0643***	-0.112***	-0.0619***	-0.109***	-0.0701***	-0.119***
	(0.00562)	(0.0130)	(0.00596)	(0.0131)	(0.00666)	(0.0141)
Constant	0.481***		0.477***		0.481***	
	(0.00473)		(0.00578)		(0.00629)	
Observations	375,914	375,914	274,450	274,450	133,562	133,562
Districts	487	487	238	238	265	265
Distributors	2750	2750	1765	1765	1140	1140
Households	34174	34174	24950	24950	12142	12142
Mean of outcome var	0.561		0.554		0.561	
Control group	Ph 3-6 & Non-policy	Ph 3-6 & Non-policy	Ph 3-6	Ph 3-6	Non-policy	Non-policy

Note: This table presents a comparison of estimates from the OLS and Poisson models for robustness check. The outcome variable (i.e., number of LPG refills per household per month) has a structure similar to Poisson distribution. 1% sample is used. Poisson estimates of the causal impact of the UID-based transfer policy are very close to OLS estimates. Month and household fixed effects are included. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses.

Table A15: Robustness check: Impact of UID-based Transfer Policy Termination on Black market Price (Demand Side)

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome variable: log(price)						
Policy termination	-0.257***	-0.452***	-0.285***	-0.234**	-0.203***	-0.444***
	(0.0192)	(0.0481)	(0.0212)	(0.115)	(0.0360)	(0.0715)
UID-based transfer	0.127***	0.193***	0.0956***	0.251**	0.188***	0.196***
	(0.0239)	(0.0551)	(0.0287)	(0.102)	(0.0371)	(0.0582)
UID-based transfer X Policy termination	-0.122**	-0.210***	-0.0940*	-0.276**	-0.175***	-0.192***
	(0.0530)	(0.0731)	(0.0541)	(0.110)	(0.0610)	(0.0736)
Constant	7.035***	7.131***	7.066***	6.907***	6.973***	7.134***
	(0.0225)	(0.00833)	(0.0275)	(0.101)	(0.0363)	(0.00787)
Observations	2,369	2,369	1,622	1,622	1,000	1,000
Firm	1,406	1,406	955	955	602	602
District	89	89	61	61	38	38
Control	Ph3-6 & Non-policy	Ph3-6 & Non-policy	Ph3-6	Ph3-6	Non-policy	Non-policy
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Day FE		Yes		Yes		Yes

Note: This table shows the impact of policy termination on the ongoing black market prices as collected from the small businesses. Outcome variable is Log(black-market price). Even numbered columns include interview date fixed effect. Robustness is checked with different combinations of control groups. Col (5) and Col(6) present the preferred specification (already provided in the paper) and are provided here for a comparison. Firm FE are included. Standard errors are clustered at the district level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses.

Table A16: Robustness check: Impact of UID-based Transfer Policy Termination on Black market Price (Refill History Data)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Outcome variable: log(price)							
Policy termination	-0.131**	-0.127**	-0.127**	-0.191**	-0.00692	0.0627**	0.0658**
	(0.0642)	(0.0620)	(0.0621)	(0.0765)	(0.0263)	(0.0264)	(0.0257)
UID-based transfer X Policy termination	-0.0902**	-0.0916**	-0.0915**	-0.0711**	-0.163**	-0.112**	-0.117***
	(0.0374)	(0.0383)	(0.0349)	(0.0345)	(0.0717)	(0.0451)	(0.0407)
Constant	6.861***	6.865***	6.866***	6.932***	6.799***	6.816***	6.820***
	(0.0201)	(0.0200)	(0.0200)	(0.0397)	(0.0223)	(0.0235)	(0.0236)
Observations	1,895	2,021	2,037	1,271	782	908	924
Firm	624	671	677	424	259	306	312
District	74	79	81	53	30	35	37
Treatment group	Ph 1	Ph 2	Ph 1 & 2	Ph 1 & 2	Ph 1	Ph 2	Ph 1 & 2
Control group		Ph 3-6 & Non-policy		Ph 3-6		Non-policy	
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: The above table shows the impact of policy termination on LPG refills prices using refill history data. In survey, each small business is asked for the date and price of last five LPG refills. Firm-date level panel is constructed with this data. Coefficient on the interaction term provide the impact of policy termination on the black market prices paid by the firms. Note that different control and treatment groups are used for robustness check. Firm and Refill date level FE are included. Standard errors are clustered at the district level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses.

Chapter 2

Mines: The Local Wealth and Health Effects of Mining in Developing Countries

Mines

The Local Wealth and Health Effects of Mining in Developing Countries*

Jan von der Goltz †‡ Prabhat Barnwal†

November 11, 2014

Abstract

Do residents of mining communities face health-wealth trade-offs? We conduct the first extensive assessment of this question using micro-data from communities near about 800 mineral mines in 44 developing countries. Households in mining communities enjoy a substantial medium-term gain in asset wealth (0.3 standard deviations), but experience a ten percentage point increase in the incidence of anemia among adult women, and a five percentage point rise in the prevalence of stunting in young children. Prior evidence links both of these health impacts to metal toxicity – and in particular, exposure to high levels of lead. We observe health impacts only near mines of a type where heavy metal pollution is to be expected, and find no systematic evidence that health is affected in ways that are not specific to exposure to such pollutants. Benefits and costs are strongly concentrated in the immediate vicinity ($\leq$ 5km) of a mine. Consistent results emerge from a range of distinct identification strategies. Baseline results come from a cross-sectional fixed effects model, and mine-level and mother-level panels. An instrumental variables approach serves as a robustness check. To demonstrate that the observed health impacts are due to pollution, we develop three difference-in-difference tests tailored to the known association of certain mine types with heavy metal pollution, and to the pathophysiology of lead toxicity. Our results add to the nascent literature on health impacts near industrial operations in developing countries.

Key Words: *Mines, Health, Development*

*The authors particularly thank Douglas Almond, Joseph Graziano, and Eric Verhoogen for their guidance and advice. We are indebted to Jesse Anttila-Hughes for his generous support and helpful discussions in the early stages of the project. We thank Ram Fishman, Alexander van Geen, Kyle Meng, Suresh Naidu, Matt Neidell, Bernard Salanié, Wolfram Schlenker, Anna Tompsett, as well as seminar and conference participants for comments. Nicole Dussault and Dheeraj Sharma provided excellent research assistance. We gratefully acknowledge financial support from the Centre on Global Economic Governance and the Center for International Business Education and Research at Columbia University. All errors are ours. We invite comments and feedback on the working paper.

†PhD program, School of International and Public Affairs, Columbia University. *email:* Jan von der Goltz (jcv2104@columbia.edu), Prabhat Barnwal (pb2442@columbia.edu).

‡Corresponding author.

Online appendices: http://www.columbia.edu/%7E jcv2104/research.shtml.

1 Introduction

This paper studies the local wealth and health impacts of mineral mining in 44 developing countries. We show that, while residents living close to mines enjoy greater wealth, there is a trade-off: life near mines exacts a price in terms of specific health burdens.

In any country and at any time, the decision to live near centers of industrial activity involves weighing the promise of economic opportunity against the risk of disamenity caused by pollution. Nowhere is this choice starker than in developing countries. More often than not, opportunities for making a good living are precious and few. At the same time, pollution tends to be poorly regulated, and information on health risks and on ways to avoid them, scarce. Poor infrastructure and inflexible housing markets commonly make commuting to avoid pollution impracticable. Yet, while "the literature on the health effects of pollution has advanced greatly in the last two decades, almost all of this research has been conducted in developed country settings." (Greenstone and Jack, 2013, p. 12)

In the following paper, we present the first systematic empirical assessment of the micro-level trade-offs between health and wealth posed by the mining industry in developing countries. We seek to add to the very limited number of broad micro data analyses of local health impacts near *any* kind of industrial operations in developing countries. For the study of industrial pollution in poor countries, the mining and mineral processing industry is an attractive test case in that it poses particularly sharp trade-offs. Single plants generate very high value – in some instances, in the hundreds of millions or billions of dollars per year. The location of ore deposits dictates where mines open, and because of transport costs, often also where smelters locate. Therefore, large operations are found in remote areas where they dwarf any other enterprises – and the economic opportunities generated by the latter. Mines and smelters therefore tend to play a conspicuous economic role. At the same time, however, they are very large polluters, and precisely because they are important sources of revenue, foreign exchange, and employment, there is a risk of weak environmental regulation and enforcement.

The importance of mining to development is reflected in a long tradition of research on the macroeconomic implications of mining and the optimal management of mineral resources. However, there is little empirical evidence on the local economic impact of mining, and on its effects on other dimensions of well-being. This particularly includes implications for the health of local communities: although there is an important literature on pollution near mines, and an extensive body of knowledge on the toxic properties of common pollutants, there is scant systematic evidence linking the two. No more than a handful of case studies have carefully assessed the actual clinical consequences of exposure to a mining environment. This paucity of empirical evidence on the local welfare effects of mining is in stark contrast to the strong passions that mining projects habitually evoke among the communities affected. In some places, projects have been supported vociferously, and people have fought over the right to work in mines. Yet, in other places, mining has been desperately opposed, as citizens feared damage to their health and environment. Our work shows that these political passions are grounded in a real trade-off. Across a broad range of settings, the local benefits of mining are real, but so are the costs.

We analyze the effect of mining activity on asset wealth, on general health, and on two specific health outcomes known to be linked to pollutants that may be present around mines in our sample, namely anemia in adults and children, and growth in young children. To study the interplay of health and wealth effects across the developing world, we compile 104 waves of Demographic and Health Surveys from 44 countries. The pooled data provide us with about 1.2m household-level records and several million individual-level observations, spanning a time period from 1986 until 2012; they also record the geo-location of each cluster of households sampled. We overlay this household data with information on the location of mining and smelting operations across the world. We use a large cross-sectional dataset of mines that records the types of minerals mined, and characteristics of the local geology; as well as two business intelligence datasets that document annual production at individual mines. Guided by prior evidence from the environmental sciences on the spatial extent of pollution around mines, we define households *within 5km* of a mine to be in its direct vicinity, and consider those to be treated. We regard households *within 5-20km* of a mine to be in its general vicinity, and rely on those in constructing control groups. Using production data, we create pseudo-panels that enable us to compare our treatment and control groups across time, namely between years when the mine was operational, and when it was dormant.

We then construct a broad range of complementary statistical tests that rely on different control groups, and offer extensive placebo tests. Table 1 provides an overview of the seven different estimators we use. In our baseline models, we estimate the effect of closeness to mines and smelters in the cross-section, and the effect of closeness and operational status in the panel. (We prefer pooling mineral mining and processing facilities, but show robustness to excluding smelters from the sample.) We weaken cross-sectional identifying assumptions by defining our control group conservatively, and by allowing for a fixed effect common to all clusters observed near the same mine, and in the same survey year. Because of the possibility of residential sorting, we argue that our cross-sectional estimates are best read as the long-run effect of mining on *communities*, much like district or county-level studies assess impacts on those units of analysis. To assess the effect of exposure to mining on individuals, we create two sets of pseudo-panels: a mine-level panel compares households observed near the same mine in different years, and a mother-level panel compares among siblings born in different years. The panels allow for common effects shared by all households observed in the same country and survey round. An instrumental variables (IV) approach further reassures us that our results are not due to endogenous choice of mine location or periods of operation; to this end, we use the location of mineral deposits and world mineral prices to instrument for the location and operational status of mines.

Beyond these standard identification frameworks, we develop three difference-in-differences tests that are tailored to prior knowledge on the toxic properties of mining pollution. Our purpose in designing these tests is two-fold. Firstly, they help bolster our claims to observing a causal impact of mining on health. In particular, we believe that they are largely immune to residential sorting. Secondly, our tests provide evidence to suggest that the observed health effects are due to pollution, not other mechanisms. To devise the tests, we (i) leverage knowledge on the association of specific mine types with lead contamination – and by extension, health impacts specific to lead – to conduct falsification tests. We show both that we only observe those health impacts that are expected from exposure to lead pollution, and that we only observe them near mine types strongly associated with the release of lead. Furthermore, we (ii) exploit detailed information on the birth history of women to describe a pattern of impaired ability to recover from blood loss after pregnancy among women living in mining communities, as compared to those living in the general vicinity. We argue that this effect is consistent with a known pathophysiological pattern of lead toxicity in adults, but not easily consistent with other mechanisms. Finally, we (iii) compare children born shortly before or after their family migrated to a mining community to those born to families migrating to the general vicinity of mines. We show that patterns of poor growth among those moving to mining towns correspond with variation in exposure to pollution *in utero* induced by the timing of conception relative to the date the family moved.

Our results show that, at the global mean, long-run asset wealth in mining communities rises by about 0.1 standard deviations of an asset index computed for the country where the community is located and the year in which the survey was taken. The medium-term wealth of households living in the vicinity of an operating mine rises by about 0.3 standard deviations. We illustrate that these are considerable effects, given the high variation in asset ownership within survey rounds. Wealth effects are strongly concentrated in the direct vicinity of the mine; there are benefits across the wealth distribution, although in the long run, the wealthiest households benefit the most; across countries, wealth gains are greatest near mines where the overall economic environment is poor.

We find clear evidence of two health impacts that are known consequences of exposure to lead and other heavy metals that may be present near mines. Thus, women in mining communities show depressed blood hemoglobin, and increases in the incidence of anemia of three to ten percentage points. They also recover more slowly from blood loss during pregnancy and delivery, a pattern consistent with prior toxicological research. Children in mining communities suffer some important adverse growth outcomes from *in utero* exposure, with a five percentage point increase in the incidence of stunting – although there is very little evidence of lower birth weight. Growth impacts weaken among older children, perhaps because of the greater wealth enjoyed by households in mining communities. We note particularly that, while our data contains no good measure of cognitive ability, lead exposure has previously been shown to cause cognitive deficits in children at exposure levels below those associated with growth retardation, and far below those associated with overt anemia. By way of contrast to these specific health impacts, we find no effects on health outcomes that are not linked to heavy metal pollution, nor are mining communities differentially affected by other known causes of anemia, or under-served by health care.

Because our paper shows reduced form impacts (that is, it allows for the effect of exposure to a mining environment on health to play out through any channel, including greater wealth), our health results should be interpreted as the *compensated* impact of mining. By implication, since living in mining communities goes hand in hand with economic benefits across the distribution, there is at least no indication that ill health is caused by deprivation. In an illustrative cost-benefit analysis, we show that in the expectation, economic gains outweigh the cost of health impacts, if gains are sufficiently permanent. However, among those who do suffer health impacts, their cost may exceed any economic gains.

This paper seeks to make three contributions to the literature. It is the first to demonstrate that residents of mining communities in developing countries face a trade-off between real economic benefits and specific health costs. Secondly, we add to the limited evidence on the consequences of industrial pollution in developing countries. Finally, we complement the toxicological and epidemiological literature by showing that the health effects of mining pollution are salient in a study of the general population near a large number of mines (rather than in local treatment effects found in case studies), and are robust to tests that require weak identifying assumptions.

The remainder of the paper is organized as follows. Section 2 discusses the prior literature on welfare in mining communities. Section 3 reviews results from environmental science and toxicology that guide the way we develop hypotheses, measure impacts, and interpret results. Section 4 discusses data, and Section 5 summarizes econometric methods. Section 6 presents results, and Section 7 concludes.

2 Existing evidence

This section reviews prior studies of mining and wealth (Section 2.1), and research on health in mining communities (Section 2.2). Section 3 summarizes the much more extensive literature on individual links in the causal chain from mining to ultimate health impacts, namely studies of (i) pollution near mines, (ii) the body burden of pollutants in residents of mining communities, and (iii) the toxic impacts of substances released near mines.

2.1 Mining and wealth

Economics has traditionally studied mineral mining in the context of optimal resource management, or in a macroeconomic context of growth and public finance.[1]

The economic impacts of mining at the local level have only recently received some attention. As of the time of writing, we are aware of only two published papers that study mining at the kind of disaggregated scale we consider. In a pioneering paper, Aragón and Rud (2013b) leverage a change in local hiring and procurement policies in a single very large gold mine in Peru to identify local economic impacts. Incomes in communities within 100km of the mine showed an elasticity of 0.3 to production at the mine, alongside significant increases in the price of housing and of locally produced agricultural output, and higher local public spending. Wilson (2012) shows that asset ownership increased among residents of copper mining communities in Zambia during a boom in the 2000s. A working paper by Aragón and Rud (2013a) investigates the impacts of gold mining in twelve operations in Ghana on agricultural productivity. It finds stark decreases in productivity (40%) in the general vicinity (less than 20km) of mines, relative to control areas farther away. Productivity losses in the general vicinity are accompanied by large increases in the poverty headcount (18 percentage points), and decreases in consumption, all driven by dire developments for rural households. The latter two papers and a working paper by Kotsadam and Tolonen (2013) use sub-sets of the micro data from the Demographic and Health Surveys also used for the present study. Kotsadam and Tolonen (2013) argue that mining activity in a comprehensive sample of African mines fosters sectoral shifts in employment out of agriculture (among women, into services, and among men, into skilled manual labor) and increases cash employment among women, but is also associated with women leaving the labor force altogether.

Long-term welfare in mining communities was also brought to the attention of the research community by Dell's (2010) work on the *mita* forced labor policy in Peru, although the focus of

[1]For a textbook-level overview of the former, see, e.g., Hartwick et al. (1986); for a survey of the latter, Frankel (2010).

the paper is on institutions and development, rather than the direct welfare impacts of mining per se. In other related work, Acemoglu et al. (2013), Dube and Vargas (2013), and Monteiro and Ferraz (2009) have recently leveraged resource revenue at a disaggregated scale as an instrument in the study of other objects of interest (health expenditure, conflict, and corruption, respectively).

2.2 Health effects of mining

Our paper asks how significant are the ultimate health effects in the general population of exposure to pollution from every-day mining and mineral processing operations. Few studies have attempted this before, and to the best of our knowledge, none considers the possible trade-off between wealth and health effects and assesses the issue across many mine sites in a manner that allows for a causal interpretation of results.

Prior work in economics on the issue is limited. Aragón and Rud (2013b) find a significant decrease in general health problems among adults with an expansion of production in the Yanacocha mine, Peru, and no effect among children. In their recent working paper on Ghana, the same authors find evidence of an adverse effect of mining activity on weight-for-height ratios and the prevalence of cough in children living in the general vicinity of twelve gold mines, but no impact on stunting and diarrhea (Aragón and Rud, 2013a). Both of these results are incidental to the main focus of the respective papers, although Aragón and Rud (2013a) note that the observed impact on cough is in line with evidence of air pollution around the mines studied. Some attention has been given to behavioral correlates of mining activity. Wilson (2012) finds that sexual risk-taking tended to decrease in Zambian copper towns during a boom. Corno and De Walque (2012) argue that in mining communities in southern Africa, there was increased risk taking and HIV infection among migrant miners, but no such effect among non-migrants.

In the field of public health, some case studies directly analyze general population health impacts in communities near smelters. Factor-Litvak et al. (1999) find impacts on "intelligence, physical growth, preschool behavior problems, renal function, blood pressure and hematopoiesis," among children of up to 7.5 years of age living in a smelter town in Kosovo (p. 14). Among school-age children living near a lead smelter in Belgium, Roels et al. (1976) find changes in sensitive biomarkers that indicate an incipient disruption of the process of blood formation, but not overt anemia. Both papers show comparisons to a matched control group in addition to dose-response relationships. Dose-response relationships alone have also been reported between blood lead (PbB) and lower blood hemoglobin (Hgb)[2], as well as reduced nerve conductivity, among

[2]The papers report hematocrit, not hemoglobin levels, but the two measures are closely correlated, and are both used to define anemia.

children living near a lead smelter in Idaho, U.S. (Landrigan and Baker, 1981; Schwartz et al., 1990). Baghurst et al. (1992) show a dose-response of IQ to PbB in children living near a lead smelter in Port Pirie, Australia. A range of papers by Hendryx and various co-authors (see for instance Hendryx and Ahern, 2008) shows cross-sectional correlations between county-level health outcomes and Appalachian coal mining, without clear causal claims.

Pollution due to mining is a special case of industrial pollution, and the latter has been analyzed in large and well-identified studies. Yet, most of these investigate developed countries (see Currie et al. (2013) for a major recent contribution); studies of developing countries – especially using large samples – remain rare. Chen et al. (2013) study reduced life expectancy from air pollution due to power generation in China; Ebenstein (2012) assesses the effect of water pollution on gastrointestinal cancer rates in China; and Rau et al. (2013) show cognitive losses from lead exposure near an abandoned toxic waste site in Chile. Hanna and Oliva (2011) describe reductions in air pollution from the closure of a large refinery in Mexico city, and an associated increase in labor demand. Studies of overall urban pollution (Arceo-Gomez et al., 2012; Greenstone and Hanna, 2011) are related, but not specific to industry, while studies of air pollution from urban traffic (e.g. Gallego et al., 2013) are less closely related. We seek to contribute to this nascent literature by presenting a multi-site micro-data study of the comparative health and wealth impacts of an important industry, across many developing countries.

3 Scientific background

This section first discusses environmental pollution near mines, and its relationship to the body burden of toxicants (Section 3.1). We then establish that metals, and in particular lead, are of most interest as pollutants in our sample, and discuss the toxic effects of lead (Section 3.2).

3.1 Environmental pollution due to mining and its relationship to the body burden of toxicants

A voluminous literature in environmental science has catalogued the pollutants emitted in the course of normal operations near mines and smelters of different types. We base the following discussion on Alloway (2013), Ripley et al. (1996), and Wright and Welbourn (2002).

Local communities can be exposed to pollution through a multitude of channels. These include dust from mining, handling and processing; mine waste water; direct exposure to abandoned mine spoils and tailings; metals leached from tailings into soil and water; and particulate and gaseous emissions from roasting and smelting. Sometimes, the material extracted is itself of concern, such as in lead, uranium, or asbestos mining. At other times, pollutants are used in processing, such as in the case of cyanide leaching of gold, or gold and silver extraction by mercury amalgamation. Finally, sometimes the concern is with toxicants co-located with the mineral mined and released either in processing or weathering of mine spoils, such as in the case of heavy metals in non-ferrous metal mining.

Two stylized facts on pollution near mines are essential to the way we analyze the health impacts of mining.

(i) The kinds of pollutants near a given mine can be predicted well from the ore mined.
Table 2 summarizes pollutants associated with common (and non-exclusively defined) mine types in our sample. The mapping is far from exact, but serves as a useful first-order approximation. We leverage the association between target minerals and toxicants to compare health effects across mine types, and to show that we find predicted health impacts only near mine types where pollutants specific to the health impact in question are found.

Of particular interest to us is the association of "non-ferrous metalliferous mining and smelting industries . . . with very high levels of heavy metal(loid) contamination of the environment." (Alloway, 2013, p. 43) Thus, 'polymetallic' mines, where any combination of copper, gold, lead, silver, and zinc are extracted, are linked with a characteristic suite of highly toxic pollutants that includes most prominently lead, but also arsenic, cadmium, and chromium. (We will refer to these metals and metalloids as 'heavy metals' – a term that is imprecise in that it does not refer to a well-defined group of chemical elements, but has the advantage of being in everyday semantics associated with the pollutants we have in mind. See Section 4.2.1 for coding notes.) Pollution near polymetallic mines is of particular concern both because heavy metals are important toxicants, but also because the minerals mined are often nested in sulfide rock. When exposed to air and water, the latter will tend to generate sulphuric acid, which in turn leaches metals from the mine's tailings; the resulting acid mine drainage can pose severe health and environmental concerns (Salomons, 1995).

(ii) The area in which highly polluted sites are found is typically small, and extends to at most a few kilometers around the mine.
Thus, for lead and in the case of smelters, high exposure ranges have been associated in the literature with distances from the point source of emissions of 0.5 to 4km. In highly exposed communities, mean blood lead levels (PbB) among children ranged from 13 μg/dL to more than 40 μg/dL.

(Table 3) In all cases, even at the mean, PbB exceeded the reference value of 5 μg/dL (the 97.5$^{\text{th}}$ percentile of blood lead levels found in the U.S.) set by the Centers for Disease Control to "trigger lead education, environmental investigations, and additional medical monitoring," (CDC, 2012) as well as the laxer and more dated 'level of concern' of 10 μg/dL (Roper et al., 1991).

In this paper, we do not directly observe environmental pollution or the body burden of toxicants. Rather, we use distance to the nearest mine as a proxy. The choice of a distance cutoff to define the treated group is therefore crucial. In line with the empirical evidence reviewed above, we look for health effects in a tightly defined treatment group, and consider only households within no more than five kilometers of a mine to have been exposed. This choice also corresponds to the extent of high-exposure buffer zones around mines in van Geen et al. (2012). It is considerably tighter than in other current studies of mining in economics, as is appropriate for our focus on health impacts.[3] A key benefit of working with our large multi-country dataset is that it allows us to restrict our treatment group in this manner, while retaining sufficient statistical power.[4]

3.2 Pathophysiological and clinical effects of lead and other metal exposure

As noted, the mines in our sample are associated with characteristic sets of pollutants. Because the latter are known to cause specific health effects, we can develop predictions for expected health impacts that are well-grounded in scientific knowledge. To the degree that we find expected health impacts, but not others, we strengthen our case that impacts are likely due to environmental pollution, rather than any other mechanism.

In our baseline investigation of health impacts, we do not distinguish between different types of mines. Yet, our main concern is with the health consequences of environmental contamination with heavy metals, and in particular, with lead. We focus on heavy metal contamination, first, because the health impacts of exposure are important and observable in our data, and second, because a large share of mines in our sample is associated with this type of pollution (40% of mines in the cross-section, and 70-90% in the panel, depending on definitions). Among heavy metal pollutants, lead takes a central role, because it is known that the lead body burdens previously measured near smelters (reported above) are high enough to cause health problems that we observe in our data.

[3]Wilson (2012) uses a cutoff of 10km, while Aragón and Rud (2013a,b) and Kotsadam and Tolonen (2013) use a baseline cutoff of 20km, with sensitivity analysis for other choices.

[4]With perfect data, we might define closeness even more restrictively. In the context of available data, a tighter cutoff would risk introducing noise, both because of the practice of jittering cluster geolocations in our socio-economic data, and because of the fact that we work with (imperfectly recorded) mine point locations, while mining operations can measure several kilometers across.

3.2.1 Sequalae of lead exposure observed in our data

The toxic properties of lead have long been studied, and are well understood.[5] The wide-ranging effects on adults include reduced blood hemoglobin (Hgb) and overt anemia, cognitive defects, hypertension, and impaired renal function. In our data, we observe only one of these conditions, namely low blood Hgb and anemia. We adduce two additional unspecific health outcomes as falsification tests, namely miscarriage and general grave illness.

For children under five years of age, we analyze two health outcomes that have previously been linked with lead exposure: anemia and growth retardation. We use for falsification tests health outcomes that have not been linked to lead (cough, fever), or linked only weakly or at very high exposure (gastro-intestinal problems and mortality). Regrettably, we do not have a good measure of impaired cognitive performance and behavioral problems due to neurological damage in children. However, while the health impacts we do observe – anemia and growth deficits – are known to require high blood lead, "there is no evidence of a threshold for the adverse consequences of lead exposure" for intellectual development (Lanphear et al., 2005, p. 899). Hence, demonstrating overt anemia or growth deficits implies a strong likelihood that the affected individuals – and presumably, others with lower PbB – also suffer some cognitive and behavioral impairment. Appendix E discusses how the health consequences we observe affect the well-being of those exposed, and what their economic cost might be.

3.2.2 Hematologic toxicity of lead

Lead depresses blood Hgb levels both by shortening red blood cell life spans, and by interfering with enzymes essential to the synthesis of the heme component in hemoglobin. Enzyme activity begins to be disrupted at very low PbB, but is not measured in our data. Effects on Hgb – which we can observe – have previously been reported at high PbB levels: in excess of 40μg/dL in children, and 50μg/dL in adults (ATSDR, 2007, pp. 69, 71f). That is, we expect the hurdle to finding impacts on Hgb to be quite high.

Therefore, we devise an additional, more sensitive test of hematotoxic effects. We build upon the insight in Grandjean et al. (1989) that, even when lead exposure is too low to reduce Hgb *levels* in adults, "increased demand on the formation of blood following blood loss could result in a delayed blood *regeneration* in individuals exposed to lead" (p. 1385 - our emphasis). Grandjean et al. demonstrate this effect by comparing Hgb recovery after blood donation in lead factory

[5] See ATSDR (2007) for a full discussion.

workers and a control group. In our study, we show that analogously, Hgb recovery is similarly impaired among women in mining communities after another kind of blood loss, namely pregnancy and delivery.

The effect of lead on children is of particular concern, since children are both more sensitive in their reaction to body burdens of lead, and (at least in the case of lead ingested with food) absorb far larger portions of lead. In the case of anemia, however, we expect effects to be *harder* to demonstrate in children than in adults. This is because, by contrast to adults, children are able to compensate for erythrocyte loss by increasing production of the hormone erythropoietin (EPO), and thus boosting the generation of red blood cells. (Factor-Litvak et al., 1998)

In summary, based on the state of scientific knowledge, we expect Hgb in residents of mining communities to be measurably affected only if there is substantial exposure to environmental lead. An effect should be detected most easily in the recovery of Hgb after blood loss, followed by Hgb levels in adult women, and least readily in Hgb levels in children.

3.2.3 Effects of lead on child growth

While there is an epidemiological link between lead and anemia, and several hematotoxic mechanisms are known, studies are in less agreement on the effect of lead on growth in children, and "the mechanism by which lead may reduce a newborn's size is unknown." (Hernandez-Avila et al., 2002, p. 486)

Correlations have been observed – including at moderate PbB on the order of 10μg/dL – between maternal or child blood lead and gestational age, as well as a wide range of measures of height and weight from birth to adolescence. (ATSDR, 2007; Bellinger et al., 1991; Hernandez-Avila et al., 2002; Sanín et al., 2001; Zhu et al., 2010) However, other studies have failed to show such correlations; indeed, it is common for a study to find impacts on some dimension of growth, but not on others, with no conclusive pattern of which indices are sensitive.

In this paper, we seek to exclude both endogeneity and small-sample variation as potential sources of ambiguous results. However, while we are able to show that *in utero* exposure affects one dimension of growth (height for age), our results mirror the existing evidence in that we find no clear effects on another key measure of growth (birth weight). In addition, in our study sites, growth effects are concentrated among infants, but abate in older children. As context for this finding, we note that, while, as a stylized fact, "blood lead levels [peak] in the age range of 1 to 3 years" (Bellinger, 2004, p. 1017), there is an important earlier path of exposure, through transfer of lead from the mother's body through cord blood and breast milk. Indeed, "infants are born with a lead body burden that reflects the burden of the mother," (ATSDR, 2007, p. 223) with correlations as high as 0.8 between maternal and infant PbB (Lauwerys et al., 1978, p. 280). Finding health impacts among infants is therefore particularly plausible if there is evidence of significant maternal lead burdens.

4 Data

4.1 Socio-economic and health data

We obtain socio-economic and health data by pooling all 104 geo-coded Demographic and Health Surveys (DHS) available (at time of writing) from countries for which we have mining data. This yields a dataset of repeated cross-sections covering 44 countries, with a total of 1.2m households, and several million individual records. About 170,000 households are within no more than 20km of a mine recorded in our data, and enter our analysis. (Table 4) Their location is shown in Figure 1.

The DHS data has some notable strengths: it covers a very broad range of developing countries; surveys have been conducted for nearly 30 years; individual surveys are fairly comparable; sampling cluster geocodes are available for many survey rounds; and there is strong data on maternal and under-five health, including anthropometrics and specifically, hemoglobin (Hgb). These features currently make DHS an obvious choice to study health and development at the micro level across multiple countries.[6]

However, the data also has some important limitations with implications for our work. (i) There is relatively little data on socio-economic status, no information on wages, and little information on employment. We therefore work with an asset index, rather than more direct measures of wealth or of income, and discuss employment outcomes only in passing. (ii) Because the surveys have kept changing and improving, very few indicators of interest to us were collected in all surveys. Indeed, working with the largest set of observations for which all indicators are available is impractical, because the number of observations is very small. On the other hand, estimating results on pair-wise common sets would lead to tedious repetition. We seek to strike a balance, and present side-by-side comparisons for core results. (iii) Finally, we stress again that the data are cross-sectional.

[6] Other data with high coverage that include both health and socio-economics are either less rich (IPUMS), or less harmonized (LSMS).

Therefore, while our difference-in-difference tests are designed to yield evidence of causal effects, they always compare across different individuals.

Our core measure of wealth is a standard asset index computed over household durables and housing characteristics. (Filmer and Pritchett (2001); see Appendix B for details.) We base it on the largest set of wealth proxies available within each survey round, but do not include slow-moving or immutable traits of the household head, such as gender, marital status, or education.

We obtain from the DHS detailed data on health among children below five years of age, and among women aged 15-49 years. There is little information on older children, men of any age, and women aged 50 years and over. Our core health indicators are blood Hgb levels and an age-adjusted height index. Hgb is adjusted for altitude, and expressed either as a continuous measure in units of grams of hemoglobin per deciliter of blood (g/dL), or as a binary indicator for the clinical condition of anemia, associated with blood Hgb below 12 g/dL in non-pregnant women, and 11 g/dL in pregnant women and in children (World Health Organization, 2011). Following standard practice, height is expressed as the difference between a respondent's height and the age-group median, normalized to standard deviations. We normalize using the median and standard deviation provided by DHS (alternative normalizations make no empirical difference). We consider the continuous height measure, as well as the clinical outcome of stunting (severe stunting), defined as a height of at least two (three) standard deviations below the median.

In addition to our core outcomes, we collect data on a range of general adult and child health outcomes, on health care, sexual risk taking, nutrition, and employment and occupation. Finally, we construct infant and under-five mortality data for all children whose births were recorded in any survey module.[7]

4.2 Mining data

We draw upon four data sources for information on the location and characteristics of mines and mineral deposits. These include a very large cross-sectional dataset that allows us to make meaningful claims about the mean effect of mining across many developing countries; two datasets of mine output that permit us to estimate panels; and an additional dataset of mine locations that serves to ensure robustness of our findings to measurement error in geo-location. In total, we observe communities near 838 mines in the cross-section, and 515 mines in the panel, though the set of mines that enters our estimating samples is generally smaller.[8] (Table 4)

[7]Because we construct these variables from birth records of all children ever born to the women in sample, the mortality variables must be interpreted as being conditional on the mother's survival until the time the survey was taken.

[8]Nearly all of those mines enter into our model when we use state-level effects (see Table 1). The number of mines near which we observe at least one community within 5km (treatment) and one within 5-20km (control) is lower, with

4.2.1 Cross-sectional data on mine location and characteristics

In the cross-section, we work with the United States Geological Survey's Mineral Resource Database (United States Geological Survey, 2005). It contains the point locations of a very large set of mines, legacies, deposits, and smelters (about 25,000 locations in total) across developing countries. The data records geological information and some basic description of the nature of the mine for a substantial subset of entries. However, there is no data on production, and start dates and status of operation are only available for very few mines.

In our baseline cross-sectional sample, we include all active mines, legacies (that is, former mines that are now dormant), and smelters. We include smelters because they are often located close to mines, and it is intuitive to think of a single mineral extraction and processing chain from mining to smelting.[9] We include legacies, because the cross-sectional data gives us little guidance in defining whether a mine was operational during a given survey round. The resulting treatment definition should be thought of as yielding 'the effect of living in a location ever exposed to mineral mining or processing'.

We extensively parse information on the types of minerals present in a given location to sort mines into larger groups that share the same expected pollutants, and hence, the same health effects. We remove from our baseline sample all quarries (see Appendix A for a definition). We do so because we seek to study the welfare impacts of mining as an industry that generates very high value added, but is potentially severely polluting. Quarries differ from mineral mines in both respects, at least as a matter of degrees. As we have argued above, we are especially interested in polymetallic mines near which we expect pollution with heavy metals, and particularly with lead. For the purposes of the present paper, we define a mine to be a 'heavy metal' mine if (i) lead is being mined or smelted, or (ii) lead, though not targeted for extraction, is known to be present in significant amounts, or (iii) any two of the metals copper, gold, silver, and zinc are being mined or processed. This definition is necessarily imprecise, but gives due recognition to the special role of lead, and seeks to exclude metal mines with different pollutant characteristics. For instance, among gold-producing mines, it would aim to exclude alluvial gold deposits, where gold is typically the only metal of interest, and we might expect mercury contamination from processing to be the primary concern, rather than lead pollution.

226 mines in the cross-section, and 175 in the panel. These are the mines that enter into our baseline mine-effects models. IV models work with larger sets of (mined and un-mined) deposits.

[9] In Appendix Table G, we show that our core results are nearly fully robust to excluding smelters. In one case (panel results on women's hemoglobin), the effect is not significant, although it is consistent in sign and approximate size.

4.2.2 Mine-level production data

Since the USGS data provides virtually no time variation, we draw additional information from two business intelligence firms: Infomine (2013), and IntierraRMG (2013) – for whose product we henceforth write ‘RMD’, for ‘Raw Materials Data’. Both sources record dates of operation and annual production, alongside diverse additional characteristics of the mines. Most mines included in the Infomine data are also available in the RMD data, but not vice versa. We therefore work with RMD as our basic data, and add those Infomine entries that are not also contained in the RMD data. RMD mines are more homogenous than those in the USGS sample: most of them are large mines, and most of those close to DHS clusters are metal mines. While the set of mines included is far smaller than for the USGS data, coverage of large mines is quite comprehensive, and the mines recorded in the dataset account for a very large share of global metal production. For instance, they account for around 80% of global gold production and 80-90% of global iron ore production in the most recent decade for which data is available.

Because there is some question as to the precision of geolocations recorded in the RMD data, we use mine geolocations from an additional dataset, Mining Atlas (2014), for three purposes. First, we add geolocations for RMD mines wherever location is missing in the original data. Secondly, we use company records and Google Earth images to investigate the small number of cases where there are very large discrepancies in location between the two sources; we discard a few records where location is plainly not recorded with any precision in either dataset. Thirdly, we use the two independent but noisy measures of location to check robustness of our results to measurement error in geolocation (see Appendix D).

4.3 Other data

For the purpose of constructing a time-varying instrumental variable, we retrieve data on mineral prices from various sources, listed in Appendix A. In order to describe how the wealth effects of mining vary with the economic environment, we obtain country-level data on GDP and governance from the World Development Indicators; data on the efforts a given country made toward compliance with the Extractive Industries Transparency Initiative (EITI) from the Initiative’s website (www.eiti.org); and state-level data on governance, geography, infrastructure, and education from Gennaioli et al. (2013).

5 Econometric Specification

5.1 Baseline treatment definition

We define exposure to mining as being geographically close to a mine in the cross-section, and as closeness interacted with the mine being active in the panel. This choice is immediate for the study of economic impacts: with transport and search cost, distance is the treatment of interest. For the purpose of studying health impacts, distance acts as a proxy for pollution – which we do not observe.

We define a cluster as being 'close', and hence, 'treated', when it is within five kilometers of the nearest mine. We will also refer to this as the 'direct vicinity' of the mine. We define a cluster as being in the control group when it is within 5-20km of the nearest mine. We will refer to this as the 'general vicinity' of the mine. As noted above, we bound our treatment group tightly, to enable us to detect health impacts within the region in which pollution is likely to occur. Bounding our control group conservatively greatly eases the stringency of assumptions required for a causal interpretation of our results. One need only consult maps of mining areas to confirm that over distances as substantial as 40km or 200km – as used elsewhere – many things other than closeness to mines change, whether in the natural and the built environment, or in institutions. The cost of working with these definitions is that we can only achieve viable sample sizes by allowing our panels to be unbalanced. We argue that this is a reasonable price to pay for the sake of working with a treatment definition that is in line with prior scientific knowledge, and a control group definition that promises to provide a credible counterfactual.[10]

In the panel, we define mining activity as a dummy variable taking value one when the mine had non-zero output, and value zero when the mine was known to have had zero output. (We conservatively impute inactivity – see Appendix A.) That is, we consider only extensive margin impacts of production. We do so because year-on-year variation in output is likely to be more weakly associated with health outcomes. In this, mines differ from sources of pollution studied elsewhere. Extracting minerals from the ground, breaking them up, and processing them generates a flow of pollution. At the same time, however, the stock of tailings dumped after processing will in many cases continue to pollute. The exact time pattern of pollution is thus hard to predict, but is bound to lie somewhere between a pure flow and a pure stock problem. We hope to do it justice by studying extensive margin variation alongside the cross-sectional 'once on, always on' measure.

[10]For the study of wealth benefits alone, a natural alternative would be to study effects of mine density in (hopefully quite balanced) panels of administrative units. This would, however, vitiate the purpose of studying health effects.

5.2 Cross-sectional model

Identification in the cross-section rests on a conservative choice of control group, and restrictive fixed effects. Because they cannot decisively address the possibility of residential sorting, the correct way to read our cross-sectional results is to view them as the long-run effect of mining on 'mining *communities*', much as a district or county-level study estimates effects on those units. As such, we believe they can be interpreted as causal; and to the degree that regional disparities matter, they are of policy interest. Our difference-in-differences models then provide evidence that impacts are unlikely to be driven by sorting, and allow us to make stronger claims about the well-being of 'people exposed to mining'.

In our baseline specification, we consider outcomes y for individuals or households i in sampling cluster j within no more than 20km of a mine, conditional on whether the cluster is *close* (within 5km) to a mine, and conditional on other covariates X. Because distance is measured between mines and sampling clusters, the treatment varies at the cluster level, not the individual level. Covariates always include an indicator for whether the cluster is in an urban or rural setting, and some appropriate measure of the age of the respondent, the respondent's mother, or the household head. Because DHS conducts repeated cross-sections, our model allows for repeated measurements of effects near the same mine, while accounting for year-specific effects in each round of measurements. We therefore use common effects γ for all observations near the same mine surveyed in the same year (mine-year effects), and account for residual correlations by clustering error terms at the mine level (not the mine-year level). Wherever the outcome of interest is binary, we model it using a linear probability model.

$$y_i = \beta_1 close_j + \beta_2 X_i + \gamma_{mine-year} + \epsilon_i \tag{1}$$

Identifying assumptions would be violated if mining towns differed from neighboring communities in geography, institutions or other characteristics in ways that correlate with potential outcomes. However, differences would have to arise even compared to locations very close by, because we restrict control locations to those no more than 20km away from the nearest mine. Identification is also only affected by such differences if they are not in some way due to the presence of the mine in long-run equilibrium (for instance, through changes in infrastructure or institutions). Conversely, we will need to offer further evidence before we make claims about the mechanisms – crucially, pollution – by which mines affect well-being.

5.3 Pseudo-panel model

We have argued that our cross-sectional setup offers valid estimates of the long-run impact of mining on communities. Still, it says less than is desirable about mechanisms of treatment transmission, and due to the possibility of sorting, it does not allow us to make claims about the impact of mining on individuals. An immediate way of addressing both challenges is to construct pseudo-panels from the repeated cross-sectional DHS surveys. We construct these in two ways. Firstly, we compare observations from households surveyed at different times, but near the same mine ('mine-level panel'). Secondly, we compare children born to the same mother at different times ('mother-level panel'). Plainly, comparisons in each case are across different individuals.

Equations 2 and 3 describe the mine-level and mother-level models. We analyze outcomes for individuals i in cluster j at time t.

$$\begin{aligned} y_{i(t)} = \beta_1 close_j + \beta_2 operating_{j(t-\tau)} + \beta_3 close_j * operating_{j(t-\tau)} \\ + \beta_4 X_{i(t-\tau)} + \gamma_{mine} + f(t) + \epsilon_{i(t)} \end{aligned} \tag{2}$$

$$\begin{aligned} y_{i(t)} = \beta_1 operating_{j(t-\tau)} + \beta_3 close_j * operating_{j(t-\tau)} \\ + \beta_4 X_{i(t-\tau)} + \gamma_{mother} + f(t) + \epsilon_{i(t)} \end{aligned} \tag{3}$$

In Equation 2, we allow for time-invariant effects γ_{mine} for each mine, and model outcomes at time t as being conditional on whether the respondent lived in a community *close* to a mine during the time period relevant for treatment, $t - \tau$, and whether the mine was *operating* at time $t - \tau$.[11] The time periods of interest t and $t - \tau$ depend on the outcome being investigated. For instance, where we analyze height-for-age in children, the outcome is measured in the survey year t, and may be modeled conditional on exposure to mining operations during the survey year (τ = 0), the birth year (τ = age), or while the child was *in utero* (τ = age + 1). The model also includes time-specific effects $f(t)$. We believe country-year dummies are sufficiently flexible and appropriate for sample size. We use these in our baseline models, and show robustness to using different time effects.

[11]For each respondent in our sample, we only observe current residence, and how long the household has been resident there. We have no information on previous residence. Therefore, the panel is inherently restricted to respondents who have lived in the location where they were surveyed for at least τ years. (Although they may have moved to their present location at a time before $t - \tau$.)

Modifications in the mother-level model are immediate (Equation 3); because we do not observe location of prior residence for migrants, no coefficient on *close* can be estimated, and because of the much smaller sample sizes in the mother-level model, we include country-specific linear trends $f(t)$ in our baseline model. We view the two panels as complementary: the mother-level panel controls well for unobservable characteristics, and is mostly impervious to sorting. Mine-level estimates require somewhat stronger sorting assumptions, namely, that sorting is sufficiently slow relative to the frequency at which the outcomes of interest are measured; however, they rest on larger samples.

5.4 Difference-in-differences tests tailored to the health conditions studied

For some indicators, our sample is small near mines where there is production information, so that the pseudo-panel tends to be highly unbalanced, in particular when using mother-level fixed effects. We therefore leverage the scientific understanding of the health conditions of interest to our study to construct additional difference in differences tests. Like the pseudo-panel, they compare the impact of mining across groups that are and are not expected to show effects. However, unlike the pseudo-panel, they do not rely on the use of time-varying production data, and hence, tend to preserve sample size better. Because they each build upon a different insight into the likely nature of exposure and the organism's reaction to it, they generate distinct control groups, and hence, further "reduce the importance of biases or random variation in a single comparison group" (Meyer, 1995, p.157).

Mine types: Firstly, we make use of the fact that, as discussed above, distinct mine types are associated with specific pollutants and health effects. This allows us to contrast differences across distance groups near mines where an effect is expected, and near mines where none is expected, as in Equation 4. (The effect of *heavy metal mine* alone is collinear with mine-year effects.)

$$y_i = \beta_1 close_j + \beta_2 heavy\ metal\ mine_j + \beta_3 close_j * heavy\ metal\ mine_j + \beta_4 X_i + \gamma_{mine-year} + \epsilon_i \tag{4}$$

Identification rests on the assumption that potential outcomes vary among those close and not close to the mine in similar ways near mines of different types. Most obviously, if wealth effects varied systematically among mine types, health results might be confounded. With respect to preference-based sorting, the assumption would be violated if respondents were aware of how mine types differ in health outcomes, and sorted accordingly. We address the issue in two ways. Firstly, we compare DiD results on health to those on wealth, and show that differences arise for health outcomes, but not wealth. Secondly, we show that there are DiD effects only on specific expected health outcomes, not general health.

Timing of birth relative to migration: Secondly, we use the fact that respondents who have long been resident in mining communities are likely to have been exposed to more pollution. Specifically, we leverage differences in exposure *in utero*, and compare among children i born to migrants before their family moved to mining communities j, and those *conceived after move* (and contrast this with the same statistic observed among those who migrated to locations slightly farther away from the mine). (See Appendix A for further detail.)

$$\begin{aligned} y_i = \beta_1 close_j + \beta_2 conceived\ after\ move_i + \beta_3 close_j * conceived\ after\ move_i \\ + \beta_4 X_i + \gamma_{mine-year} + \epsilon_i \end{aligned} \tag{5}$$

The identifying assumption is hence that potential health outcomes do not vary systematically with the exact timing of birth relative to the move, comparing those who moved to mining communities to those who moved to communities farther away. It would be violated by sorting if, for instance, pregnant women were more likely to hold off on moving when they are about to move to mining communities, perhaps because they are weary of pollution, or conversely, if pregnant women were more likely to speed up relocation, perhaps because they hope for good economic opportunities.

Maternal Hgb recovery: Finally, we develop a DiD test based on the observation that in lead-exposed adults, the recovery of Hgb after blood loss is even more readily affected than the steady-state level of Hgb. As discussed above, this result was previously proven by studying Hgb recovery after donating blood. Of course, we cannot identify blood donors in our sample. We do, however, observe one population group that experiences dramatic drops in Hgb: women who are pregnant, or have recently given birth. This allows us to formulate a test that asks whether differences in Hgb between women i in mining and control communities j are particularly stark during pregnancy and postpartum. In our preferred specification, we estimate the model with state-year effects, since the number of women we observe within the time period of interest is borderline too small for allowing for mine-year effects. (We discuss identifying assumptions and extensive robustness checks below, in Section 6.4.)

$$y_i = \beta_1 close_j + \beta_2 pregnant\ or\ postpartum_i + \beta_3 close_j * pregnant\ or\ postpartum_i + \beta_4 X_i + \gamma_{state-year} + \epsilon_i \tag{6}$$

5.5 IV models

Finally, we use both cross-sectional and panel IV strategies to study wealth effects. Our purpose for the IV estimates is somewhat narrow: they provide reassurance against endogenous choice of location (even within 20km) in the cross-section, and endogenous decisions to produce in the panel. However, we have argued that residential sorting is the key concern in our baseline cross-sectional model, and to a somewhat lesser degree, in our mine-level panel. IV estimates do not address this issue. We therefore discuss them relatively briefly, and for wealth results only – for health impacts, we instead rely on the additional difference-in-differences tests described above, because they address sorting, and preserve sample size better.

5.5.1 Cross-sectional IV

In the cross-section, to instrument for whether a cluster is within 5km of a mine, we use the dummy (Wald) instrument *deposit* that simply indicates whether there is a mineral deposit within 5km of a given cluster (Equation 7).[12] The sample is restricted to clusters within no more than 20km of a deposit.

$$\begin{cases} y_i = \beta_1 close_j + \beta_2 X_i + \gamma_{state-year} + \epsilon_i \\ close_j = \phi deposit_j + \delta_{state-year} + \eta_j \end{cases} \tag{7}$$

Because coverage of deposit locations in the cross-sectional data is very broad, we can think of our IV estimates as general population effects. Because there can be no mine without a mineral deposit, there are neither 'defiers' nor 'always-takers', and we can interpret IV estimates as the effect of treatment on the treated. (Imbens and Wooldridge, 2009) Unsurprisingly, the dummy instrument is exceedingly strong. Since the true global distribution of mineral deposits is exogenous to human activity, the instrument is also exogenous, as long as there is no preferential *prospecting* for minerals. We believe this is likely the case, since all anecdotal evidence suggests that mining companies will seek out promising deposits in virtually any location, regardless of geographic or

[12]This is similar in spirit to the geographic instrument in Duflo and Pande (2007).

political obstacles. We also believe that the instrument satisfies the exclusion restriction. The most likely violations would be due to topographical features such as land quality, gradient, or water availability. Because we work at small spatial scales and across many countries, potential violations are hard to test directly. Yet, since we strongly restrict our analysis in space, characteristics would have to vary systematically over small scales to cause any problems.

5.5.2 Panel IV

Our cross-sectional IV strategy extends very naturally to the panel setting, by interacting the presence of mineral deposits with world minerals prices. Our panel data does not have very high coverage of mineral deposits, but it does include some deposits that are being explored or prepared for exploitation. We adjust the panel IV sample to include such deposits. Hence, in Equation 8, we treat the variable *deposit* that records whether cluster j was within 5km of any deposit as exogenous.

$$\begin{aligned} y_{i(t)} = \beta_1 deposit_j + \beta_2 operating_{j(t-\tau)} + \beta_3 deposit_j * operating_{j(t-\tau)} \\ + \beta_4 X_{i(t-\tau)} + \gamma_{mine} + f(t) + \epsilon_{i(t)} \end{aligned} \tag{8}$$

We then instrument for whether the mine was *operating*, and for the interaction of closeness and operating status, using world mineral prices *price*, and their interaction with *deposit*. (See Appendix A for a full description of the instrument.)

6 Results

6.1 Effects on wealth

Mining towns are wealthier than neighboring communities, both in the long run and the medium term

Households in mining communities are at the mean considerably wealthier in terms of asset ownership than control households. The magnitude of the cross-sectional effect at the global average is on the order of 0.11 standard deviations of the asset index. (Table 5) In the mine-level panel, the DiD coefficient on the effect of living close to a mine in a year when it is operating is 0.26

standard deviations of the asset index in our preferred specification. (Table 6, Column 2) Since survey rounds are generally about five years apart, we interpret this as a medium-term effect.

The effect size is considerable, given that in the countries in our sample, there is generally great within-country variation in asset ownership. In the linear index, the magnitude of the cross-sectional effect is comparable to that of owning a car or motorbike in the case of Peru in the year 2000, and to the effect of owning a radio or a watch in the case of Burkina Faso in the year 2010. The panel effect is comparable to the impact on the index of having an electricity connection or living in a dwelling with finished flooring in the case of Peru in the year 2000, and to the effect of owning a motorbike or mobile phone in the case of Burkina Faso, in the year 2010. (See Appendix B for a description of the index and for examples of factor loadings.)[13]

We argue below that, because of the spatial pattern of long-run wealth effects, the cross-sectional baseline estimate should be interpreted as a lower bound. In Appendix C, we show that our unweighted baseline estimates are smaller than estimates obtained by (i) weighting each mine equally, or (ii) weighting by estimates of the mine-year population. In Appendix D, we use two independent measures of the geolocation of mines to instrument with one distance measure for the other, and show that our baseline results likely carry substantial attenuation bias – in our preferred specification, some 18% of the estimate. Cross-sectional IV estimates yield results that are close to and not statistically different from both our baseline results, and OLS estimates on the IV sample. Panel IV estimates are larger than the benchmark, but not significantly different. (Table 7) We also note that our cross-sectional and mine panel results cross-validate each other quite well. In the panel, the single-difference coefficients in distance tend to be weakly negative. Thus, communities within the direct vicinity of mines were wealthier than those in the more general vicinity only when the mine was operating. Indeed, because we observe communities near inoperative mines mostly *before* the mine ever reports production, we can to the first order say that mining communities were wealthier only *after* the mine began operating. This is reassuring for the validity of our cross-sectional analysis.

We have argued that, if the object of interest is the effect of mining on household welfare, rather than on the spatial distribution of wealth, the most salient identification concern in the cross-section is residential sorting. Panel results can be presumed to be more robust, but with about

[13] Regrettably, the DHS surveys have no wage data, and limited coverage of employment. The number of men living near mines in our sample for whom employment data was collected is small. In consequence, an in-depth analysis of effects on these core dimensions of welfare is not possible. Appendix Table L.3 shows that, in the cross-section, unemployment among men is virtually unaffected, consistent with long-run general equilibrium. As is intuitive, the sectoral share of agriculture decreases alongside ownership of agricultural land. In the panel, employment effects tend to be adverse in sign – consistent with queuing – but the estimates are noisy and not stable. We refer the reader to Kotsadam and Tolonen (2013) for a detailed discussion of effects on women and sectoral shifts in sub-Saharan Africa.

five years between survey rounds, there is still the possibility that sufficiently rapid sorting could influence results. We therefore separately study results for households that report never having moved from their current location. Effects are somewhat smaller and weaker (if not significantly different) among never-movers in both the cross-section and the panel. (Table 5, Column 2; and Table 6, Columns 5-7) We interpret this as limited evidence of sorting of migrants with better potential socio-economic outcomes into mining communities, or sorting of previous residents with better potential outcomes out of mining communities.[14]

Spatial extent of the wealth effect

Wealth effects decay steeply with distance to the nearest mine. In the panel, effects are limited to those living within 5km; in the cross-section, there is a gradient in wealth up to a distance of 15-20km. (Figures 2 and 3) That is, in the long-run, communities in the general vicinity are economically affected to some degree, although less so than those in the direct vicinity. Hence, the cross-sectional treatment effect in our baseline model is smaller than the wealth effect on the direct vicinity of mines, as compared to those living *outside* of the general vicinity, within 20-40km of a mine (0.4σ – results not shown). Conversely, it is larger than the *average* effect of living either in the direct or general vicinity of the mine, as opposed to living at 20-40km (0.05σ).

The difference in spatial patterns between the cross-section and the panel allows for a number of explanations. If both patterns are well-identified, one would argue that the discrepancy reflects the contrast between medium-term and long-run impacts, with further diffusion of wealth effects over time. If we were not convinced of identification in the cross-section, we might feel that the pattern suggests that mines tend to locate in places that are already wealthier than their surroundings. (Of course, panel results come from a smaller sample, and simply might be more attenuated.)

We note that, even in the cross-section, the estimated spatial extent of treatment effects is smaller than in the case study analyzed in Aragón and Rud (2013b, p. 26), who find "positive and significant [income effects] for households located within 100km of Cajamarca city," the community closest to the mine studied. The discrepancy could be due to the fact that Aragón and Rud study a policy change that can be presumed to be very favorable for local welfare; or the fact that they consider the case of a very large mine in a region with reportedly high transport cost. In addition, Aragón and Rud have income data available; presumably, a more sensitive measure of well-being than our asset index.

[14]For background, we note that there is only weakly more migration in mining communities than in neighboring communities. However, in both mining and control communities, the share of migrant households is very high: around 60% of households migrated at some time, and about 23% migrated within the five years preceding the survey. Sorting could therefore easily explain cross-sectional differences, if the characteristics of migrants (including those unobserved households who left the communities) are sufficiently different.

Effects on the distribution of asset wealth

Mining is associated with wealth benefits across the distribution, though in the long run, there are much higher gains for the top quantiles, and a mild increase in wealth inequality. Benefits are more evenly distributed among never-movers. The distributional pattern might, for instance, reflect slow sorting of high-income households into mining communities, or the gradual emergence of economic opportunities that are open only to a select few.

Quantile regressions suggest that closeness to mines raises long-run asset wealth quite evenly across the distribution, with effect sizes for most quantiles close to the mean effect. (Figure 4 – see Appendix A for details on implementation.) That said, the top 5-10% benefit the most, with gains about three times as large as those at the median. Gains at the top are more limited among never-movers. In the panel, if anything, benefits are progressive, and the top quantiles gain less than others (Figure 5); this pattern is comparable to the distribution of income effects found in Aragón and Rud (2013b).

Secondly, we directly consider effects on a simple measure of within-cluster inequality, namely the absolute deviation of a household's asset index value from the cluster mean.[15] In the cross-section, the mean absolute deviation increases moderately among all households, by 0.03 standard deviations of the asset index, or one-fourth of the cross-sectional wealth effect. (Table 5, Column 3) There is no effect among never-movers, nor in the panel. (Table 6, Column 4)

Correlates of long-run effects across countries

Long-run wealth effects vary greatly across mining communities. Table 8 shows correlations of mine-level wealth effects with measures of the larger economic, geographic and policy environment. Gains are greatest where the economic environment is weak, across a range of indicators – GDP, education, access to infrastructure, some dimensions of remoteness, and (directionally only) measures of institutional quality.[16] While these correlations cannot be interpreted as causal relationships, they raise the question whether the local economic effect of mining might be driven not by the interaction of mining with other economic activity, but by the opportunities mining provides in areas where there is a paucity of other options.[17] With the same caveat regarding causal interpretation, we also note that we do observe stronger wealth effects in surveys conducted in countries

[15]This simple index seems more appropriate than more familiar inequality indices both due to the small number of households in many clusters, and to the nature of the mean-zero standardized asset index.

[16]Appendix M shows the distribution of treatment effects across world regions and countries; correlations with measures of overall development empirically supersede regional patterns.

[17]We emphasize that, because we study effects purely at the local level, the correlation between local benefits and a weak economic environment cannot be read to contradict findings from the resource curse literature. Our findings have no implications for whether, beyond the local level, resource revenue creates corrupt structures or drives Dutch disease.

at a time when the country had completed a report for the Extractive Industries Transparency Initiative,[18] or (weakly) when it had participated in the EITI in any way.

6.2 Evidence of hematologic toxic effects

We have argued above that exposure to lead among residents of mining communities may affect the hematopoietic system and reduce red blood cell survival. In the DHS data, we observe only a single indicator of potential hematologic toxicity – blood Hgb concentrations. As argued in Section 3.2.2, we would expect most strongly to see a reduced ability to *recover* from blood loss in adults, perhaps alongside depressed Hgb levels. In children, we might expect to see reduced blood Hgb levels, though in the age group we observe, children are likely able to compensate for lead exposure. Our results confirm this expectation: we find strong evidence of lower Hgb levels and slower Hgb recovery after blood loss in adult women, and weaker evidence of lower Hgb levels in children.

Hemoglobin levels in adult women are strongly depressed in mining communities

In the cross-section, blood hemoglobin (Hgb) levels are depressed among women living in mining communities by about 0.09 g/dL. The effect among never-movers is larger (0.13 g/dL), consistent with longer exposure to environmental lead, although (on this smaller sub-sample) it is just below significance ($t = 1.56$). Considering directly the clinical outcome of anemia, we find that prevalence is significantly elevated by three percentage points among all households, and by five percentage points among never-movers (Table 9). Appendix M shows the distribution of mine-level effects across countries.

Panel results confirm these patterns. Point estimates are larger, with DiD coefficients of a 0.33 g/dL decrease in blood Hgb, and a ten percentage point increase in the incidence of anemia in our preferred specification. (Table 10, Columns 3 and 6) A number of causes could account for the larger point estimate in the panel; notably, the share of metal mines associated with lead pollution is higher in the panel sample (and, as we show below, the treatment effect is concentrated near such mines). In the long-run, there might also be more adaptation to avoid pollution.

The size of the effect on Hgb levels can be compared, for instance, to changes in Hgb on the order of 1g/dL associated with treating anemic pregnant women with a course of iron supplementation (Sloan et al., 2002). That is, we obtain a general population effect estimate on the order of

[18] See www.eiti.org. The EITI describes itself as "a global coalition of governments, companies and civil society working together to improve openness and accountable management of revenues from natural resources.'

one-tenth to one-third of the effect of a targeted intervention in a highly susceptible population. Another point of comparison is the drop in Hgb during pregnancy and the first year post-partum, estimated in our sample to be on the order of 0.44 g/dL (compared to women who gave birth two or three years ago, and among women living at least 20km away from any mine). The increase in the incidence of anemia is a large effect in absolute terms, though it must be seen in the context of a baseline proportion of anemic women of about 36% in control locations. That is, the cross-sectional effect amounts to a 7% relative increase in incidence, and the panel effect, to a 27% relative increase.

We note that the single difference coefficient in distance suggests that when the mine is not operational, residents of mining communities have higher Hgb levels than the control group. This is perhaps surprising, given that our wealth results showed a zero or weak negative effect in mining communities when the mine is not operational. (Table 6) However, it further reassures us against any concerns that geographic features, for instance altitude, might be driving cross-sectional results.

We adduce two additional tests, both to further bolster identification, and to help establish that pollution, rather than other possible causes, is the likely cause of depressed blood hemoglobin. (i) Firstly, we show that Hgb effects are only observed near mines where the combination of minerals mined suggests that lead contamination is likely to be present. (ii) Secondly, we provide direct evidence of reduced ability to recover Hgb after blood loss — an effect that is hard to reconcile with causes other than lead toxicity.

We observe effects on hemoglobin levels only near mines where we expect heavy metal pollution

Table 11 shows that the effect on Hgb levels of living in mining communities are statistically zero (and mildly negative) in women living near mines where there is less reason to expect heavy metal contamination. However, in mines where there is a high likelihood of such contamination, Hgb levels are strongly and significantly depressed – by about 0.22 g/dL relative to women living farther away from the same mines, and by 0.19 g/dL compared to women living near non-heavy metal mines. (Column 3) Correspondingly, the incidence of anemia is five percentage points higher compared to women living near non-heavy metal mines (compared to women living further away from the same mines, it is six percentage points higher). (Column 4) The size of the cross-sectional effect near heavy metal mines is far closer to the panel effect than the average effect in the cross-section.[19] As noted (in Section 4.2.1), our definition of heavy metal mines is best thought of as a

[19] A similar test is hard to construct for the panel, since mines that are potentially associated with heavy metal contamination make up a large part of the sample.

meaningful but far from perfect proxy of the presence of lead and other toxic metals. In consequence, DiD estimates are likely attenuated.

The DiD effect is robust to including interactions of the treatment dummy with region indicators (hence allaying any concerns over geographical clustering of heavy metal mines), as well as to including an interaction of the treatment with a pregnancy dummy. (Columns 5-6) We note that there is a significant negative effect of living near *any* mine in Latin America (the base category for the region interaction), perhaps due to the imperfect nature of our definition of heavy metal mines. The effect near any mine is statistically zero for the other regions.[20] We further estimate the DiD model for the asset index, and confirm that there is no differential wealth impact of living close to a heavy metal mine, as opposed to any mine. (Column 7) Finally, we do not observe similar differential effects of living near a mine associated with heavy metal contamination on two general indicators of ill health among women, namely miscarriage, and grave sickness (Columns 8-9).

The trajectory of maternal Hgb recovery after birth in mining communities corresponds with known pathophysiological patterns

The left panel in Figure 6 shows the pattern of recovery from blood loss during pregnancy and delivery among women living close to heavy metal mines, and those living in adjacent areas. Hgb levels conspicuously diverge during pregnancy, and stay apart during the first one and one-half years of the child's life. However, thereafter, they converge to an apparent noise pattern about a common mean. (The right panel shows the same data, with effects smoothed out for the nine months from conception to birth, and each year of the newborn's life, thereafter.) The pattern is characteristic of a pollution-induced decrease in the ability to recover Hgb after blood loss, as described in Grandjean et al. (1989) and discussed above (in Section 3.2.2), but not of other causes of anemia.

While the pattern is visually striking, given limited sample size, it is too strong a test to assess the difference between coefficients for the two distance groups in each individual trimester. Instead, we test for the difference in differences between the groups across two time periods: pregnancy and the first year of the infant's life (when there is the clear impression of divergence), and the second and third years of the child's life (when there is not). The results presented in Table 12 show that the DiD coefficient is negative, large (0.25 g/dL), and significant. (Column 1) That is, the difference in Hgb levels between women exposed to mining and other women is far greater during and after blood loss due to pregnancy and delivery, than after some time has passed since

[20] As a further robustness check, Appendix Figure M.3 demonstrates that the median difference between heavy metal mines and non-heavy metal mines is always at least weakly negative in each individual country for which sufficient mine-level estimates can be computed.

delivery. The single difference in distance is negative, but not stable on the small sub-sample of women in the model. As expected, Hgb is dramatically lower in all women during pregnancy and in the first year post-partum.

The pattern is similar when we estimate the model with mine-level fixed effects, as shown in Column (2). Mine-level results do not always reach significance, but are as stable as the state-level results when we include controls, vary the treatment definition, or conduct placebo tests. Because of the small sample size and strong identification from the DiD setup, we prefer the state-level model. In our baseline model, we consider a postpartum period of three years. This seems more appropriate than shorter periods because the detailed time pattern of Hgb recovery shown in Figure 6 suggests that differences even out only in the second year of the child's life. It seems more appropriate than longer periods because the more we extend the time window, the stronger are the identifying assumptions required. Results are robust to extending the post-partum control period to four or five years; they are directionally consistent but insignificant when we shorten it to just two years. (Results not shown.)

Alternative explanations for the pattern of Hgb recovery are harder to come by than those for cross-sectional differences in Hgb levels. Because the test uses as a counterfactual women whose most recent birth lies at most three years in the past, identification requires only that the precise timing of pregnancies is ignorable within a limited time window. However, somewhat complex behavior patterns could generate the observed effect. Perhaps most simply, wealth could be associated with different child bearing choices in mining communities and control locations. For instance, it might be that wealthier women (with higher baseline Hgb levels) tend to have fewer children or space out births more in mining communities than in communities farther afield – perhaps because of better earnings opportunities. The DiD effect could then be due to comparing (relatively) poorer women in mining towns to richer controls in the pregnancy and post-partum group, and (relatively) wealthier women in mining towns to poorer controls for the following years.

To conclusively assess this concern, we first (i) note that Column (7) shows that there are no significant DiD effects on wealth. Secondly, (ii) the DiD effect is robust to controlling directly for the woman's height as a slow-moving wealth proxy, or for whether she gave birth in an 'improved' setting. (Columns 3 and 4). Finally, we (iii) show a placebo regression to test whether a similar recovery pattern emerges when we compare mothers in households in the bottom wealth quintile (placebo treatment) to those in the top quintile (placebo control). We generate two samples: a small sample designed to match the baseline sample particularly tightly, and a larger sample designed to allow for more power. Both placebo samples include women who are pregnant or have given birth within the past three years, and reside at least 20km away from the nearest mine. The small sample is restricted to observations in the same state-year pairs as those observed in the main model, and the large sample, to observations within the same survey rounds. As expected, Columns (5) and (6) show that women in poor households always have lower Hgb levels than those in wealthy households – but there is no indication of an adverse time pattern around pregnancy and postpartum, with placebo DiD coefficients either near zero, or with an opposite sign.

In summary, we obtain two DiD tests by disaggregating effects, first among mine types, and then with respect to recent pregnancy. The results are instructive both regarding mechanisms of treatment transmission and regarding identification. In terms of mechanisms, they offer strong evidence that the observed health effect is caused by pollution, not other facets of life near mines. For instance, if the observed effect on Hgb were due to iron deficiency or malaria infection, then nutritional behavior and infection rates would have to vary across distance groups in systematically different ways near metal and non-metal mines, and among pregnant and non-pregnant women – despite the fact that socio-economic outcomes do not vary in such ways. The results also provide reassurance on identification, most importantly because they are very hard to explain with sorting. Because mine types differ in Hgb impacts, but not in wealth and non-specific health impacts, one would have to hypothesize that in their migration decisions, people not only take mine type into account, but also differentially sort on their potential health and wealth outcomes. (We have discussed above the corollary for Hgb recovery.) This would require an extraordinary level of sophistication.

Residents of mining communities are not differentially affected by causes of anemia other than lead exposure, do not bear a higher burden of disease unrelated to pollution, and are not under-served by health care

The high dimensionality of the DHS data allows for diverse falsification tests that could yield evidence against our contention that the observed hematologic effects are due to pollution, not other mechanisms. Across a range of tests, we find no such evidence.

Firstly, we show in Appendix F that here is no conclusive pattern in mining communities in the leading causes of anemia other than lead toxicity (nutritional iron deficiency, malaria, and intestinal worm infections). Secondly, we test whether residents of mining communities suffer ill health that is unlikely to be attributable to pollution. Significance patterns are very sparse in the cross-section, and there are no significant adverse health impacts at all in the panel (Table 13).[21] Appendix Tables

[21] It is noteworthy that, among never-movers, infant and child mortality rates decrease significantly and strongly, by 0.7 and one percentage points, respectively. This result is of course consistent with greater wealth in mining communities. Panel results do not show a significant decrease in mortality, although the sign on the DiD coefficient is negative.

J and K show additional specifications with similarly sparse patterns.[22]

Finally, we note that residents of mining communities are at least as well off in terms of health care as those living farther afield. As Appendix Tables L.1 and L.2 demonstrate, in the long run, women are more likely to have health insurance coverage, and to give birth with some level of skilled assistance. Panel results suggest that such benefits, along with access to health care, may extend beyond the immediate vicinity of the mine. The one potential exception to this pattern is that in our mother-level panel, we find significant decreases in the share of women who gave birth in an improved setting in mining communities when the mine was operational. The cross-sectional and mine-level panel evidence contradicts this finding. However, we mention it here because it is at odds with our otherwise consistent evidence on wealth. We note that our discussion of maternal Hgb recovery explicitly sought to exclude the potential effect of differences in maternal health care.

Patterns of anemia among children mirror those among women, but are less conclusive

Our data shows patterns of anemia among children in mining communities that resemble those found among adult women. However, significant results are hard to come by. This may be because the true treatment effect is weaker – we have noted above (Section 3.2.2) that children can effectively compensate for the hematologic toxicity of lead by increasing production of EPO and red blood cell production. It may also be due to small sample size (for children, we only have about half the number of observations in the women's sample). In the cross-section, we observe insignificant decreases in Hgb on the order of 0.07 g/dL (Table 14, Column 1); the effect is strongly concentrated near heavy metal mines, but the DiD coefficient is again not significant. (Column 2) The panel shows statistically insignificant losses from current exposure to mining, but is highly sensitive to changes in the treatment definition. (Results omitted.)

Next, we explore whether infants might be more strongly affected by pollution than older children. There are two reasons to expect this. Firstly, we have attributed anemia among women – and particularly, pregnant women – in mining communities to lead exposure, and it is known that children are born with a lead burden mirroring that of their mothers. Secondly, it has been previously shown that compensatory over-production of EPO and Hgb in lead-exposed children does not quite start at birth, but at some point during infancy. (Wasserman et al., 1992) When we consider impacts on infants only, we find a larger but insignificant effect on Hgb (0.13 g/dL), relative to infants living farther away from the mine. (Column 3) However, the differential impact on infants near heavy metal mines (Column 4) is both significant and large. The triple-difference

[22] We also show in Appendix Table L.4 that we find no indication of greater alcohol abuse among men or women, and at most a mild indication of increased sexual risk taking, consistent with Wilson (2012).

coefficient shows a 0.60 g/dL difference in Hgb levels, with a nearly identical and significant difference in differences between the effect on infants near heavy metal mines and infants near other mines. Falsification results show that infants near these mines are not indiscriminately less healthy. (Columns 6-8) However, we caution that infants born in the direct vicinity of heavy metal mines tend to live in poorer households. (Column 5) As shown above, we did not find such a correlation between mine type and wealth in our analysis of hematologic toxic effects among women living near heavy metal mines. The fact that we do find it here makes it somewhat less compelling to interpret the difference among mine types as evidence that the health impacts are due to pollution.

6.3 Evidence of adverse growth outcomes

As noted, exposure to environmental lead has previously been linked to decreased growth early in life. However, the evidence is mixed. In the following, we consider impacts on height for age and the incidence of stunting and severe stunting (height more than two or three standard deviations below the age-appropriate median, respectively). We find evidence of lower height among children exposed to a mining environment *in utero*, but also evidence of a compensatory positive growth effect of living in mining communities after birth. Appendix Table I reports that we observe an effect on birth weight in the mother-level panel only, and lack corroborating evidence from our other models.[23]

Without regard to time patterns of exposure, children in mining communities grow taller than their peers

In the simple cross-section, we observe *better* outcomes for height among children of less than five years of age in mining communities than in the controls. (Table 15, Column 1) This may not be surprising: growth is strongly linked with nutrition (both the mother's and the child's), and with greater wealth in mining communities, there may also be better diets. There is also no differential impact near 'heavy metal' mines. (Column 3)

However, the evidence is somewhat more subtle. Firstly, as Column 2 makes obvious, there is no indication of a positive effect among never-movers. Secondly, although infants are not more affected than older children when we consider *all* types of mines (Column 4), there is at least some indication of an adverse effect on infants of living near a 'heavy metal' mine. The triple-difference effects are adverse, and significant for stunting. The DiD comparing the treatment

[23] Prior studies have observed that adverse conditions *in utero* can impair long-run well-being without being reflected in birth weight. (Schulz, 2010)

effect of closeness on infants near metal mines and other mines amounts to a loss of 0.1 standard deviations in the height measure, and a four and two percentage point increase in the incidence of stunting and severe stunting, respectively, although none of these effects reach significance. (Columns 5-7) There are no significant differences between mine types in the economic status of families with infants. (Column 8)

The cross-sectional evidence alone is thus not easy to read. There clearly are growth benefits to be had for children in mining communities, and it seems obvious to connect these to the wealth increases enjoyed by residents. However, not all children appear to benefit. The question is whether this is because some children are simply left out from economic gains, or whether they suffer countervailing direct health damage. The absence of a DiD effect between mine types and of a differential effect on infants near all mines may suggest the former. Yet, the appreciable effect on infants near mines associated with heavy metals points toward the latter. Similarly, the difference between never-movers and the general population is consistent with lower economic benefits among never-movers. (See Table 6.) However, since the differential in wealth effects is not very large, it is reasonable to note that children born to never-movers are also more likely to have been exposed to pollution, particularly *in utero*, through the maternal body burden of lead. We look to the panel for more conclusive evidence of the impacts of different exposure patterns.

Panel evidence shows that *in utero* exposure to mining increases the incidence of stunting

Results from the mine-level panel confirm that there is an effect of mining activity on height, that the effect is chiefly due to exposure *in utero*, and that it attenuates with age. It also allows us to at least suggest that there are genuinely positive effects of life in mining communities on growth in older children, so that children do not simply 'out-grow' *in utero* effects without further exposure, as earlier reported by Shukla et al. (1991).

The DiD effect of *in utero* exposure among all children under five years of age shows a loss of 0.14 standard deviations in the height index, and a five percentage point increase in the incidence of stunting and severe stunting. (Table 16, Columns 1-3) The effect on the discrete outcomes is significant; the one on the continuous measure not significant (t = 1.39), but stable. With a baseline incidence of 23% and 8%, respectively, the impact on stunting is appreciable, and the impact on severe stunting dramatic.

We next note that, in the case of the continuous index and of stunting, the effect of *in utero* exposure is larger and stronger when we estimate it for infants only. (Columns 4-6) This points either to a balancing effect – perhaps due to household wealth – in older children, or a spontaneous attenuation of *in utero* impacts with time. We shed some further light on this question by studying the effect of different exposure patterns. Thus, the estimated effect of exposure during the first year of life alone is centered near zero. (Column 7) Results when estimating *in utero* and birth-year effects jointly are more instructive. We find robust and large adverse effects of *in utero* exposure on the continuous index (0.5σ), alongside beneficial effects of birth-year exposure. (Column 8) This is at least consistent with exposure to maternal lead loads *in utero*, alongside positive effects from the socio-economic benefits of mining, once the child is born. It points less toward a mere attenuation of impacts. While it is attractive to allow *in utero* and birth-year effects to jointly enter into the model, the sample of children born just before and just after a mine opened or closed is small (conversely, operational status is highly serially correlated).[24] To further solidify the result, we therefore show that a similar pattern emerges when we first estimate separately the effect of the mine operating during the survey year (Column 9), and then compare this estimate to the one obtained when we include also the effect of the mine operating during gestation. (Column 10)

Finally, when we estimate the effects of *in utero* and birth-year exposure with mother-level effects, the results match the pattern in the mine-level panel, but do not reach significance. (Columns 11-13) This is perhaps to be expected: although we observe more than 2,000 women near mines in our sample for whom our data records child growth outcomes for at least two children born within five years of each other, there are few mothers with recorded births both while the mine was operational and while it was not operational.

Patterns among children born to migrants provide further evidence of adverse *in utero* effects on growth

Because the panel evidence suggests that it is *in utero* exposure to mining that matters for growth outcomes, we are able to leverage differences in the timing of exposure among children born to migrants for an additional test. More precisely, we compare measures of growth in children born to migrants (i) before and after migrating (in our preferred specification, within no more than four years of the move), and (ii) born to mothers who moved to locations close to a mine and those who moved to the general vicinity. (For details on how exposure groups are defined, refer to Appendix A.)

The DiD estimate suggests that children whose families moved to a mining community before conception – and who hence were exposed to a mining environment *in utero* – experience a decrease in their on height-for-age score of 0.17 standard deviation, controlling for mine-year fixed effects. (Table 17, Column 1) This is reasonably close to the mine panel estimate of the *in utero*

[24]The DHS surveys record only health data from children born no more than five years before the survey time. This helps identification, but limits sample size, in particular where we use mother-level effects.

effect. When controlling for mother fixed effects, we find an insignificant adverse effect of similar size. (Column 2) Appendix H shows that the results are reasonably stable when we change the window around the time of migration; for completeness, we note that they are not robust to including observations with ambiguous treatment status omitted from our baseline sample.

Columns (3-6) report falsification tests. First, as a meaningful measure of well-being that varies at the birth year, we show in Columns (3-4) that mothers were at least weakly more likely to have given birth in an improved setting if their children were born after they moved to a mining community.[25] Children born to migrants in mining communities also suffered at least weakly lower infant mortality rates. (Columns 5-6)

7 Conclusion

We present the first systematic empirical assessment of the health-wealth trade-off facing mining communities, using micro-data from 44 developing countries. In communities in the vicinity of mines, we find important economic benefits, alongside serious health impacts, namely increases in the incidence of anemia in adult women, and of stunting in young children. These health impacts are consistent with exposure to lead contamination, and have previously been observed at body burdens of lead that are known also to cause cognitive deficits in children.

We obtain estimates of long-run effects from a cross-sectional fixed effects model; medium-term estimates come from mine-level and mother-level panels. We confirm our wealth results with an IV approach that uses deposit location and world mineral prices to instrument for mine locations and operating times. We then develop additional difference-in-difference tests that exploit (i) the association of certain mine types with lead pollution, (ii) known pathological patterns of Hgb recovery in adults exposed to lead, and (iii) variation in *in utero* exposure among children born to migrants induced by the exact timing of conception relative to the date of the move. These additional tests are intended both to allow for weaker identifying assumptions, and to demonstrate that the observed health impacts are due to pollution, rather than other mechanisms.

The economic benefits to mining communities in the long run are on the order of 0.1 standard deviations of a country and year-specific asset index. Medium-term benefits to households in communities near operating mines are larger, on the order of 0.3σ. Benefits are strongly concentrated

[25]Children in the treatment group also live in families with weakly higher asset index values. However, the asset index is measured only once for every mother – at survey time – while treatment status varies with the birth years of her children.

within the immediate vicinity (5km) of mines, and we find no asset wealth effects at all beyond some 15-20km. Wealth rises quite evenly across the distribution, with modest increases in inequality in the long run. Benefits in terms of health care may extend beyond the most direct vicinity of mines, although mining communities do at least as well as communities farther afield.

The evidence conclusively reveals that the real economic benefits generated in mining communities go hand in hand with increases in the incidence of anemia, by three to ten percentage points in adult women. The ability to recover hemoglobin levels after blood loss due to pregnancy and delivery is particularly impaired. There is weaker but consistent evidence of hematologic toxic effects in children. Children in mining communities are not disadvantaged in all aspects of physical growth. Yet, young children exposed to a mining environment *in utero* are more likely to be stunted or severely stunted than those born in control groups, with an increase in incidence of five percentage points. There is very limited evidence of reduced birth weight, and increases in stunting are clearly strongest among infants, and may not persist. By way of contrast to these specific health impacts, there is no general pattern of ill health in mining communities.

We conclude by highlighting some conceptual and policy implications of our results.

Firstly, the presence of adverse compensated health impacts in a generally wealthier population poses an important question. The most straightforward explanation might be to suggest that the cost of avoiding exposure to pollution is high, perhaps due to the structure of settlements and the quality of public transport. We can speculate whether the decision on living in mining towns in developing countries might resemble less the choice of an optimal distance along a continuum, and more a discrete choice between two stark options – namely living either in relatively unpolluted communities outside of a reasonable commuting distance to the mine, or in a highly polluted but bustling community adjacent to the mine. The high spatial concentration of medium-term economic benefits is certainly consistent with such a situation, as is the fact that we observe the greatest wealth effects near mines in environments that are economically less active. An alternative explanation might point to limited information. Pollutant levels near mines vary greatly, even over small distances (van Geen et al., 2012). Hence, contamination may not always be easily observed. In addition, the health impacts of pollution may not be widely known. The fact that we find strongly raised wealth levels, but only weakly better health care among households in the direct vicinity of mines at least suggests that residents are not making very decisive health investments to compensate for exposure to pollution.

Secondly, while a direct comparison of our estimates of health cost and wealth benefits requires strong assumptions and involves considerable uncertainty, we can offer some observations with reasonable confidence. (Appendix E provides background on the following discussion.)

We first note that the cost to affected individuals of the health consequences we observe is very high. The contemporaneous productivity loss due to anemia in adults has been estimated to be on the order of 5-17%. The persistent economic impact of stunting is known to be dramatic (if childhood stunting persists through adulthood), perhaps as large as an annual 53% loss in adult wages. (For completeness, we note that the imputed permanent annual productivity loss due to cognitive effects may be on the order of 1.6-13%.)[26] – Yet, it is also clear that the health burden imposed by mining pollution is very unequally distributed: at least in our compensated reduced-form estimates, relatively small population groups are affected. In consequence, the *expected* cost of health impacts is modest, while the individual cost on those afflicted with health problems is steep.

For the sake of comparing wealth effects to these productivity losses, we suggest in Appendix E that wealth benefits may be roughly of an order associated with a 7-37% increment in household income. Without relying on the exact magnitude of either of these estimates, we can thus conclude the following. (i) If they reflect changes in *permanent* income, the wealth benefits are of similar magnitude as the health cost borne by *affected* residents. Both considerably affect well-being. At the same time, (ii) if income changes are fully persistent, they outweigh *expected* cost by an order of magnitude. However, (iii) the cost-benefit balance tilts dramatically toward costs if economic gains are less than permanent. Certainly, the cost to affected individuals begins to outweigh gains as soon as the latter are not fully permanent. Similarly, (iv) the balance tilts toward costs if legacy effects of pollution after operations cease outlast economic benefits.

These observations imply that, if economic benefits are sufficiently persistent, we may view the decision to live in mining communities as a rational but risky choice to locate close to economic opportunity. (We can draw the same conclusion if residents believe them to be permanent, or are sufficiently impatient.) We also note that, while we have shown that economic gains are quite equally distributed, the *net* benefits of mining look to be very unequally distributed. Thus, mining makes winners and losers not only between communities that benefit and communities that suffer consequences, but also *within* mining communities.

From a policy perspective, our evidence suggests that – on the global average – residents of mining communities can expect to benefit from the industry. (This is of course not to say that there are not instances of egregious local environmental damage and gross welfare loss.) Still, the presence of a health externality due to normal operations at mines in our sample that is observable in compensated health outcomes suggests that the management of mining pollution deserves

[26]While we do not find strong evidence of an effect of mining on the prevalence of other health conditions recorded in our data, mining communities may obviously suffer health impacts – or enjoy health benefits – that we do not observe.

renewed scrutiny. Our results yield two leads as to what effective interventions might look like. One, health concerns are most acute in the immediate vicinity of mines. Proven but expensive engineering solutions to contain and remediate pollution therefore might deserve a second look. Similarly, policy approaches need not be too broad in spatial scope to allow residents to live away from the worst pollution, while still working in or near the mine. At least for some countries in our sample, there may be a case for experimentation with programs to improve public transport, road infrastructure, or flexibility in local housing markets. Secondly, the highly uneven distribution of damages may imply that there is a premium on interventions that reduce risk. We note that the uneven distribution of costs mirrors the great spatial variation in pollution around mines described in van Geen et al. (2012), and it is tempting to posit that it might be causally related. If so, then *testing* of pollution levels in residential areas might enable residents to avoid the most dangerous sites, at a comparatively low cost.

References

Acemoglu, D., Finkelstein, A., and Notowidigdo, M. J. (2013). Income and health spending: evidence from oil price shocks. *Review of Economics and Statistics*, 95(4):1079–1095.

Alloway, B. J. (2013). Sources of heavy metals and metalloids in soils. In *Heavy Metals in Soils*, pages 11–50. Springer.

Aragón, F. M. and Rud, J. P. (2013a). Modern industries, pollution and agricultural productivity: Evidence from Ghana. Working paper.

Aragón, F. M. and Rud, J. P. (2013b). Natural resources and local communities: evidence from a Peruvian gold mine. *American Economic Journal: Economic Policy*, 5(2):1–25.

Arceo-Gomez, E. O., Hanna, R., and Oliva, P. (2012). Does the effect of pollution on infant mortality differ between developing and developed countries? Evidence from Mexico City. Working Paper 18349, National Bureau of Economic Research.

ATSDR (2007). Toxicological profile for lead. Technical report, US Department of Health and Human Services. Public Health Service, Atlanta, GA, USA.

Baghurst, P. A., McMichael, A. J., Wigg, N. R., Vimpani, G. V., Robertson, E. F., Roberts, R. J., and Tong, S.-L. (1992). Environmental exposure to lead and children's intelligence at the age of seven years: the Port Pirie Cohort Study. *New England Journal of Medicine*, 327(18):1279–1284.

Bellinger, D., Leviton, A., Rabinowitz, M., Allred, E., Needleman, H., and Schoenbaum, S. (1991). Weight gain and maturity in fetuses exposed to low levels of lead. *Environmental Research*, 54(2):151–158.

Bellinger, D. C. (2004). Lead. *Pediatrics*, 113(4 Suppl):1016–1022.

CDC (2012). Low level lead exposure harms children: A renewed call of primary prevention. *Atlanta, GA: Centers for Disease Control and Prevention*.

Chen, Y., Ebenstein, A., Greenstone, M., and Li, H. (2013). Evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River policy. *Proceedings of the National Academy of Sciences*, 110(32):12936–12941.

Corno, L. and De Walque, D. (2012). Mines, migration and HIV/AIDS in southern africa. *Journal of African Economies*, 21(3):465–498.

Currie, J., Davis, L., Greenstone, M., and Walker, R. (2013). Do housing prices reflect environmental health risks? Evidence from more than 1600 toxic plant openings and closings. Working Paper 18700, National Bureau of Economic Research.

Dell, M. (2010). The persistent effects of Peru's mining *mita*. *Econometrica*, 78(6):1863–1903.

Dube, O. and Vargas, J. (2013). Commodity price shocks and civil conflict: Evidence from Colombia. *The Review of Economic Studies*.

Duflo, E. and Pande, R. (2007). Dams. *The Quarterly Journal of Economics*, 122(2):601–646.

Ebenstein, A. (2012). The consequences of industrialization: Evidence from water pollution and digestive cancers in China. *Review of Economics and Statistics*, 94(1):186–201.

Factor-Litvak, P., Slavkovich, V., Liu, X., Popovac, D., Preteni, E., Capuni-Paracka, S., Hadzialjevic, S., Lekic, V., LoIacono, N., Kline, J., et al. (1998). Hyperproduction of erythropoietin in nonanemic lead-exposed children. *Environmental Health Perspectives*, 106(6):361.

Factor-Litvak, P., Wasserman, G., Kline, J. K., and Graziano, J. (1999). The Yugoslavia Prospective Study of environmental lead exposure. *Environmental Health Perspectives*, 107(1):9.

Filmer, D. and Pritchett, L. H. (2001). Estimating wealth effects without expenditure dataor tears: An application to educational enrollments in states of India. *Demography*, 38(1):115–132.

Frankel, J. A. (2010). The natural resource curse: a survey. Working Paper 15836, National Bureau of Economic Research.

Gallego, F., Montero, J.-P., and Salas, C. (2013). The effect of transport policies on car use: Evidence from Latin American cities. *Journal of Public Economics*, 107(0):4762.

Gennaioli, N., La Porta, R., Lopez-de Silanes, F., and Shleifer, A. (2013). Human capital and regional development. *The Quarterly Journal of Economics*, 128(1):105–164.

Grandjean, P., Jensen, B. M., Sandø, S., Jøgensen, P., and Antonsen, S. (1989). Delayed blood regeneration in lead exposure: an effect on reserve capacity. *American Journal of Public Health*, 79(10):1385–1388.

Greenstone, M. and Hanna, R. (2011). Environmental regulations, air and water pollution, and infant mortality in India. Working Paper 17210, National Bureau of Economic Research.

Greenstone, M. and Jack, B. K. (2013). Envirodevonomics: A research agenda for a young field. Working Paper 19426, National Bureau of Economic Research.

Hanna, R. and Oliva, P. (2011). The effect of pollution on labor supply: Evidence from a natural experiment in Mexico City. Working Paper 17302, National Bureau of Economic Research.

Hartwick, J. M., Olewiler, N. D., et al. (1986). *The Economics of Natural Resource Use*. Harper & Row.

Hendryx, M. and Ahern, M. M. (2008). Relations between health indicators and residential proximity to coal mining in West Virginia. *American Journal of Public Health*, 98(4):669–671.

Hernandez-Avila, M., Peterson, K. E., Gonzalez-Cossio, T., Sanin, L. H., Aro, A., Schnaas, L., and Hu, H. (2002). Effect of maternal bone lead on length and head circumference of newborns and 1-month-old infants. *Archives of Environmental Health: An International Journal*, 57(5):482–488.

Imbens, G. W. and Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1):5–86.

Infomine (2013). Company and property mining intelligence.

IntierraRMG (2013). Raw materials data.

Kotsadam, A. and Tolonen, A. (2013). Mineral mining and female employment. *Oxcarre Research Papers*, 114.

Landrigan, P. J. and Baker, E. L. (1981). Exposure of children to heavy metals from smelters: epidemiology and toxic consequences. *Environmental Research*, 25(1):204–224.

Lanphear, B. P., Hornung, R., Khoury, J., Yolton, K., Baghurst, P., Bellinger, D. C., Canfield, R. L., Dietrich, K. N., Bornschein, R., Greene, T., et al. (2005). Low-level environmental lead exposure and children's intellectual function: an international pooled analysis. *Environmental Health Perspectives*, 113(7):894.

Lauwerys, R., Buchet, J., Roels, H., and Hubermont, G. (1978). Placental transfer of lead, mercury, cadmium, and carbon monoxide in women: I. comparison of the frequency distributions of the biological indices in maternal and umbilical cord blood. *Environmental research*, 15(2):278–289.

Meyer, B. D. (1995). Natural and quasi-experiments in economics. *Journal of Business & Economic Statistics*, 13(2):151–161.

Mining Atlas (2014). Mining industry data and maps.

Monteiro, J. and Ferraz, C. (2009). Does oil make leaders unaccountable? Mimeo, Pontificia Universidade Catolica do Rio de Janeiro.

Rau, T., Reyes, L., and Urzúa, S. S. (2013). The long-term effects of early lead exposure: Evidence from a case of environmental negligence. Working Paper 18915, National Bureau of Economic Research.

Ripley, E. A., Redmann, R. E., Crowder, A. A., et al. (1996). *Environmental Effects of Mining*. St. Lucie Press Delray Beach, FL.

Roels, H., Bucket, J.-P., Lauwerys, R., Hubermont, G., Bruaux, P., Claeys-ThOreau, F., Lafontaine, A., and Overschelde, J. V. (1976). Impact of air pollution by lead on the heme biosynthetic pathway in school-age children. *Archives of Environmental Health: An International Journal*, 31(6):310–316.

Roper, W. L., Houk, V., Falk, H., and Binder, S. (1991). Preventing lead poisoning in young children: A statement by the Centers for Disease Control, October 1991. Technical report, Centers for Disease Control, Atlanta, GA (United States).

Salomons, W. (1995). Environmental impact of metals derived from mining activities: processes, predictions, prevention. *Journal of Geochemical Exploration*, 52(1):5–23.

Sanín, L. H., González-Cossío, T., Romieu, I., Peterson, K. E., Ruíz, S., Palazuelos, E., Hernández-Avila, M., and Hu, H. (2001). Effect of maternal lead burden on infant weight and weight gain at one month of age among breastfed infants. *Pediatrics*, 107(5):1016–1023.

Schulz, L. C. (2010). The Dutch Hunger Winter and the developmental origins of health and disease. *Proceedings of the National Academy of Sciences*, 107(39):16757–16758.

Schwartz, J., Landrigan, P. J., Baker Jr, E. L., Orenstein, W. A., and Von Lindern, I. (1990). Lead-induced anemia: dose-response relationships and evidence for a threshold. *American Journal of Public Health*, 80(2):165–168.

Shukla, R., Dietrich, K. N., Bornschein, R. L., Berger, O., and Hammond, P. B. (1991). Lead exposure and growth in the early preschool child: a follow-up report from the Cincinnati lead study. *Pediatrics*, 88(5):886–892.

Sloan, N. L., Jordan, E., and Winikoff, B. (2002). Effects of iron supplementation on maternal hematologic status in pregnancy. *American Journal of Public Health*, 92(2):288–293.

United States Geological Survey (2005). Mineral Resources Data System.

van Geen, A., Bravo, C., Gil, V., Sherpa, S., and Jack, D. (2012). Lead exposure from soil in peruvian mining towns: a national assessment supported by two contrasting examples. *Bulletin of the World Health Organization*, 90(12):878–886.

Wasserman, G., Graziano, J., Factor-Litvak, P., Popovac, D., Morina, N., Musabegovic, A., Vrenezi, N., Capuni-Paracka, S., Lekic, V., Preteni-Redjepi, E., others, et al. (1992). Independent effects of lead exposure and iron deficiency anemia on developmental outcome at age 2 years. *The Journal of Pediatrics*, 121(5):695–703.

Wilson, N. (2012). Economic booms and risky sexual behavior: Evidence from Zambian copper mining cities. *Journal of Health Economics*.

World Health Organization (2011). Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity.

Wright, D. A. and Welbourn, P. (2002). *Environmental Toxicology*, volume 11. Cambridge University Press.

Zhu, M., Fitzgerald, E. F., Gelberg, K. H., Lin, S., and Druschel, C. M. (2010). Maternal low-level lead exposure and fetal growth. *Environmental Health Perspectives*, 118(10):1471.

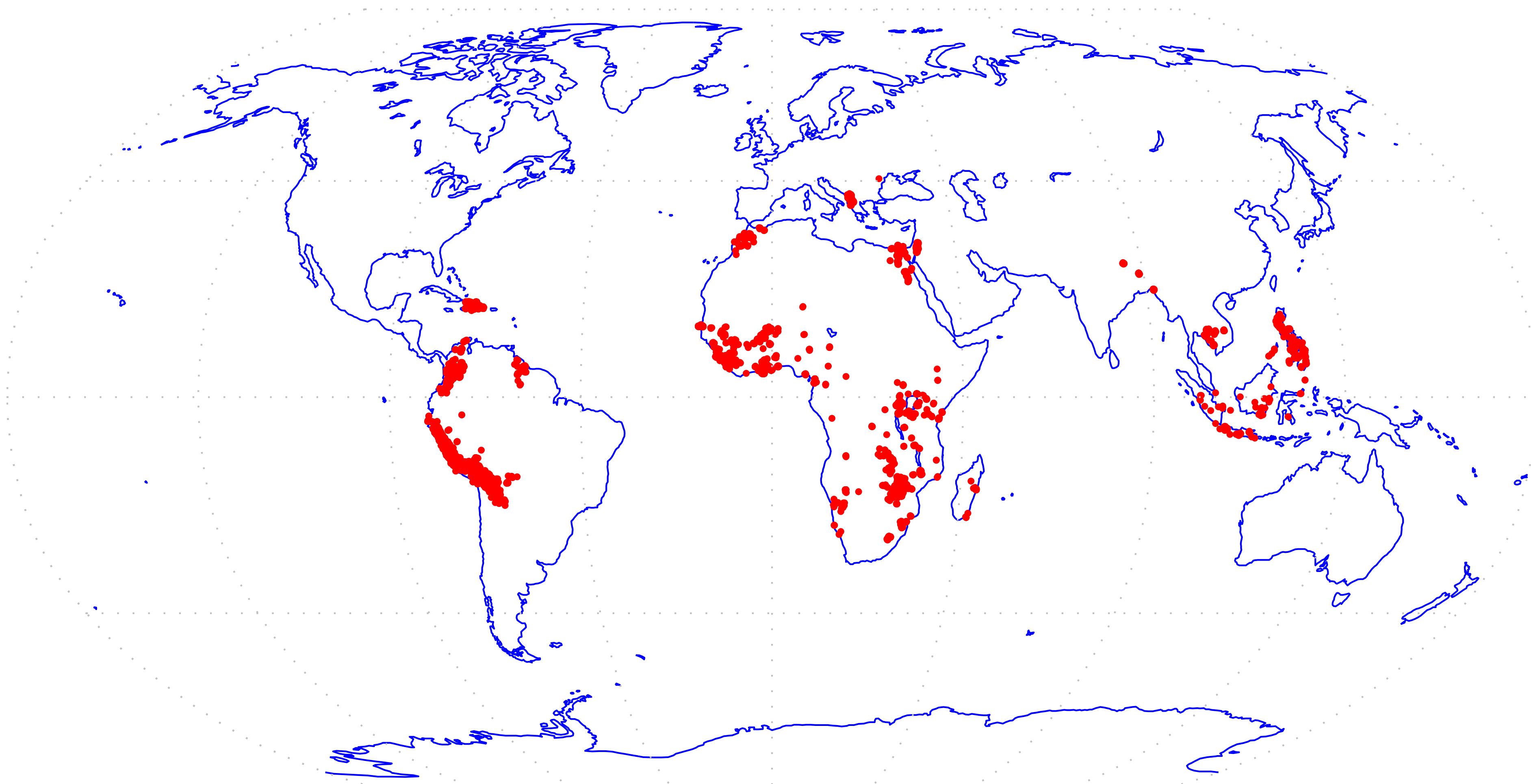


Figure 1: DHS clusters within no more than 20km of a mine in the sample

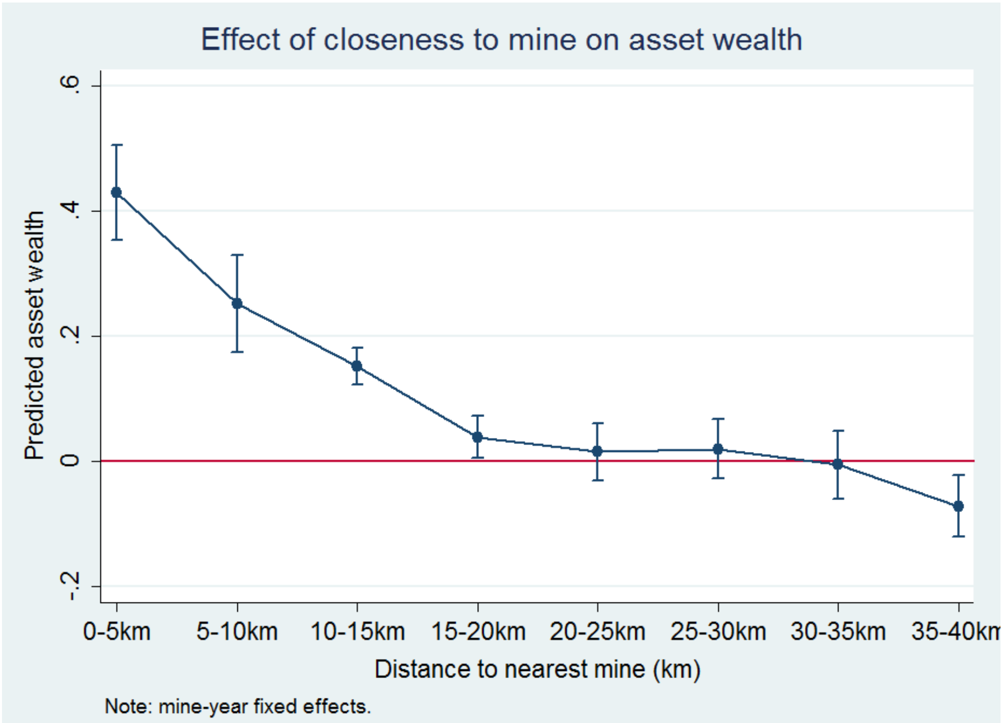


Figure 2: Effect of closeness to mine on asset wealth in the cross-section

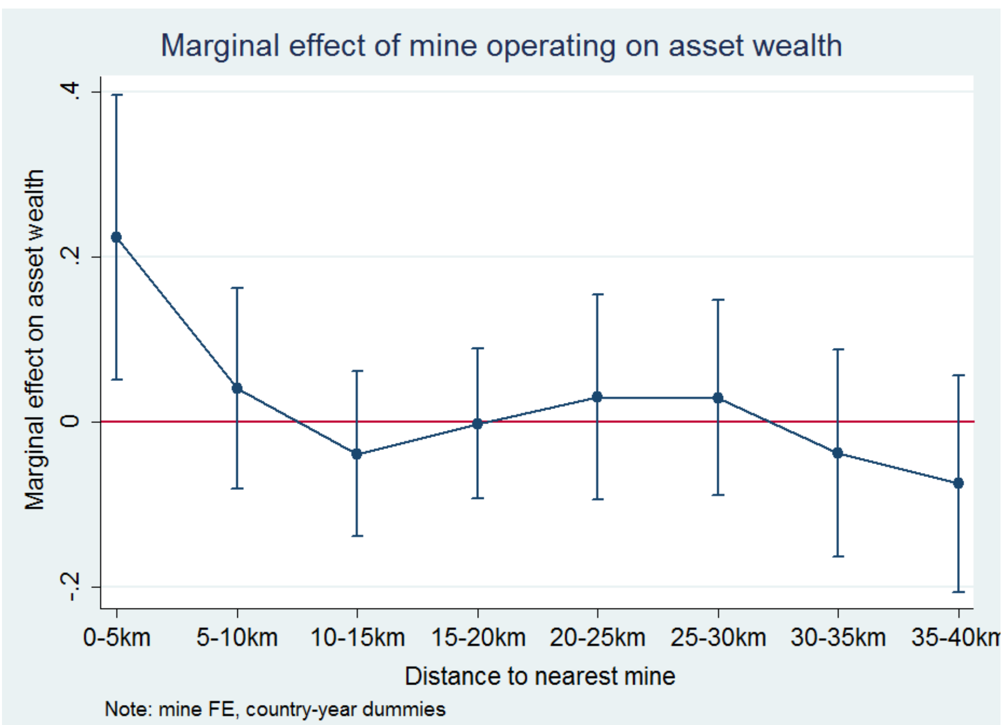


Figure 3: Marginal effect of mine operating on asset wealth in the panel

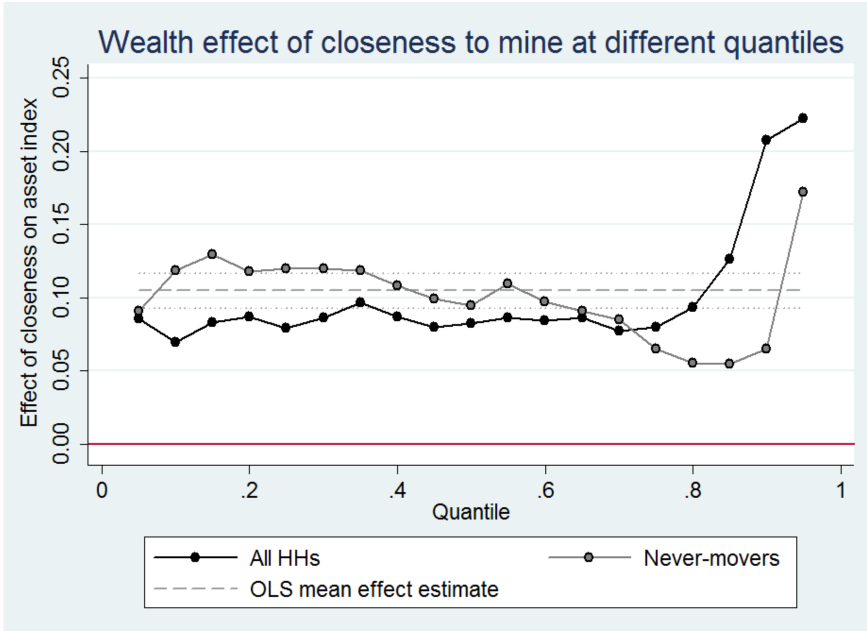


Figure 4: Cross-sectional effect of closeness to mine on asset wealth at different quantiles of the wealth distribution

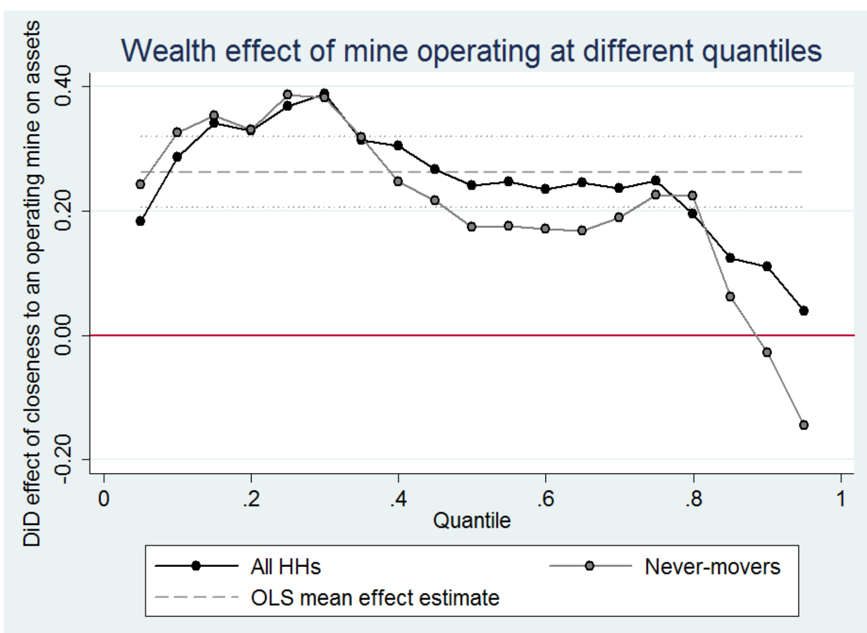


Figure 5: Panel effect of mine operating on asset wealth at different quantiles of the wealth distribution

Figure 6: Effect of closeness to heavy metal mine on maternal Hgb recovery

Table 1
Overview of estimators

Estimator	Uses variation in distance to mine and…	DHS respondent selection criteria	Fixed effects (baseline specification)	Time effects (baseline specification)
Baseline estimators				
Cross-section	n/a	≤ 20km from mine	Mine-year	n/a
Mine-level panel	Operating status	≤ 20km from mine	Mine	Country-year
Mother-level panel	Operating status	≤ 20km from mine	Mother	Country-specific linear trends
IV estimators				
Cross-sectional IV	n/a	≤ 20km from deposit	State-year	n/a
Mine-level panel IV	Operating status	≤ 20km from deposit	Mine	Country-year
Tailored DiD estimators				
Mine-type DiD	Type of nearest mine	≤ 20km from mine	Mine-year	n/a
Maternal Hgb recovery	Recency of pregnancy	Women ≤ 20km from heavy metal mines with a birth within last three years	State-year	n/a
Migrant DiD	Timing of conception relative to the time of migration	Children ≤ 20km from heavy metal mines, born to migrants within the last four years	Mine-year	n/a

Note. Country-year time effects are implemented as indicator variables for each survey-round. This is because in many surveys, a small number of observations was taken in the months just before or after the main survey year.

Table 2

Mine types, associated pollutants, and health effects

Mine type	Pollutants of concern	Health effects
Polymetallic mines	Heavy metals, especially lead	Neurodevelopmental damage, anemia, growth deficits, renal disease
Small-scale gold and silver mining	Mercury	Renal disease, neurological conditions
Large-scale gold mining	Cyanide	Heart irregularities, thyroid problems
Bulk metal mines, gemstone mines	Particulates	Respiratory disease, GI problems associated with turbid water
Phosphate rock	Radionuclides	Lung cancer and non-malignant respiratory disease
Coal	Particulates, radionuclides	Respiratory disease, GI problems, lung cancer
Metal smelters	*Heavy metals,* SO_2	*As shown for polymetallic mines, and respiratory disease*

Notes. Not all mine types are mutually exclusive. Mapping based on ATSDR Toxicological Profiles for the respective pollutants, Alloway (2013), Ripley (1996), and Wright and Welbourn (2002). Health effects as reported from chronic low-level environmental exposure.

Table 3

Prior literature on blood lead levels in communities near smelters

	Distance to smelter	Mean PbB
Fontúrbel et al., 2011	0.5-1.8km	n/a
Roels et al., 1980	1-2.5km	13-30 μg/dL
Recio-Vega et al., 2012	2km	14-19 μg/dL
Factor-Litvak et al., 1999	2-4km	28-39 μg/dL
Benin et al., 1999	3km	20-40 μg/dL
Landrigan and Baker, 1981	4km	*≥ 40 μg/dL in 87% of subjects*

Notes. The table summarizes prior studies of lead levels in communities near smelters. It shows the maximum distance between the smelter and the communities considered highly exposed, alongside mean blood lead in highly exposed communities. Ranges of mean PbB refer to means for population groups that differ in age, gender, and other characteristics. Incidence for Landrigan and Baker summarized by the authors. In the case of Benin et al. (1999), PbB was predicted from observed environmental pollution; in all other studies, PbB was measured directly.

Table 4
Sample size

	Surveys with observations within 20km of a mine
Survey rounds	104
Countries	44
Interview years	25

	Number of households		
	Full sample	Within 5km of a mine	Within 5-20km of a mine
Households	1,192,492	37,608	132,797
% of total		*3.2%*	*11.1%*
Children under five years of age	1,364,156	31,964	121,519
Women aged 15 and over	2,877,024	87,234	310,096
Men aged 15 and over	2,717,928	82,973	294,723

	Mines and smelters near DHS sampling clusters		
	USGS data	RMD data	Infomine data
DHS cluster within 20km	838	508	7
DHS cluster within 0-5km	339	225	4
DHS cluster within 5-20km	687	455	6
DHS cluster in both distance categories	226	172	3

Notes. Sample size based on all types of mines, smelters, and legacies, excluding quarries. Not all variables used in this study are available for the entire sample. The count of locations from Infomine includes only those mines not covered in the RMD data.

Table 5
Cross-sectional effect on mean asset wealth and wealth disparities

	Mean asset wealth		Mean absolute wealth deviation	
	All HHs	Never-movers	All HHs	Never-movers
	(1)	(2)	(3)	(4)
HH close to mine	**0.105***** **(0.035)**	**0.0784*** **(0.0423)**	**0.0274*** **(0.0156)**	**0.00746** **(0.0151)**
Number of households	90,319	31,079	90,319	31,079
Number of groups	1,562	1,371	1,562	1,371
R-squared	0.094	0.081	0.010	0.004

Notes. The table reports estimates of equation (1), using indicator variables for each mine-year pair as fixed effects. The dependent variable in Columns (1) and (2) is the asset factor index, with units expressed in standard deviations. In Columns (3) and (4), it is the absolute deviation of a household's asset factor index from the sampling cluster mean, in units of standard deviations. Controls include a quadratic in the household head's age and an indicator for urban/rural status. Columns (2) and (4) restrict the sample to households with at least one respondent who had always been resident in the current location at survey time. Standard errors are clustered at the mine level. Significant at * 10%, ** 5%, *** 1%.

Table 6
Effects on mean asset wealth and wealth disparities in the panel

	All households			
	Mean asset wealth			Mean absolute wealth deviation
	(1)	(2)	(3)	(4)
HH close to mine	-0.0555	-0.113	-0.0874	-0.0451
	(0.0878)	(0.089)	(0.0954)	(0.0362)
Mine operating	0.0651	-0.0296	0.0639	-0.00611
	(0.0566)	(0.0348)	(0.133)	(0.0300)
Mine operating * HH close (DiD)	**0.159***	**0.262*****	**0.229****	**0.0384**
	(0.0821)	**(0.0958)**	**(0.105)**	**(0.0380)**
Area effects	State	Mine	Mine	Mine
Time effects	Country*year	Country*year	State*year	Country*year
Number of households	22,579	22,579	22,579	22,579
Number of area effects	141	218	218	218
R-squared	0.153	0.13	0.152	0.029

	Never-movers			
	Mean asset wealth			Mean absolute wealth deviation
	(5)	(6)	(7)	(8)
HH close to mine	-0.0381	-0.0352	0.0128	-0.0492
	(0.0758)	(0.0892)	(0.0863)	(0.0493)
Mine operating	0.0822	-0.0585*	0.261	-0.0303
	(0.0663)	(0.0354)	(0.161)	(0.0255)
Mine operating * HH close (DiD)	**0.126***	**0.173***	**0.106**	**0.0223**
	(0.0697)	**(0.0897)**	**(0.0879)**	**(0.0505)**
Area effects	State	Mine	Mine	Mine
Time effects	Country*year	Country*year	State*year	Country*year
Number of households	9,459	9,459	9,459	9,459
Number of area effects	136	205	205	205
R-squared	0.171	0.141	0.167	0.034

Notes. The table reports estimates of equation (2), with area and time fixed effects as indicated. The dependent variable in Columns (1-3) and (5-7) is the asset factor index, with units expressed in standard deviations. The baseline specification is shown in columns (2) and (6). In Columns (4) and (8), the dependent variable is the absolute deviation of a household's asset factor index from the sampling cluster mean, in units of standard deviations. Columns (5-8) restrict the sample to households with at least one respondent who had always been resident in the current location at survey time. Controls include a quadratic in the household head’s age and an indicator for urban/rural status. Standard errors are clustered at the state level in columns (1) and (5), and at the mine level, otherwise. Significant at * 10%, ** 5%, *** 1%.

Table 7
Instrumental variables estimates of the effect on asset wealth in the cross-section and in the panel

	Cross-section		Panel				
					IV robustness checks		
	OLS benchmark	IV	FE benchmark	IV	Full sample	Baseline with all deposits	Baseline with smelters
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
HH close	**0.0758****	**0.0925****	-0.191**	-0.288*	-0.242	-0.284	-0.530**
	(0.0322)	**(0.0448)**	(0.0962)	(0.153)	(0.167)	(0.182)	(0.257)
Mine operating in survey year			-0.0514	0.115	0.267	0.101	-0.00301
			(0.0442)	(0.150)	(0.251)	(0.146)	(0.172)
HH close * mine operating in survey year			**0.370*****	**0.524****	**0.465****	**0.653***	**0.728*****
			(0.110)	**(0.208)**	**(0.232)**	**(0.396)**	**(0.281)**
Number of households	126,434	126,434	14,671	14,671	14,735	20,593	19,780
Number of Groups	618	618	187	187	188	253	200
R-squared	0.181		0.157				
First-stage F statistic		77.45		10.5	2.54	7.33	3.28
Mine-level first stage relationship		0.686***		0.215***	0.174***	0.217***	0.189***
		(0.0265)		(0.0447)	(0.0361)	(0.0445)	(0.0422)
Number of mines		1616		487	491	490	523
R-squared		0.64		0.041	0.042	0.042	0.034

Notes. The table shows cross-sectional and panel IV estimates of wealth effects, alongside fixed effects benchmark estimates computed for the IV sample, and first-stage relationships. Column (2) shows cross-sectional IV estimates following Equation (7); the instruments is an indicator recording whether there is a mineral deposit within 5km of a sampling cluster. Column (1) shows results from the cross-sectional baseline model in Equation (1), estimated on the same sample. For the panel, Columns (4-7) show IV results as in Equation (8). The instrument is an index of the world market price of minerals produced at a given mine, weighted by the preceding year's production. Column (3) shows the benchmark. The dependent variable is the asset index, expressed in units of standard deviations. The baseline panel IV sample in column (4) includes all mines and deposits that are in a stage of exploration or development. It excludes one outlier with an extremely high value of the instrument. Columns (5-7) show IV estimates using modified samples. Column (5) includes the outlier; Column (6) includes a broader range of un-mined deposits; Column (7) includes smelters, despite the poor first-stage fit. All regressions include a quadratic in the household head's age, and an indicator for urban/rural status. The cross-sectional regressions include state-year indicators, and the panel, mine fixed effects and country-year indicators. In addition to first-stage F statistics, we show simple mine-level (as opposed to household-level) OLS results on the first stage relationship between the presence of a deposit and the presence of a mine in the cross-section, and between the price index and operational status in the panel. Standard errors are clustered at the state level in the cross-section, and at the mine level in the panel. Significant at * 10%, ** 5%, *** 1%.

Table 8

Correlation of mine-level wealth effects with measures of development

Country log GDP	-0.0928*	State average years of education	-0.0794**
	(0.0493)		(0.0328)
Number of mines	228	Number of mines	135
R-squared	0.015	R-squared	0.043
State inverse distance to coast	-1.004**	State power line density (log)	-0.127**
	(0.483)		(0.0622)
Number of mines	137	Number of mines	137
R-squared	0.031	R-squared	0.030
Travel time to nearest city	0.0885	Access to land is an obstacle	0.504
	(0.0599)		(0.488)
Number of mines	137	Number of mines	66
R-squared	0.016	R-squared	0.033
State institutional quality	-0.771	Country completed an EITI report	1.500**
	(0.781)		(0.644)
Number of mines	70	Number of mines	228
R-squared	0.023	R-squared	0.039
Country ever participated in EITI	0.122		
	(0.0898)		
Number of mines	228		
R-squared	0.023		

Notes. The dependent variable in the table is composed of mine-level estimates of the cross-sectional effect of closeness to a mine on asset wealth, obtained by estimating equation (1) for each mine separately. (See Appendix C for details.) The regressors either record characteristics of the country in which the mine is located, during the year in which the survey was taken, or of the state in which the mine is located, during the year closest to the survey time for which data was available. All regressions also include log GDP. OLS estimates with conventional standard errors are shown throughout. We find no correlations with the World Bank's CPIA, and omit results for conciseness. Conventional standard errors. Significant at * 10%, ** 5%, *** 1%.

Table 9
Hematologic toxic effects on women in the cross-section

	Altitude-adjusted hemoglobin (g/dL)		Anemia	
	All HHs	Never-movers	All HHs	Never-movers
	(1)	(2)	(3)	(4)
HH close to mine	**-0.0863****	**-0.131**	**0.0262****	**0.0495***
	(0.0438)	**(0.0838)**	**(0.0126)**	**(0.0268)**
Number of women	38,217	13,506	36,225	13,204
Number of groups	934	785	934	784
R-squared	0.0001	0.001	0.0003	0.001

Notes. The table reports estimates of equation (1), using mine-year fixed effects. Columns (1-2) show effects on hemoglobin levels at survey time among adult women; columns (3-4) show results for the incidence of anemia (defined as Hgb below 12 g/dL in non-pregnant women, and Hgb below 11 g/dL in pregnant women). Controls include a quadratic in the respondent's age, and an indicator for urban/rural status. Columns (2) and (4) restrict the sample to respondents who had always been resident in the current location at survey time. Standard errors are clustered at the mine level. Significant at * 10%, ** 5%, *** 1%.

Table 10
Hematologic toxic effects on women in the panel

	Altitude-adjusted hemoglobin (g/dL)			Anemia		
	(1)	(2)	(3)	(4)	(5)	(6)
HH close to mine	0.261*	0.368***	0.396***	-0.0697**	-0.0976***	-0.107***
	(0.151)	(0.137)	(0.146)	(0.0332)	(0.0284)	(0.0292)
Mine operating in survey year	-0.0234	0.0217	0.0852	-0.00281	-0.0208	-0.0277
	(0.117)	(0.121)	(0.136)	(0.0246)	(0.0313)	(0.0309)
Mine operating * HH close (DiD)	**-0.280**	**-0.298***	**-0.330***	**0.0757***	**0.0882****	**0.0966****
	(0.189)	**(0.170)**	**(0.173)**	**(0.0446)**	**(0.0392)**	**(0.0390)**
Area effects	State	Mine	Mine	State	Mine	Mine
Time effects	Country*year	Year	Country*year	Country*year	Year	Country*year
Number of women	9,845	9,845	9,845	9,845	9,845	9,845
Number of area effects	69	122	122	69	122	122
R-squared	0.008	0.011	0.007	0.006	0.007	0.006

Notes. The table reports estimates of equation (2), with area and time fixed effects as indicated. The baseline specification is shown in columns (3) and (6). Columns (1-3) show effects on hemoglobin levels at survey time among adult women; columns (4-6) show results for the incidence of anemia (defined as Hgb below 12 g/dL in non-pregnant women, and Hgb below 11 g/dL in pregnant women). Controls include a quadratic in the respondent's age, and an indicator for urban/rural status. Standard errors are clustered at the state level in columns (1) and (4), and at the mine level, otherwise. Significant at * 10%, ** 5%, *** 1%.

Table 11

Hematologic toxic effects on women near different mine types

	Benchmark		Effect near 'heavy metal' mines		Additional interactions		Falsification tests		
	Hgb (g/dL)	Asset index	Hgb (g/dL)	Anemia	Hgb (g/dL)		Asset index	Miscarriage	Grave illness
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
HH close to mine	-0.0863**	0.145**	-0.0317	0.0123	-0.161*	-0.0285	0.140*	0.00347	0.00366
	(0.0438)	(0.0576)	(0.0533)	(0.0155)	(0.0886)	(0.0507)	(0.0756)	(0.00519)	(0.00595)
HHs close to a 'heavy metal' mine (DiD)			**-0.192****	**0.0466***	**-0.253*****	**-0.192****	**0.0176**	**-0.00377**	**-0.00286**
			(0.0944)	**(0.0247)**	**(0.0876)**	**(0.0902)**	**(0.101)**	**(0.0109)**	**(0.00898)**
Additional interactions					Region	Pregnancy			
Number of women	38,217	25,695	38,217	36,225	38,217	36,225	25,695	117,118	11,022
Number of groups	934	932	934	934	934	934	932	1,469	151
R-squared	0.0001	0.111	0.001	0.0004	0.001	0.027	0.111	0.061	0.011

Notes. Columns (1) and (2) show estimates of equation (1), and columns (3-9) estimates of equation (4), using indicators for each mine-year pair as fixed effects. In columns (3-9), the treatment variable is interacted with an indicator recording whether there is a high expectation of environmental contamination with heavy metals at the nearest mine. Columns (1), (3), (5), and (6) show effects on hemoglobin levels at survey time among adult women. Columns (2) and (7) show effects on asset wealth in households in which respondents in the health sample live. Column (4) shows results for the incidence of anemia (defined as Hgb below 12 g/dL in non-pregnant women, and Hgb below 11 g/dL in pregnant women). Columns (8) and (9) show effects on the incidence of two health conditions not specific to lead exposure. The dependent variable in column (8) is an indicator for whether a woman of reprductive age has ever suffered a miscarriage; in column (9), it is an indicator for whether the respondent was gravely ill for three months or more in the year preceding the survey. Where the dependent variable is a health condition, controls include a quadratic in the respondent's age at survey time and an indicator for urban/rural status. Where the dependent variable is the asset index, a quadratic in the household head's age replaces the respondent's age. Columns (5) and (6) show results from including additional interactions of the treatment variable, as indicated. Standard errors are clustered at the mine level. Significant at * 10%, ** 5%, *** 1%.

Table 12

Recovery of maternal Hgb after birth near heavy metal mines

	Altitude-adjusted hemoglobin (g/dL)						Asset index
	Baseline	Mine-year fixed effects	Control for height-for-age	Control for delivery setting	Placebo treatment (small sample)	Placebo treatment (large sample)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Pregnancy and infancy	-0.558***	-0.581***	-0.547***	-0.429***	-0.470***	-0.652***	-0.0644**
	(0.0421)	(0.0375)	(0.0436)	(0.0454)	(0.0771)	(0.0621)	(0.0255)
HH close to mine	-0.0277	-0.00274	-0.0143	-0.0639			0.0304
	(0.0791)	(0.101)	(0.0839)	(0.0860)			(0.0523)
Pregnancy and infancy * HH close (DiD)	**-0.253****	**-0.185***	**-0.262****	**-0.330*****			**-0.0122**
	(0.0982)	**(0.107)**	**(0.104)**	**(0.116)**			**(0.0562)**
Placebo - HH in lowest wealth quintile					-0.169*	-0.232***	
					(0.0862)	(0.0551)	
Pregnancy and infancy * placebo					-0.0231	0.0783	
					(0.0825)	(0.0538)	
Number of women	5,004	5,004	4,700	3,928	6,851	14,857	4,892
Number of groups	167	521	161	158	139	269	167
R-squared	0.045	0.044	0.043	0.037	0.028	0.031	0.125

Notes. The table shows estimates of equation (6), using indicators for state-year pairs as fixed effects in all columns except column (2). Column (2) shows results using mine-year indicators. The dependent variable is Hgb at survey time among adult women in columns (1-6), and the asset index in column (7). In all columns, the sample consists of observations near mines where heavy metal contamination is to be expected. Where Hgb is the dependent variable, the sample is restricted to women who are currently pregnant, or have given birth within the three years preceding the survey, and who are known to have been resident in the current location since conception. In column (7), the sample is restricted to households in which the women included in the regression in column (1) reside. Columns (5) and (6) give results from a placebo regression, in which the treatment variable is replaced with a placebo indicator that takes value one if the respondent's household is in the bottom wealth quintile, and value zero if it is in the top wealth quintile. In column (5), the placebo sample is restricted to women who are pregnant or have given birth in the past three years, but live in households at least 20km from the nearest mine, and are observed in state-year pairs also represented in the sample in (1). In column (6), it is restricted to women who are pregnant or have given birth in the past three years, live at least 20km from the nearest mine, and are observed in country-year pairs also represented in the sample in (1). Controls include a quadratic in the respondent’s age at survey time and an indicator for urban/rural status in columns (1-6). In addition, the model in column (3) includes the respondent's height for age z-score, and that in column (4) includes an indicator for whether she most recently gave birth in an improved setting. In column (7), a quadratic in the household head's age replaces that in the respondent's age. Standard errors are clustered at the mine level in column (2), and at the state level, otherwise. Significant at * 10%, ** 5%, *** 1%.

Table 13
Health outcomes not specifically linked to heavy metal pollution

	Panel A: Child health outcomes - cross-section				
	Infant mortality	Under-five mortality	Diarrhea	Cough	Fever
	(1)	(2)	(3)	(4)	(5)
HH close to mine	**-0.00246**	**-0.00305**	**0.0112***	**0.00480**	**0.00191**
	(0.00223)	**(0.00270)**	**(0.00579)**	**(0.00963)**	**(0.00788)**
Number of children	298,373	298,373	61,567	60,305	59,494
Number of groups	1,566	1,566	1,510	1,503	1,384
R-squared	0.002	0.003	0.029	0.007	0.01
	Panel B: Child health outcomes - panel				
	Infant mortality	Under-five mortality	Cough	Diarrhea	Fever
	(1)	(2)	(3)	(4)	(5)
Mine operating in exposure period * HH close (DiD)	**-0.00499**	**-0.00819**	**0.00392**	**-0.00260**	**-0.0234**
	(0.00745)	**(0.00864)**	**(0.0299)**	**(0.0282)**	**(0.0258)**
Exposure period	In utero	In utero	Survey year	Survey year	Survey year
Number of observations	43,057	43,057	15,325	15,449	15,576
Number of mines	259	259	236	237	230
R-squared	0.003	0.006	0.025	0.034	0.021

	Panel C: Adult health outcomes - cross-section			
	Ever miscarried	Night blindness during pregnancy	Female respondent very sick	Male respondent very sick
	(6)	(7)	(8)	(9)
HH close to mine	**0.00263**	**0.00254**	**0.00328**	**0.0120**
	(0.00460)	**(0.0104)**	**(0.00527)**	**(0.00977)**
Number of respondents	117,118	29,317	11,022	9,808
Number of groups	1,469	1,185	151	151
R-squared	0.061	0.001	0.011	0.011
	Panel D: Adult health outcomes - panel			
	Ever miscarried	Night blindness during pregnancy	Female respondent very sick	Male respondent very sick
	(6)	(7)	(8)	(9)
Mine operating in exposure period * HH close (DiD)	**-0.00236**		**-0.00845**	
	(0.0152)		**(0.0119)**	
Exposure period	Survey year	n/a	Survey year	n/a
Number of observations	29,666		4,111	
Number of mines	202		63	
R-squared	0.065		0.005	

Notes. The table reports estimates of equation (1) in the rows marked 'cross-section', and estimates of equation (2) in the rows marked 'panel'. In the latter, treatment variables are defined using the time period of exposure to pollution most appropriate to each health condition, as indicated. Only the difference in differences coefficient is reported. Cross-sectional models use indicator variables for each mine-year pair as group fixed effects; panel models, mine fixed effects and survey round dummies. The dependent variable in columns (1) and (2) is an indicator for whether a child died within the first year and the first five years after birth, respectively. In the other columns, it is an indicator for whether the respondent suffered the condition indicated - over the two weeks preceding the survey (3-5); at any point during her reproductive life (6); during the most recent pregnancy (7); or for three months or more during the year preceding the survey (8-9). Controls in columns (1-5) include an indicator for urban/rural status in all columns, a quadratic in the mother's age at birth, an indicator for gender, birth-order indicators, as well as indicator variables for the child's age (columns 3-5 only). In columns (6-9), they include an urban/rural indicator, and a quadratic in the respondent's age at survey time. In cells marked 'n/a', the model could not be estimated. Standard errors are clustered at the mine level. Significant at * 10%, ** 5%, *** 1%.

Table 14
Hematologic toxic effects on children

	All children under five years of age		Infants					
	Cross-section	Mine-type DiD	Cross-section	Mine-type DiD	Mine-type DiD falsification tests			
	Hgb (g/dL)	Hgb (g/dL)	Hgb (g/dL)	Hgb (g/dL)	Asset index	Diarrhea	Cough	Fever
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HH close to mine	-0.0653	-0.0418	-0.0487	-0.0464	0.179***	0.00739	-0.00328	0.00196
	(0.0460)	(0.0516)	(0.0505)	(0.0569)	(0.0647)	(0.00732)	(0.0116)	(0.00923)
HHs close to a 'heavy metal' mine (DiD)		-0.109		-0.0184	0.0887	-0.000835	0.0492**	0.00500
		(0.108)		(0.119)	(0.115)	(0.0139)	(0.0216)	(0.0192)
Child in infancy			-0.358***	-0.359***	0.0318	0.0339***	0.0107*	0.0132*
			(0.0598)	(0.0651)	(0.0240)	(0.00579)	(0.00613)	(0.00699)
HH close to mine, and child in infancy			-0.0823	0.0643	-0.0175	0.0120	-0.00890	-0.00822
			(0.102)	(0.110)	(0.0617)	(0.0134)	(0.0120)	(0.0134)
Nearest mine (≤ 20km) is a 'heavy metal' mine, and child in infancy				0.00959	0.0850	-0.0198*	0.00249	0.00593
				(0.155)	(0.0750)	(0.0112)	(0.0119)	(0.0126)
HH close to a 'heavy metal' mine, and child in infancy (triple difference)				**-0.597*** **	**-0.309** **	**-0.0163**	**-0.0212**	**-0.0122**
				(0.199)	**(0.147)**	**(0.0240)**	**(0.0262)**	**(0.0234)**
Number of children	18,029	18,029	18,029	18,029	12,697	61,567	60,305	59,494
Number of mines	907	907	907	907	901	1,510	1,503	1,384
R-squared	0.068	0.068	0.021	0.022	0.120	0.006	0.002	0.002
DiD effect on infants				**-0.616*** **	**-0.221**	**-0.0171**	**0.028**	**-0.0072**
				(0.201)	**(0.158)**	**(0.0232)**	**(0.0283)**	**(0.0261)**

Notes. Columns (1) and (3) show estimates of equation (1); columns (2) and (4-8) show estimates of equation (4). In column (3), the treatment variable in equation (1) is interacted with an indicator for whether the child was in her first year of life at survey time. In columns (4-8), the treatment variable and its interaction in equation (4) are interacted with the indicator variable for infancy. All columns indicators for each mine-year pair as fixed effects. Columns (1-4) show effects on hemoglobin levels at survey time. Columns (5-8) show effects on asset wealth and health conditions not specific to lead exposure; in columns (6-8), the dependent variables record whether a child suffered from the respective condition in the two weeks preceding the survey. Where the dependent variable is a health condition, controls include an indicator for urban/rural status in all columns, a quadratic in the mother's age at birth, an indicator for gender, birth-order indicators, as well as indicator variables for the child's age. Where the dependent variable is the asset index, controls include a quadratic in the household head's age, and an urban/rural indicator. The row labeled "DiD effect on infants" shows the sum of the coefficients "HHs close to a 'heavy metal' mine" and "HH close to a 'heavy metal' mine, and child in infancy", and tests the hypothesis that the sum is zero. Standard errors are clustered at the mine level. Significant at * 10%, ** 5%, *** 1%.

Table 15
Growth effects on children in the cross-section

	All children under five years of age			Infants				
	Benchmark	Never-movers only	Effect near 'heavy metal' mines	Benchmark	Effect on infants near 'heavy metal' mines			
	Height for age	Height for age	Height for age	Height for age	Height for age	Stunting	Severe stunting	Asset index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HH close to mine	0.0828**	-0.00153	0.0754*	0.0760*	0.0598	-0.0116	-0.00884	0.111**
	(0.0389)	(0.0497)	(0.0430)	(0.0418)	(0.0464)	(0.0103)	(0.00742)	(0.0469)
HHs close to a 'heavy metal' mine (DiD)			0.0378		0.0726	-0.0263	-0.00668	-0.0384
			(0.0988)		(0.102)	(0.0241)	(0.0163)	(0.0929)
Child in infancy				0.805***	0.781***	-0.164***	-0.0717***	0.0156
				(0.0368)	(0.0432)	(0.0112)	(0.00675)	(0.0156)
HH close to mine, and child in infancy				0.0304	0.0719	-0.00459	0.00597	0.00945
				(0.0528)	(0.0631)	(0.0136)	(0.00824)	(0.0403)
Nearest mine (≤ 20km) is a 'heavy metal' mine, and child in infancy					0.102	-0.0571***	-0.0175	-0.0263
					(0.0729)	(0.0217)	(0.0108)	(0.0317)
HH close to a 'heavy metal' mine, and child in infancy (triple difference)					**-0.170**	**0.0696****	**0.0280***	**-0.0710**
					(0.110)	**(0.0327)**	**(0.0166)**	**(0.0693)**
Number of children	40,552	16,927	40,552	40,552	40,552	40,552	40,552	28,540
Number of groups	1,244	1,041	1,244	1,244	1,244	1,244	1,244	1,243
R-squared	0.064	0.054	0.064	0.062	0.062	0.036	0.015	0.084
DiD effect on infants					**-0.097**	**0.043**	**0.021**	**-0.109**
					(0.131)	**(0.029)**	**(0.018)**	**(0.107)**

Notes. Columns (1), (2) and (4) show estimates of equation (1), and columns (3) and (5-8), estimates of equation (4). In column (4), the treatment variable in equation (1) is interacted with an indicator for whether the child was in her first year of life at survey time. In columns (5-8), the treatment variable and its interaction in equation (4) are interacted with the indicator variable for infancy. All models use indicators for each mine-year pair as fixed effects. Columns (1-5) show effects on height for age z-scores. The dependent variable in Column (6) is the prevalence of stunting, defined as a height of two σ or more below the median; in Column(7), it is the prevalence of severe stuting, defined as a height of more than three σ below the median. Column (8) shows effects on asset wealth in households in which children in the height-for-age sample live. Where the dependent variable is a health condition, controls include an indicator for urban/rural status in all columns, a quadratic in the mother's age at birth, an indicator for gender, birth-order indicators, as well as indicator variables for the child's age. Where the dependent variable is the asset index, controls include a quadratic in the household head's age, and an urban/rural indicator. The row labeled "DiD effect on infants" shows the sum of the coefficients "HHs close to a 'heavy metal' mine" and "HH close to a 'heavy metal' mine, and child in infancy", and tests the hypothesis that the sum is zero. Standard errors are clustered at the mine level. Significant at * 10%, ** 5%, *** 1%.

Table 16
Comparative growth effect of in utero and birth-year exposure in the panel

	In utero exposure			In utero exposure, infants only			In utero vs. birth year exposure		In utero vs. survey-year exposure		In utero and birth year exposure, mother fixed effects		
	Height	Stunting	Severe stunting	Height	Stunting	Severe stunting	Height		Height		Height		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
HH close to mine	0.191**	-0.0501**	-0.0424***	0.535**	-0.126***	-0.0448	0.0972	0.141	0.0743	0.137			
	(0.0808)	(0.0251)	(0.0110)	(0.209)	(0.0472)	(0.0346)	(0.0902)	(0.0896)	(0.107)	(0.0963)			
Mine operating during pregnancy	0.0212	0.000151	-0.00102	0.0930	-0.0369	-0.0135		0.0950		0.0953	0.175		0.0855
	(0.0614)	(0.0178)	(0.0124)	(0.124)	(0.0307)	(0.0141)		(0.0712)		(0.0826)	(0.161)		(0.173)
Mine operating during pregnancy * HH close	**-0.136**	**0.0534***	**0.0496*****	**-0.371***	**0.149*****	**0.0512**		**-0.494*****		**-0.417****	**-0.460**		**-0.422**
	(0.0981)	**(0.0274)**	**(0.0140)**	**(0.217)**	**(0.0510)**	**(0.0346)**		**(0.139)**		**(0.185)**	**(0.361)**		**(0.359)**
Mine operating in second exposure period							-0.00443	-0.0843	-0.00185	-0.0908		0.197	0.169
							(0.0595)	(0.0692)	(0.0546)	(0.0922)		(0.154)	(0.172)
Mine operating in second exposure period * HH close							-0.00860	0.423***	0.0474	0.356*		-0.163	-0.0540
							(0.110)	(0.149)	(0.130)	(0.189)		(0.424)	(0.418)
Second exposure period							Birth year		Survey year			Birth year	
Number of children	11,629	11,629	11,629	2,426	2,426	2,426	11,629	11,629	11,155	11,321	11,629	11,629	11,629
Number of fixed effects	200	200	200	186	186	186	200	200	188	191	9,408	9,408	9,408
R-squared	0.113	0.072	0.055	0.091	0.100	0.058	0.113	0.114	0.073	0.117	0.204	0.204	0.205

Notes. Columns (1-10) report estimates of equation (2), and Columns (11-13), estimates of equation (3). Columns (1-10) use indicators for each country-year pair as time effects, and columns (11-13) use country linear time trends. In columns (1-6) and (11), treatment is defined as exposure to mining *in utero*. Columns (7-10) and (12-13) compare this to the effect of exposure during a second exposure period, as indicated. The dependent variable in columns (1), (4), and (7-13) is the height-for-age z-score. In columns (2) and (5), it is the prevalence of stunting, defined as a height of two σ or more below the median; in Columns (3) and (6), it is the prevalence of severe stuting, defined as a height of more than three σ below the median. Controls include an indicator for urban/rural status in all columns, a quadratic in the mother's age at birth, an indicator for gender, birth-order indicators, as well as indicator variables for the child's age. For consistency across models, the sample is restricted to those observations where the operating status of the mine is known both in the birth year and during gestation (this removes about 3% of observations). Standard errors are clustered at the mine level. Significant at * 10%, ** 5%, *** 1%.

Table 17
Growth effects on children born to migrants

	Height for age		Falsification tests			
			Improved delivery		Infant mortality	
	(1)	(2)	(3)	(4)	(5)	(6)
HH moved to within 5km of a mine	0.360***		-0.00902		0.00911*	
	(0.0988)		(0.0274)		(0.00542)	
Child born after move	0.0541	0.214	-0.00479	0.0438	-0.00836***	0.00700
	(0.0455)	(0.163)	(0.0117)	(0.0420)	(0.00251)	(0.00860)
Child born after move * HH now within 5km of mine	**-0.169***	**-0.190**	**0.0492***	**0.0691**	**-0.00547**	**-0.0142***
	(0.0971)	**(0.336)**	**(0.0276)**	**(0.0747)**	**(0.00564)**	**(0.00842)**
Group effects	Mine	Mother	Mine	Mother	Mine	Mother
Number of children	7,882	7,882	8,404	8,404	56,652	56,652
Number of groups	891	6,940	896	7,395	1,756	37,598
R-squared	0.043	0.131	0.035	0.043	0.012	0.016

Notes. The table shows estimates of equation (5), with fixed effects as indicated. The sample is restricted to children born to migrants, within no more than four years of the time of migration, and living in locations no more than 20km from a heavy metal mine, smelter, or legacy. Child age is recorded in months, but residence in years. Whether a child is born after the family moved is therefore ambiguous for some children; we exclude these from the sample. The treatment variable takes value zero if a child was with certainty not exposed at any time during gestation, and value one when she was with certainty exposed for the entire gestation period. The dependent variables are the height-for-age z-score in columns (1) and (2). Falsification tests are shown for whether birth was given outside the home, with some assistance (columns 3-4), and for infant mortality (Columns 5-6). Controls include a quadratic in the mother's age at birth, indicator variables for the child's gender, birth order indicators, and indicators for the child's age at survey time, and an indicator for urban/rural status. Standard errors are clustered at the mine level. Significant at * 10%, ** 5%, *** 1%.

Chapter 3

Sensitivity to price and demand for arsenic testing of drinking-water wells: Evidence from Bihar, India

Sensitivity to price and demand for arsenic testing of drinking-water wells: Evidence from Bihar, India

Prabhat Barnwal (Columbia University, USA)

Jan von der Goltz (Columbia University, USA)

Chander Kumar Singh (TERI University, India)

Alexander van Geen (Columbia University, USA)

Abstract: Groundwater contaminated with arsenic of natural origin poses a serious threat to the health of tens of millions of villagers across South and Southeast Asia. With a field experiment conducted in Bihar, this study estimates the demand for testing well water after this service was offered at difference prices in 26 villages. The test relies on a field kit and requires less than 15 min. We find that demand is highly sensitive to price in the 10-50 Rs. range. We further study whether the use of information provided after testing is sensitive to the price paid. We also estimate additional demand after a repeated round of campaign was conducted in the same villages and explore wealth and learning as potential factors. Finally, the study provides empirical evidence that households proactively try to hide bad news regarding the status of their well with respect to arsenic.

Key words: *Development, Public Health, Arsenic*

I. Introduction

The high social benefits associated with health products – such as insecticide-treated bed nets to prevent malaria infection, or water filters to get rid of microbial pathogens – form the basis for a compelling case for providing a full subsidy in low-income settings, where willingness to pay is limited even for very effective health interventions (Dupas 2014a). Yet, public provision is beset with difficulties, from slow and unreliable provision to poor targeting of the free good toward intended beneficiaries. Innovation in products and delivery is commonly stifled. Cost-sharing is often suggested as a way to reduce dependency on public programs, but has often been found to significantly affect take-up. (Kremer and Miguel 2007; Dupas 2014a).

This study considers the effect of fee-based provision on demand in the case of tests to ascertain the arsenic content of tubewell water. Arsenic tests are a highly efficient health intervention: the cost of a test provided through our program was a mere USD 2. The information that tests provide is not substitutable: the safety of a well cannot be determined or even 'guessed' without a test. The distribution of arsenic incidence in groundwater is difficult to predict, and varies greatly even over small distances. A well that meets the WHO guidelines for arsenic in drinking water may be found in immediate neighborhood of a very unsafe well (*Figure 1*). Within shallow (<100 m) aquifers tapped by most private wells, there is no systematic and predictable relationship between and arsenic and the well depth[1]. At the same time, precisely because arsenic contamination varies greatly in space, well tests make available an effective way to avoid exposure, namely switching to nearby safe wells.[2] (Ahmed et al. 2006; Opar et al. 2007; Madajewizcz et al. 2007; Chen et al. 2007; Bennear et al. 2012; George et al. 2012, Pfaff et al. 2015).

Finally, the health consequences avoided by ending chronic arsenic exposure are dramatic. Chronic exposure to arsenic by drinking groundwater at over 10 times the level of the current World Health Organization guideline of 10 microgram per liter has recently been shown to double all-cause deaths in a large cohort study conducted in Bangladesh (Argos et al. 2010). Arsenic in tubewell water has also been associated with impaired intellectual and motor function in children (Wasserman et al. 2004; Parvez et al. 2011). In consequence, arsenic has been found to have a significant effect on income and labor supply: Pitt et al. (2013) estimate that lowering the amount of retained arsenic among Bangladesh prime-age males to levels encountered in uncontaminated countries would increase earnings by 9%.

[1] Malgosia (2007) show that arsenic incidence is uncorrelated with household characteristics, a finding which we confirm later in this paper

[2] For example, Opar et al. (2007) find that 68% households are likely to switch, if there is a safer well within 50m.

Matching households to arsenic exposure, Carson et al. (2011) find that overall household labor supply is 8% smaller due to arsenic exposure.

Because of their low cost and important health benefits, tens of millions of arsenic well test have been carried out through public provision in rural communities across the Indo-Gangetic Plain (Fendorf et al. 2010). However, these important programs may need complementing. Thus, after a single blanket testing covering 5 million wells by the government of Bangladesh in 2000-2005, no such country-wide public programs have been carried out. In consequence, recent estimates suggest that more than half of currently used tube wells in Bangladesh have never been tested for arsenic (van Geen et al., 2014). This prompts the question whether a cost-shared provision might be sustainable, and whether there is the prospect of a market for arsenic tests in which local entrepreneurs would have an incentive to seek out untested wells (Miller and Babiarz 2013).

In this paper, we use a randomized control trial to estimate demand for arsenic testing of water wells, when offered at a price. We investigate the determinants of the demand, as well as households' behavioral response to the information regarding arsenic status of private wells. In order to estimate the price elasticity of demand, we randomize assignment of price for arsenic testing of tubewells at village levels in 26 villages in Bihar, India. Five different levels of prices are assigned between Rs. 10 to Rs. 50, with the highest level approximately equal to one day of per capita income in Bihar.[3] The program offered to test household's tube wells for arsenic contamination, if they agreed to pay the assigned price. The testing campaign was carried out over two years, with test being offered twice in the same villages and at the same pre-assigned prices – first in 2012, and later in 2014. After the first phase of

[3] In 2011-12, Per capita daily income in Bihar was Rs. 45 (http://www.indiaenvironmentportal.org.in/files/file/Economic-Survey-2014-bihar.pdf).

testing, we conducted a follow up survey on how the information about arsenic was being used by the households. In the second phase, we carried out a detailed survey of the households, including their recall of the information provided by the diagnostic test.

Two limitations arising from the study's implementation are worth highlighting. A review of the field work finds that the first phase of test sales campaign was not geographically complete and did not entirely cover some of the villages. The missing area is quite large but is mainly concentrated in five villages. This is likely because second phase involved a more systematic door-to-door campaign. Secondly, an attempt to create a well owner-level panel was unsuccessful, since the well tags attached during the first phase proved to be less durable than expected, and could not be comprehensively tracked.

We find that there is a considerable demand for arsenic testing: at the mean across price groups, 47% of households purchased the test. However, the willingness to pay for arsenic testing is highly sensitive to price, and demand drops steeply with price. Our findings align well with other studies on cost-sharing of preventive health care products which has found relatively high price sensitivity of demand despite large private returns (e.g., Kremer and Miguel 2007, Cohen and Dupas 2010, Meredith et al. 2013). We record significant additional demand for arsenic test when the test is offered again after two years in the same villages. At the same time, more than half of the households decided to not purchase the test. It is not surprising that household wealth comes out as a major determinant of the decision to purchase the test. We further find three additional results. We don't find the use of information – i.e. switching to be sensitive to the price paid. In a follow up after three months, about 30% households who had an arsenic-high well, self-report to having switched to a safer tube well for their drinking and cooking water needs. Secondly, despite arsenic testing being a non-experience good (such as a mosquito net) and existing constraints to switching, we find a significant increase in demand when the test was offered again. Finally, we find evidence on selective retaining of water source quality information. We document households' proactive behavior to discard information on "bad news", when the diagnostic test result is positive,. Stigma and restrictions on water access based on affiliation to social groups may explain this.

Recent studies have looked at the question of price vs. subsidies of products with large private and public benefits. Cost-sharing (i.e. pricing) is often favored by development practitioners to ensure sustainability, to reduce the burden of 'entitlement effect' of subsidies and for better self-targeting, but it is found that take-up drops steeply with prices for products, when people have low private valuation (Kremer and Miguel 2007, Cohen and Dupas 2010, Meredith et al. 2013). Furthermore, pricing a product can also change the likelihood how the product will be used through screening, sunk cost or signal of quality (Ashraf, Berry and Shapiro 2010, Cohen and Dupas 2010). However, most of the studies focus on products which are repeatedly used, such as drinking water disinfectant, mosquito net and deworming pills, and all such products are already known to the people. To our knowledge, no study has attempted to estimate demand curve for diagnostic testing of water source quality for arsenic. One related study is conducted by George et al. (2013) who look at the impact of education and media campaign of increasing adoption of fee-based arsenic testing at a single price. Our study further contributes to this literature by investigating how household respond to the information regarding arsenic status of their well and whether there is any effect of price paid on switching to safer water sources.

At the policy level, this is the first study which shows that it is feasible to provide arsenic testing in a cost-shared way. Cost-sharing has its own limitations because of overexclusion[4] concerns, but similar to vaccination programs which are generally available outside of the public health system, diagnostic testing of arsenic can complement the government programs. We here also propose a different aspect on how continuously offering a product increases demand over time -- the coverage increases when the product is offered second time. This is important from a policy perspective, when discussion on

[4] Dupas (2014a) define 'overexclusion' as the number of people who would use the product and become healthier, do not take it up because of cost sharing.

continuing subsidies or switching to cost-sharing is essentially dominated by studies which offer priced product only once or within a short time window. [5]

The rest of the paper is structured as follows. Section 2 discusses the details of the experiment and data. Results are discussed in Section 4 and Section 5 concludes.

II. Details on Experiment, Data and Methodology

For the purpose of this study, we focus on a region in the Indo-Gangetic plains in Bihar, India, where geologic factors suggest arsenic levels could be elevated in a significant proportion of wells. The area has not previously been covered by government-sponsored blanket testing of wells, although one study reports arsenic testing of about 5000 water wells in the study district and reported that 26% of wells were unsafe (Nickson et al. 2014). As elsewhere in South Asia, arsenic tests are not available on the private market. Within the general study area, we selected Bhojpur district to conduct our intervention; the district contains 1045 villages according to the 2011 Census. Within this district, we chose 26 villages for this study based on a shared high probability of arsenic incidence, as indicated by the distance from the river and within four blocks (sub-districts) in the Bhojpur districts.[6] These villages are of moderate size with population varying from 50 to about 400 households. Our endline survey estimates 4084 well-owner households in these 26 villages, which is about 75% of total number of private well-owner households as counted in 2011 Census.

To elicit demand, we use a revealed preference method, namely, making take-it-or-leave-it offers of arsenic tests at a certain price to households in the sample villages. As is immediately obvious, a take-it-

[5] Some studies which only look at one time or limited time window offer (Meredith et al. 2013) and three months voucher (Dupas 2014b).

[6] The original intention was to work in a sample of 25 villages, i.e., five villages in each price group. However, since enumerators erroneously worked in two villages of the same name during initial field work, we included the additional village for the rest of the program.

or-leave-it offer elicits only a bound of each household's willingness to pay. For instance, if a household accept to purchase a test at USD 1, we can only infer that its willingness to pay was at least USD 1. Similarly, a case of rejection only suggest that the willingness to pay is less than the offered price.

We then randomly assigned each village to one of five price levels – Rs. 10, Rs. 20, Rs. 30, Rs. 40 and Rs.50 – at which households were offered arsenic tests for purchase. It was felt that offering different prices in the same village would be seen as violating fairness norms, and would deter purchases of the tests. We therefore chose not to randomize our prices within villages. The highest price was chosen based on initial focus group discussion at local level. It is slightly higher than the daily per capita average income of Rs. 45 in Bihar during 2011-12. In a pilot study, tests had been offered free of charge in four villages in the intervention area; consistent with experience in other settings, uptake was close to 100%. We therefore did not add a treatment arm that would have offered tests free of charge.

Testers were recruited and trained prior to the roll-out of the campaign. In the early phase of testing in 2012, a focus group meeting was organized in each village. A sample poster showing a satellite image of a pilot village along with color markers indicating the arsenic status of tested wells was put on display to increase awareness about the arsenic issue (Figure 2). Following the focus group meetings, testers began to offer tests using a standard arsenic test kit. GPS locations of households approached with a test offer were collected, along with basic data on the household. However, testers did not record data from all households that did not purchase a test, as we find in the recall of 2012 offers by households in 2014. During the initial wave of test offers, a total of 1,212 tests were sold across the 26 sample villages. The results of each test were posted on each pump-head with an easy-to-read metal placard color coded red for unsafe wells (>50 ug/L arsenic), green for 'borderline safe' wells where arsenic is of some concern (>10-50 ug/L), and blue for safe wells (<10 ug/L) (*Figure 3*).The cut-off values were chosen to correspond with the Indian national safety standard for arsenic of 50 ug/L that was current as of the time of the test campaign, and the WHO guideline of 10 ug/L (the government of India has since matched its standard for arsenic in drinking water to the WHO guideline). Smaller placards with a unique well ID were also attached to each pump-head in anticipation of a future response survey. In 2012-13, right after first wave of arsenic testing was completed, village level maps were exhibited in the villages with the geo-locations of safe (Blue), borderline safe (Green) and unsafe (Red) wells. During home visits, households were also alerted to the fact that switching from unsafe or borderline safe wells to any neighboring safe wells would be an effective way to avoid exposure. The first phase of the project concluded with a follow up survey conducted approximately three months later. Enumerators visited all households that had purchased a test, and collected information on their behavioral response to the information on arsenic status of their well – and in particular, on whether households now drew water from neighboring safe wells.

In a second phase, commencing in 2014 – some two years after the initial visits – we offered the test again in the same set of villages, and at the same price assigned initially. A total of 4084 households were approached with the intention of making a sales offer across the 26 villages.[7]\ Data was collected systematically from every household where a respondent could be interviewed, including from households that did not wish to buy the tests. Each house was visited at least two times to ensure maximum coverage. After two visits, about 13.6% of households could not be surveyed because no adult member was present or willing to answer questions. The resulting sample contains data from 3528 households. A total of 719 tests were sold in this second phase. The household survey administered in the second round gathered detailed socio-economic and demographic information, along with GPS locations of the wells. It also collected information on recall of tests being offered and purchased in 2012, along with test results. This recall data allows us to work around some of the constraints posed by

[7] We cross-checked the number of households approached against 2011 Census data for 21 out of 26 villages that could be matched to the census. For these villages, the census shows 4497 households that own a hand pump, whereas we record 3322 attempted sales in the same 21 villages – that is, 74% of the census population

the fact that households visited in 2012 and in 2014 cannot be matched with confidence, since the names of residents and the exact address were not comprehensively collected, and only a small percentage of well tags placed in 2012 were still attached to previously tested wells in 2014.

Summary statistics from the 2015 survey reflect moderately well-off village communities. Households are of moderate size (3.9 members on average). Most households (89%) own at least one mobile phone, and most (70%) live in houses made from durable building materials ("pucca"). Ownership of bikes (68%) and cows (67%) is common, though fewer households own household durables or have access to sanitation, and very few own cars (*Table 1*).

A randomization check shows that the price category dummies are jointly significant for two out of the eleven variables tested (*Table 2*). The two instances where there are significant differences (ownership of cars and access to sanitation) appear isolated, and there is no indication that the price groups in question are generally any more or less wealthy than the other groups.

In the empirical analysis, ordinary least square regression (OLS) is used. In all the regressions, we report cluster bootstrapped standard errors.

III. Results:

3.1 Demand for well arsenic testing

We find overall substantial demand for arsenic tests – but demand is highly sensitive to price. Overall, after adjusting for repeat purchases, a total of 1931 tests were sold at randomly assigned prices across the 26 sample villages over the entire duration of the program (2012-2014). This implies that arsenic testing covered about 47% of households approached for sales (Col 12 in *Table 3*).[8] A map displaying the proportion of safe, unsafe and untested wells in each village is shown in *Figure 4*.

Digression: Comparison of recorded and recall sales data

As discussed earlier, during the first offer phase in 2012, testers did not systematically collect data from households that did not want purchase the test (Column 2 and 4, *Table 3*). In addition, there is qualitative evidence that testers offered tests less systematically in parts of the villages where people showed strong reservations against the idea of arsenic tests being offered for a fee (rather than free of charge) during focus group meetings.

We hence face a considerable challenge in reliably assessing baseline demand, since the number of households to whom the test was offered in 2012 cannot be directly ascertained. *Figure 5* shows the number of offers and sales in 2012 survey and in the endline survey in 2014. We address this challenge with the following strategy. (1) We first estimate demand based on the assumption that as many households were approached during the 2012 campaign as during the 2014 campaign (Column 9, *Table 3*). This estimate is correct to the degree that (i) sales approaches were comprehensive in 2012, and (ii) the number of households has remained constant between survey rounds. To assess whether these assumptions – and the resulting estimate – are reasonable, we (2) collect in the follow-up survey recall data both on whether households were offered the test at baseline, and on whether they purchased the test at baseline (Panel (II) in *Table 3*). As is evident, the estimates never significantly diverge (e.g. total demand is 30% in recorded data (Col 9) and 27% in recall data (Col 10)), and are very well aligned in the Rs. 10-30 groups, and reasonably aligned at higher prices. There is a good match between the ratio of *recalled* 2012 sales to *recorded* 2012 sales (0.65) on the one hand, and the ratio between *recalled* 2012

[8] This final figure if adjusted for repeat purchases to avoid any double counting would be 46%. In total, 74 households who recalled having bought the test in 2012 purchased another test in 2014.

offers and *recorded* 2014 sample size on the other (0.60). We also compare calculated demand with recorded data and recall data at all the price levels, and find no significant difference in estimates except at higher prices (*Figure 6*). This pattern indicates that the count of households (as ascertained by their GPS location) who were not interested in buying the test in 2012 is far from complete in the recorded data. In the following, we discuss results based on recall data – arguably, the more internally consistent estimate.

Price sensitivity of demand

In line with prior research, we find that demand is highly sensitive to price. The mean elasticity across sales at different prices in our data is -0.36, a point estimate very close to the mean elasticity -0.37 provided in Cohen and Dupas (2010), and well in line with other recent studies on the demand of similar products in developing countries (Meredith et al. 2013).[9]

To estimate the demand curve after two offer phases, we rescale recalled purchases with recall loss rate of 35%. This is combined with the demand observed in 2014. The resulting estimate is shown in *Figure 7*. At the lowest price (Rs. 10) about 67% of households purchase the test. While our experiment did not include an arm with zero price offer, uptake of free tests can be assumed to be almost 100%.[10] Thus, while there is significant demand at a moderate price of Rs. 10 (USD 0.15), charging this small amount, rather than offering the test for free, reduces coverage by about one-third. For comparison, Kremer and

---

[9] Alternative techniques, such as the Becker-DeGroot-Marschak (BDM) mechanism and other auction-based methods might have provided richer information than our take-it-or-leave-it design. However, they would have been significantly challenging to implement in the field. More fundamentally, auctions would have been unlikely to be efficient mechanisms, given the potential buyers' uncertain and likely correlated beliefs over the value of arsenic tests. As noted earlier, tests are not sold in the market, so that households are quite unfamiliar with the technology – and it is in any case difficult to appreciate the long-term benefits of reducing arsenic exposure.
[10] We didn't have a zero price arm because of "effect of price 0" is well established and is found in lab (Ariely and Shampan'er 2001, as cited in Cohen and Dupas 2011) and field experiments (Kremer and Miguel 2007). Nevertheless, we have 4 pilot villages where these tests were offered for free, before we started the experiment with randomized prices in 26 villages. While survey data is not systematically collected from those villages, we observed almost 100% uptake.

Miguel (2007) find introduction of a small fee for deworming drugs led to an 80 percent decrease in take-up. Demand further drops precipitously at higher prices, and reduces to 21% at the highest price offered, Rs. 50.

Demand at repeat offer

In 2014, in each village, tests were offered again at the same price as the initial offer. The second phase was carried out approximately two years after the first round. Our purpose behind re-offering the arsenic test was to assess whether additional demand (i.e. from households who did not purchase in the first phase) would materialize after a two-year delay – in the distinct context of a product whose characteristics are not familiar to potential customer ex-ante. We find that repeating the offer again after a two-year delay did indeed generated significant additional demand. Particularly, about 25% of those households who were offered the test in 2012 but didn't purchase it, decided to purchase the test in 2014. When compared to the number of tests sold in 2012, purchases increase by 59% following the second offer. This implies an 18 percentage point (pp) increase in total coverage (defined as take-up as a percentage of the total number of households). About 70% of the new purchases in 2014 (i.e. 502 out of 719 tests) are made by households who recall being offered the test in 2012 but didn't purchase. We cannot completely rule out the concern that the additional demand and second phase comes from the households who were not covered during 2012 campaign. However, after accounting for even a modest loss in recall, failure to reach all the households in 2012 can only be attributed to 30% of new purchases in 2014.

Overall, the key patterns in demand in second phase are similar as observed in the first round of sales: demand sensitivity to price is higher at left tail (*Figure 8*). While the total coverage -- after the repeat offer-- reaches 47%, this aggregate result masks very stark differences across price groups, which arise for the following reason. At any price point, there is substantial demand at the repeat offer, from 74% of the original sales at Rs. 10 to 53-65% at higher prices. Yet, because of the steep drop in demand at higher prices, this leads to even bigger divergences in achieved coverage across the price levels. From a policy perspective, the effect of making a repeat offer is remarkable: price matters greatly for demand, but at *any* price level that may be desirable for policy or business reasons, repeating the offer strongly increases coverage (and from a business perspective, sales). Irrespective of the channels – learning, awareness or marketing intensity, this simple finding underscores the need for a more careful assessment of experimental evidence generated with one time offers or offers given for a short period.

Selection effect

We test how household wealth is correlated with prices, among the households who decide to purchase the test, using end line survey data from 2014. OLS regression for key asset variables are shown, which primarily indicate that selection was mainly limited to high price levels. Investment in sanitation – i.e. having a Latrine facility in the house, is highly correlated with purchase decisions at high price levels, which suggest that hygiene conscious households have higher willingness to pay This is worth to note that there is no large difference between buyers in Rs. 10 and Rs. 20 group, although we see a large drop in take-up. Separate analysis of first phase buyers also suggest selection at higher prices (Results not shown).

Why is there substantial demand at the time of the repeat offer?

The high level of demand (between 53-74% of demand at baseline) at the time of a repeat offer made within the relatively short time frame of two years raises the question why so many additional customers are accept to purchase the test when offered again at a substantial gap of time. Kremer, Leino, Miguel and Zwane (2011) et al little evidence on any significant change in valuation of households within one year of lag in spring protection treatment, though they mention that one year may be insufficient for learning.

We explore two possible explanations, namely (i) changes in wealth between the first and second offer, and (ii) learning. A gap of two years between two phases of the program provide us a rich setting where any substantial demand at repeat offer can be explained by household's correction of their view about the valuation of arsenic testing or increase in income/wealth. On the other hand, per capita real income in Bihar increased at the rate of about 10% per year between 2012 and 2014, which can also cause additional demand by pure income effect.[11] Because our nominal prices remained constant in two phases, an effect in similar direction can be expected because of inflation. Since we don't have income data in our sample, we compare wealth in early and later purchasing households using information collected on assets in the endline survey.[12] We also look at differences in demography, education and social status.

(i) Wealth effects

The data suggests clearly that 2012 buyers were at least less well-off than 2014 buyers when *both* are observed in the 2014 survey (i.e., households that *recall* buying a test at the first offer, and those that

[11] State Gross GDP growth table for India is available at http://planningcommission.nic.in/data/datatable/data_2312/DatabookDec2014%20157.pdf

[12] Regrettably, wealth questions asked in the 2012 and 2014 survey rounds were meaningfully different, so that it is hard to assess whether *at the time of purchase*, households that bought the tests in 2012 were wealthier than those who bought in 2014.

bought a test at the time of the second offer) (*Table 6*). We select this test, rather than comparing data collected in 2014 to data collected in 2012 because (a) it removes shared growth effects between the two offers (buyer groups may have different growth trajectories, of course), and (b) We note that, strictly speaking, we draw a comparison between one group observed pre-treatment (2014 buyers) and one group observed post-treatment (2012 buyers). However, since the health effects of Arsenic are long-term, one would not expect a strong treatment effect a mere two years after the test, even conditional on households effectively avoiding exposure.[13] We acknowledge that households that recall buying the test may differ in wealth from those buyers that did not purchase the test.

As is evident, across a range of dimensions, second-round buyers are somewhat worse-off than their counterparts who bought the tests in first round. Difference in ownership of durables such as TV and other white goods are significant. Difference in house type, latrine ownership and other assets such as motor-cycle also suggest the same, though the difference is not statistically significant. In addition, in line with our observation on assets, we find that second round buyers have significantly less education than first round buyers. We also find some differences in caste composition. Proportion of upper caste families, the second year buyer sample, is less than the same in first round, though middle and lower level caste households buy it more in the second round. This suggests a shift from upper to middle to lower caste households.

Overall, the comparison of household characteristics of first and second round buyers indicate that second round buyers are not as well off as first round buyers. It therefore appears unlikely that a wealth shock might have caused significant demand at the time of the repeat offer. This prompts the question

[13] We cannot exclude possible social effects of learning one's Arsenic status at the time of the first offer. However, any adverse social effects would lead us to underestimate the wealth of the 2012 buyers, and hence, to downward-bias our results.

whether learning, over the period of two years, might have led the marginal households to decide to purchase the test in 2014.

(ii) Learning

In the endline survey, about 80% households report to have heard about arsenic issue. We test for evidence on increased learning after first phase in the following way. First, (a) we consider the correlation between village-level demand in first round and purchases in second round. Secondly, (b) we consider differential learning from exogenous variation in *results* of tests conducted at baseline. Note that learning is not that straightforward in this study, since the degree of learning can also be affected by the distribution of arsenic incidence within a village.

(a) Since price was randomized at the village level, variation in village baseline demand is exogenous; by way of contrast, variations at the neighborhood or network level are prone to the well-known reflection problem. Yet, there is a mechanical effect by which a very high proportion of households purchasing tests in any village in 2012 leaves little margin left for have additional purchases in 2014.

The empirical evidence appears to reflect both effects. Villages with a low to medium level of demand in the first round reflect a positive effect of baseline uptake on second round uptake. (At very high levels of baseline purchases, the direction of the relationship flips, with few households on already covered.) This may indicate that households in villages with higher initial adoption of the technology became more aware of the value of arsenic testing during the two year period. Since arsenic tests are not available for sale in outside market, it reassures against any concern about unaccounted purchase of tests through other sources.[14].

[14] On the other hand, unaccounted sales may be a concern in demand calculations in case of other preventive health care products – such as mosquito nets – which are readily available in outside market.

(b) Secondly, we show that second-phase purchases react to the share of unsafe levels of arsenic wells in the first phase. As Table 8 shows, we observe an effect across a range of model specifications. Sample size is marginal in this village-level regression, and results are not significant with cluster bootstrap standard errors. One might expect such a differential effect of 'bad news' for other households in the village if it corrects initial subjective beliefs on the probability that an untested well is safe.

While this study does not distinguish between specific types of learning, one could suggest that, since arsenic tests are distinctly a non-experience good, learning might be chiefly driven by increased awareness of the probability of arsenic contamination, and the risks of exposure. Households may learn and become more aware about arsenic by observing other individuals in the village. While individual signals received from first round buyer or non-buyer households in the network may be noisy, the revealed evidence on additional purchases made by households at lower asset levels also indicate a possibility of positive learning. It should be noted that this model is different from learning from successful outcomes since time horizon to have any detectable impact of arsenic is much larger.

4.3 Well switching: Behavioral response to the information:

On average, we observe about 31% of Households reported that they have switched to safer water sources after purchasing the test in 2012. Other studies have found switching in the range of 30 to 70% in Bangladesh. One potential reason for non-switching could be limited number of identified safe wells because of low take-up, though it is also likely that an arsenic high well users may switch to an untested well.

Does higher price simulate more switching?

It is important to understand how people use the test results and whether price of the test affects their response. Our estimates suggest that switching rates are largely constant, irrespective of the price paid to purchase the test.   This is counter-intuitive since most of the current literature in preventive health care products have focused on screening and sunk cost effects. Both effects are in the same direction and will increase usage. In our case, we do not find any increase in usage by the households who purchase the test at higher price.

4.4.Hiding the bad news

We find evidence that households prefer to not report the bad news regarding the arsenic contamination in their wells during follow up, and they also take proactive action to remove displayed test results...

Table 9 shows the marginal difference in “recalled” arsenic status with “recorded” arsenic status, by comparing proportion of each well type (Red, Green and Blue) in the test outcome (2012) and the follow up on test outcomes (2014). We have three different measures of “recalled” arsenic status – (1) Placard still fixed on the well, (2) Placard removed from the well, but still kept in the house, (2) Placard neither in house nor kept aside, rather household is able to recall the “arsenic contamination status”. Consistently, the estimated proportion of recall in red wells is more than 9 percent points (pp) lower than the expected proportion based on recall. The magnitude is significant since about 50% of the households intentionally want to hide the outcome that their well has high arsenic. In case of green wells, we see an increase, and it suggests that households prefer to claim a medium arsenic level in their highly contaminated wells.

These findings are consistent with the theoretical and experimental studies on “self-serving bias” and “over-confidence”. Eli and Rao (2011) have found that negative feedback on personal intelligence and beauty were not used by people to update one’s prior, as one would expect from the Bayesian updating. ‘Good news’, on the other hand, is well received and inference conforms closely to Bayes’ rule. Ours is a unique context where environmental disasters affect private values and we show that the people try to intentionally ignore and hide bad news. Further, attached stigma of having a contaminated well and a high cost of switching, particularly if there are restriction on sharing water based on affiliation to social proximity and religion, can further increase the tendency to hide positive arsenic diagnosis. [15] Underscoring the role of social proximity, our survey shows that more than 90% households report that they prefer to exchange water within their own caste or relation.

Policy Discussion and Conclusion

This study measures willingness to pay for a water quality diagnostic product. Demand is highly sensitive to prices, as we also know from other studies in preventive health care research that cost-sharing reduces take-up significantly. The study measures how cost-sharing would affect demand. Since the demand is greatly affected by the extent of cost-sharing, the role of subsidy remains critical in ensuring maximum coverage.

Further, we find that switching rate is not affected by the price paid to buy the test. However, overall switching rates are relatively low. Increasing awareness about adverse effects of arsenic exposure can definitely increase switching rates (George et al., 2012), though the time period to experience the effect of switching to safe water source is typically much longer which restricts the ability to learn by doing. At the same time, when travel cost is high or there are barriers to access safe water wells, a lower switching rate would be observed.

Our empirical evidence on households' preference to hide the outcome of arsenic testing underscore further need for research. There may be restrictions put by socio-economic structure with in the village and lack of awareness feeding to stigma attached to having a contaminated well. This is highly policy

[15] In an adjacent state, caste is particularly found to be one of the major factor in impending water trade within a village (Anderson 2011).

relevant, particularly when ex-ante decision to purchase a test is also affected by any motivation to avoid bad news.

References

Ahmed, M. F., Ahuja, S., Alauddin, M., Hug, S. J., Lloyd, J. R., Pfaff, A., ... & Van Geen, A. (2006). Ensuring safe drinking water in Bangladesh. *SCIENCE-NEW YORK THEN WASHINGTON-*, *314*(5806), 1687.

Anderson, S. (2011). Caste as an Impediment to Trade. *American Economic Journal: Applied Economics*, *3*(1), 239-263.

Argos, M., Kalra, T., Rathouz, P. J., Chen, Y., Pierce, B., Parvez, F., ... & Ahsan, H. (2010). Arsenic exposure from drinking water, and all-cause and chronic-disease mortalities in Bangladesh (HEALS): a prospective cohort study.*The Lancet*, *376*(9737), 252-258.

Ashraf, N., Berry, J., & Shapiro, J. M. (2010). Can Higher Prices Stimulate Product Use? Evidence from a Field Experiment in Zambia. *American Economic Review*, *100*(5), 2383-2413.

Bennear, L., Tarozzi, A., Pfaff, A., Balasubramanya, S., Ahmed, K. M., & Van Geen, A. (2013). Impact of a randomized controlled trial in arsenic risk communication on household water-source choices in Bangladesh. *Journal of environmental economics and management*, *65*(2), 225-240.

Carson, R. T., Koundouri, P., & Nauges, C. (2011). Arsenic mitigation in Bangladesh: A household labor market approach. *American Journal of Agricultural Economics*, *93*(2), 407-414.

Chen, Y., Graziano, J. H., Parvez, F., Liu, M., Slavkovich, V., Kalra, T., ... & Ahsan, H. (2011). Arsenic exposure from drinking water and mortality from cardiovascular disease in Bangladesh: prospective cohort study. *Bmj*, *342*, d2431.

Cohen, J., Dupas, P. (2010). Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment. *Quarterly Journal of Economics*.

Dupas, P. (2014a). Getting essential health products to their end users: Subsidize, but how much?. *Science*, *345*(6202), 1279-1281.

Dupas, P. (2014b). Short-Run Subsidies and Long-Run Adoption of New Health Products: Evidence From a Field Experiment. *Econometrica*, *82*(1), 197-228.

Eil, D., & Rao, J. M. (2011). The good news-bad news effect: asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics*, *3*(2), 114-138.

Fendorf, S., Michael, H. A., & van Geen, A. (2010). Spatial and temporal variations of groundwater arsenic in South and Southeast Asia. *Science*,*328*(5982), 1123-1127

George, C. M., Graziano, J. H., Mey, J. L., & van Geen, A. (2012). Impact on arsenic exposure of a growing proportion of untested wells in Bangladesh.*Environmental Health*, *11*(7).

George, C. M., Inauen, J., Rahman, S. M., & Zheng, Y. (2013). The effectiveness of educational interventions to enhance the adoption of fee-based arsenic testing in Bangladesh: a cluster randomized controlled trial. *The American journal of tropical medicine and hygiene*, *89*(1), 138-144.

Kremer, M., & Miguel, E. (2004). *The illusion of sustainability* (No. w10324). National Bureau of Economic Research.

Kremer, M., Leino, J., Miguel, E., & Zwane, A. P. (2007). Spring cleaning: a randomized evaluation of source water quality improvement. *Quarterly Journal of Economics*.

Madajewicz, M., Pfaff, A., Van Geen, A., Graziano, J., Hussein, I., Momotaj, H., ... & Ahsan, H. (2007). Can information alone change behavior? Response to arsenic contamination of groundwater in Bangladesh. *Journal of development Economics*, *84*(2), 731-754.

Meredith, J., Robinson, J., Walker, S., & Wydick, B. (2013). Keeping the Doctor Away: Experimental Evidence on Investment in Preventative Health Products.*Journal of Development Economics*, *105*, 196-210.

Miller, G., & Babiarz, K. S. (2013). *Pay-for-performance incentives in low-and middle-income country health programs* (No. w18932). National Bureau of Economic Research.

Opar, A., Pfaff, A., Seddique, A. A., Ahmed, K. M., Graziano, J. H., & Van Geen, A. (2007). Responses of 6500 households to arsenic mitigation in Araihazar, Bangladesh. *Health & place*, *13*(1), 164-172.

Parvez, F., Wasserman, G. A., Factor-Litvak, P., Liu, X., Slavkovich, V., Siddique, A. B., ... & Graziano, J. H. (2011). Arsenic exposure and motor function among children in Bangladesh. *Environmental health perspectives*,*119*(11), 1665.

Pfaff, A., Schoenfeld, Amy, Ahmed, Kazi Matin and Geen Alexander van (2015), Arsenic exposure reduction in Bangladesh limited by insufficient testing and awareness. Unpublished working paper.

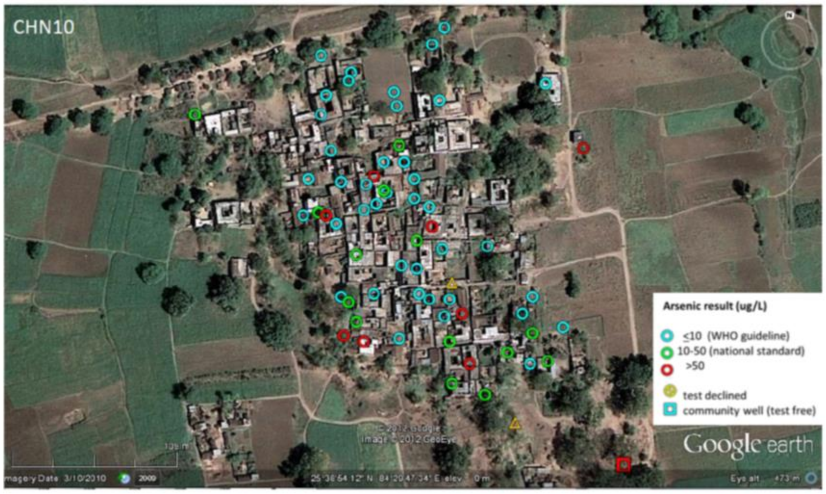
Pitt, M., Rosenzweig, M. R., & Hassan, N. (2012). Identifying the hidden costs of a public health success: arsenic well water contamination and productivity in Bangladesh. In *PSTC Working Paper Series 2012–02*.

van Geen, A., Ahmed, E. B., Pitcher, L., Mey, J. L., Ahsan, H., Graziano, J. H., & Ahmed, K. M. (2014). Comparison of two blanket surveys of arsenic in tubewells conducted 12years apart in a 25km 2 area of Bangladesh. *Science of The Total Environment*, *488*, 484-492.

Wasserman, G. A., Liu, X., Parvez, F., Ahsan, H., Factor-Litvak, P., van Geen, A., ... & Graziano, J. H. (2004). Water arsenic exposure and children's intellectual function in Araihazar, Bangladesh. *Environmental health perspectives*, 1329-1333.

Tables and Figures

Figure 1: Showing Arsenic Incidence in a village in Bhojpur district, Bihar (India)



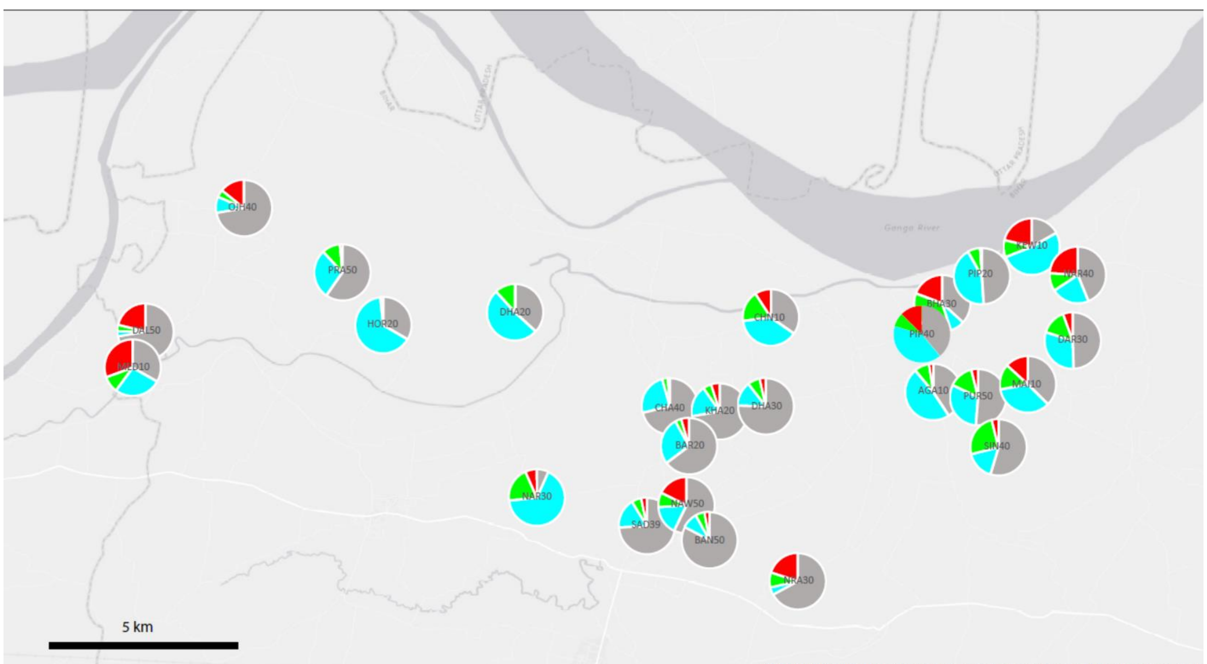
Note: A map sample village from the study is shown with the outcome of arsenic testing. Red circles denote high arsenic drinking water wells (> 50 microgram per liter). Green wells are relatively high but still can be used for drinking and cooking purpose, as per the national standard. Blue villages are low in arsenic and safe to drink water from.

Figure 2: Google maps from nearby villages were shown in focus group meetings

Note: Village level meetings and exhibition of posters showing safe and unsafe wells from near by villages. The geo-location of wells were jittered because of privacy concerns.

Figure 3: Metal Placard shoiwng arsenic stautus after testing

Figure 4: Map showing village locations with the arsenic test outcomes

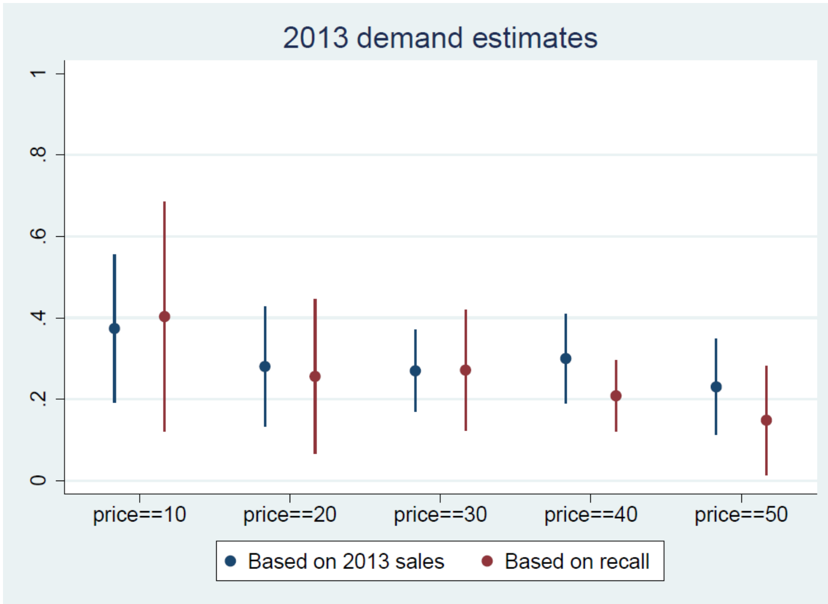


Note: This map shows the location of villages, take-up and outcome of the arsenic testing in subject area. Red (Arsenic high), Green (Arsenic medium) and Blue (Arsenic safe) colors show the outcome of arsenic testing. Grey color shows the proportion of untested wells.

Figure 5: Sales and Outcome of Arsenic tests

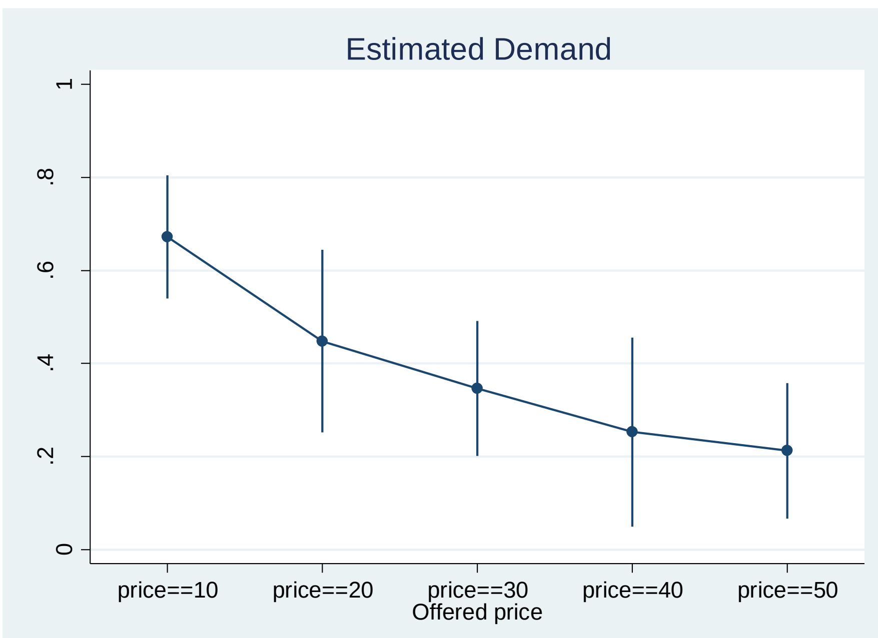
Note:. Recalled number of offers as well as purchases are less than the sample size recorded in endline survey (2014) and sales recorded in 2012 respectively, but such loss in recall is of similar magnitude (0.35 and 0.40, respectively). Red (Arsenic unsafe) and Green (Arsenic safe)) colors show the outcome of arsenic testing in the recorded data in 2012 and 2014. For "Recalled 2014", the colors show the households perception of their well quality in the endline survey. Yellow color represents the number of households who do not know the outcome of the previously purchased test

Figure 6: Comparison of demand estimate from first phase data and recall



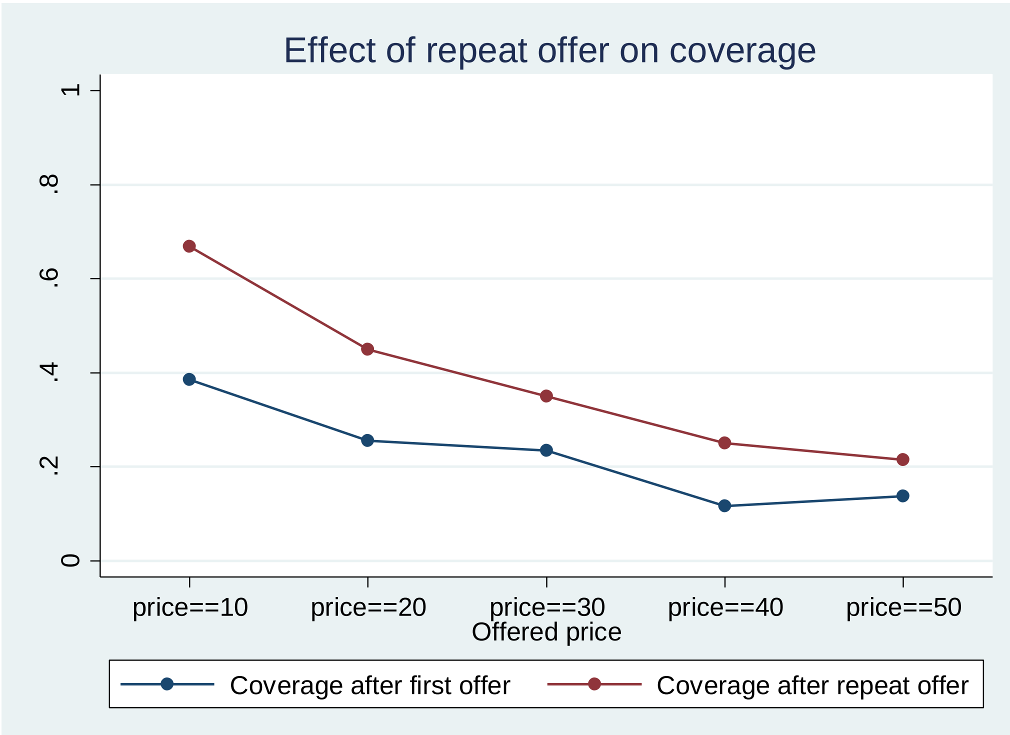
Note: Above plot shows the difference in demand estimates between recorded data in 2012-13 and recalled data in 2014-15. Except higher prices (Rs. 40 and Rs. 50) coefficients are of similar magnitude.

Figure 7: Estimated Demand



Note: Above demand curve includes sales in 2012 and 2014 and shows total coverage at the base of total number of households estimated in endline survey (2014). Recall data is used for 2012 sale and is scaled up considering a constant loss in recall.

Figure 8: Effect of repeat offer on total coverage



Note: This plot shows the comparison of demand pattern in first phase (2012) and second phase (2014). Percentage increase in demand is comparable at all price levels, but overall demand becomes steeper at low prices because of downward slope of the curve.

Table 1: Summary Table

Price group	(1)	(2)	(3)	(4)	(5)	ALL
has_car	0.038	0.027	0.026	0.011	0.039	0.029
	(0.007)	(0.005)	(0.006)	(0.004)	(0.009)	(0.003)
has_cell	0.855	0.867	0.908	0.917	0.912	0.885
	(0.012)	(0.011)	(0.011)	(0.012)	(0.013)	(0.005)
has_several_cells	0.080	0.189	0.024	0.171	0.223	0.133
	(0.009)	(0.012)	(0.006)	(0.016)	(0.019)	(0.006)
has_tv	0.198	0.228	0.219	0.190	0.158	0.204
	(0.014)	(0.013)	(0.016)	(0.017)	(0.017)	(0.007)
has_bike	0.722	0.694	0.675	0.585	0.664	0.676
	(0.015)	(0.015)	(0.018)	(0.021)	(0.022)	(0.008)
has_motorbike	0.214	0.162	0.216	0.244	0.278	0.213
	(0.014)	(0.012)	(0.016)	(0.018)	(0.021)	(0.007)
has_cow	0.638	0.632	0.763	0.639	0.677	0.665
	(0.016)	(0.015)	(0.017)	(0.020)	(0.022)	(0.008)
has_several_cows	0.171	0.339	0.284	0.276	0.377	0.283
	(0.013)	(0.015)	(0.018)	(0.019)	(0.022)	(0.008)
has_whitegoods	0.209	0.236	0.221	0.191	0.201	0.215
	(0.014)	(0.014)	(0.016)	(0.017)	(0.019)	(0.007)
pucca	0.795	0.568	0.757	0.653	0.764	0.700
	(0.014)	(0.016)	(0.015)	(0.020)	(0.019)	(0.007)
latrine	0.278	0.211	0.304	0.444	0.548	0.326
	(0.015)	(0.013)	(0.018)	(0.021)	(0.023)	(0.008)
num_hh_members	3.741	4.419	3.012	4.009	4.180	3.893
	(0.074)	(0.085)	(0.075)	(0.104)	(0.132)	(0.042)
infants	0.242	0.325	0.304	0.383	0.418	0.322
	(0.022)	(0.029)	(0.042)	(0.040)	(0.041)	(0.015)
children	0.492	0.730	0.358	0.555	0.649	0.564
	(0.029)	(0.034)	(0.027)	(0.043)	(0.051)	(0.016)
N	960	1105	815	653	551	4084

Note: Price group 1 to 5 denote Rs. 10 to Rs. 50 price

Table 2: Randomization balance among different price groups

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Households Members			House		Other Assets					
VARIABLES	Total Adults	Infants	Children	Concrete	Has_latrine	Car	Cell	TV	Bike	Motorbike	Cow
Price=Rs. 10	0.678	0.0830	0.238	-0.227	-0.0667	-0.0110	0.0124	0.0308	-0.0277	-0.0515	-0.00546
	(0.672)	(0.104)	(0.250)	(0.154)	(0.0855)	(0.0196)	(0.0735)	(0.105)	(0.0518)	(0.0578)	(0.0944)
Price=Rs. 20	-0.729	0.0618	-0.134	-0.0372	0.0257	-0.0127	0.0532	0.0214	-0.0469	0.00206	0.125
	(0.584)	(0.148)	(0.228)	(0.102)	(0.102)	(0.0166)	(0.0634)	(0.114)	(0.107)	(0.0367)	(0.0852)
Price=Rs. 40	0.268	0.141	0.0633	-0.142	0.166	-0.0276*	0.0623	-0.00814	-0.137	0.0297	0.00104
	(0.749)	(0.126)	(0.242)	(0.100)	(0.112)	(0.0146)	(0.0565)	(0.128)	(0.128)	(0.0321)	(0.0795)
Price=Rs. 50	0.439	0.176	0.157	-0.0304	0.270**	0.000127	0.0576	-0.0392	-0.0583	0.0644	0.0387
	(0.998)	(0.167)	(0.312)	(0.108)	(0.125)	(0.0223)	(0.0699)	(0.0766)	(0.0793)	(0.0574)	(0.0875)
N	3,526	3,528	3,522	3,758	3,528	3,527	3,528	3,528	3,528	3,528	3,527
R-squared	0.040	0.004	0.019	0.040	0.059	0.003	0.007	0.003	0.009	0.009	0.011
Mean at Price=Rs. 10	3.741	0.242	0.492	0.795	0.278	0.0384	0.855	0.198	0.722	0.214	0.638
Joint significance											
Wald chi2(df)	4.848	2.295	2.464	3.558	15.08	9.317	1.752	0.811	1.685	3.793	4.509
Prob > chi2	0.303	0.682	0.651	0.469	0.00455	0.0536	0.781	0.937	0.793	0.435	0.342

Note: Above plot shows difference in mean value of key demographic and assets variables across different price groups. Joint significance test result is reported in the bottom two rows. Cluster bootstrap standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Sales and Demand for arsenic tests

		First Phase (2012)				Second Phase (2014)		Demand			
		As recorded in 2012		As recalled in Endline survey (2014)		Endline survey (2014)		Demand in First Phase		Demand in Second phase	Total coverage
		(I)		(II)		(III)		(IV)			
Price level (Rs.)	Total well-owners visited in endline survey	Recorded sales	Recorded offers	Recalled sales	Recalled offers	Sales	Offered	Recorded in 2012	Recalled in Endline survey (2014)		
P	N	S12_a	n12_a	S12_b	n12_b	S14	n14	(S12_a)/N	(S12_b)/(n12_	S14/N	(S12_a+S14)/N
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
10	960	361	431	249	615	288	860	45%	40%	30%	68%
20	1105	310	423	206	804	183	985	38%	26%	17%	45%
30	815	218	352	125	460	117	662	43%	27%	14%	41%
40	653	196	327	92	441	86	554	50%	21%	13%	43%
50	551	127	289	52	350	45	467	52%	15%	8%	31%
All	4084	1212	1822	724	2670	719	3528	30%	27%	18%	47%

Note: This table summarizes the number of offers, sales and demand in both phases. Difference between demand calculated with recorded data (2012) and recalled data (2014) is shown in column 9 and 10. Note that N (column 2) denotes the total number of households as counted in the endline survey (2014).

Table 4: Estimated demand

	(1)	(2)
	Demand in 2014	Total Coverage (2012 and 2014)
Price=Rs. 20	-0.134*	-0.226**
	(0.0702)	(0.111)
Price=Rs. 30	-0.156*	-0.324***
	(0.0849)	(0.102)
Price=Rs. 40	-0.168**	-0.423***
	(0.0807)	(0.101)
Price=Rs. 50	-0.218***	-0.461***
	(0.0702)	(0.1000)
Observations	4,084	4,084
R-squared	0.037	0.110
Mean at Price=Rs.10	0.300	0.673

Note: Above table shows estimated additional demand in 2014 and total demand at the end of the two year. Recalled data is scaled up assuming constant recall loss in number of offers and purchases in order to estimate demand in 2012. Cluster bootstrap standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Selection at high price levels

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	House Type		Other Assets					
VARIABLES	Concrete	Has latrine	Car	Cell	TV	Bike	Motorbike	Cow
Panel A: Linear specification								
price	-0.00162	0.00747**	0.000154	0.00156	0.00180	-0.000673	0.00296***	9.18e-05
	(0.00294)	(0.00302)	(0.000374)	(0.00164)	(0.00324)	(0.00197)	(0.00100)	(0.00237)
Panel B: Breakdown by price levels								
20.price	-0.189	-0.0350	-0.00346	-0.0304	0.0459	-0.0366	0.0297	-0.0546
	(0.136)	(0.114)	(0.0164)	(0.120)	(0.135)	(0.0705)	(0.0742)	(0.0873)
30.price	-0.0367	0.0171	0.00884	0.0121	0.0394	0.0425	0.0279	0.0882
	(0.119)	(0.136)	(0.0184)	(0.0757)	(0.143)	(0.0778)	(0.0422)	(0.0805)
40.price	-0.173	0.254**	-0.0121	0.107***	0.0837	-0.0805	0.115***	-0.0501
	(0.118)	(0.116)	(0.0135)	(0.0407)	(0.183)	(0.146)	(0.0428)	(0.102)
50.price	0.0112	0.334**	0.0168	0.00559	0.0489	-0.0168	0.116***	-0.0221
	(0.0922)	(0.135)	(0.0235)	(0.0733)	(0.150)	(0.0824)	(0.0417)	(0.107)
mean at Price= Rs. 10	0.803	0.330	0.0267	0.886	0.223	0.789	0.221	0.685
Observations	1,301	1,366	1,365	1,366	1,366	1,366	1,366	1,365

Note: Above table shows correlation between higher wealth and purchase of test at higher prices. Cluster bootstrap standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Households characteristics in 2012 and 2014

	2015 buyers	2013 recall	2015 vs. 2013 recall		2015 buyers	2013 recall	2015 vs. 2013 recall
	(1)	(2)	(1) - (2)		(1)	(2)	(1) - (2)
Household characteristics				*Asset ownership*			
Number of HH members	4.919	4.311	0.608	HH has whitegoods	0.225	0.301	-0.0766*
	(0.367)	(0.325)	(0.382)		(0.0404)	(0.0563)	(0.0405)
Infant living in HH	0.302	0.223	0.0798**	Has cell phone	0.912	0.861	0.0507
	(0.0459)	(0.0246)	(0.0370)		(0.0230)	(0.0578)	(0.0460)
Child living in HH	0.488	0.438	0.0497	Has several cell phones	0.163	0.145	0.0175
	(0.0585)	(0.0618)	(0.0657)		(0.0532)	(0.0453)	(0.0469)
Several children living in HH	0.201	0.140	0.0612	Has TV	0.208	0.298	-0.0905**
	(0.0447)	(0.0437)	(0.0496)		(0.0372)	(0.0573)	(0.0424)
Education				Has bicycle	0.783	0.811	-0.0285
					(0.0187)	(0.0402)	(0.0382)
HH head has no formal education	0.134	0.100	0.0336*				
	(0.0225)	(0.0244)	(0.0194)	Has motorbike	0.248	0.261	-0.0131
					(0.0254)	(0.0243)	(0.0260)
HH head has primary education	0.610	0.632	-0.0225				
	(0.0691)	(0.0495)	(0.0637)	Has cow	0.680	0.680	6.24e-05
					(0.0417)	(0.0319)	(0.0353)
At least one HH member has more	0.429	0.445	-0.0152				
	(0.0923)	(0.0731)	(0.0738)	Has several cows	0.259	0.219	0.0398
					(0.0527)	(0.0475)	(0.0587)
Caste				*Housing characteristics*			
Scheduled caste or tribe	0.0163	0.0386	-0.0223	House pucca	0.701	0.756	-0.0553
	(0.00852)	(0.0240)	(0.0226)		(0.0556)	(0.0504)	(0.0391)
Other backward caste	0.227	0.127	0.0995**	Well depth (ft)	82.51	80.80	1.702
	(0.0518)	(0.0298)	(0.0411)		(3.692)	(3.922)	(2.806)
Kshatriya	0.0767	0.124	-0.0473	Has latrine	0.330	0.408	-0.0778
	(0.0309)	(0.0455)	(0.0371)		(0.0551)	(0.0496)	(0.0553)
Brahmin	0.251	0.388	-0.137***				
	(0.0658)	(0.0646)	(0.0510)				
Baniya	0.297	0.203	0.0940*				
	(0.0670)	(0.0446)	(0.0537)				

Note: This table compares socio-economic characteristics of early and late buyers. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Effect of price paid on the behavioral response to information

	(1)	(2)
Switched	from Red to Green	from Red to Blue or Green
20.price	0.227	0.227
	(0.146)	(0.249)
30.price	0.00227	0.00227
	(0.0930)	(0.211)
40.price	0.0292	0.0292
	(0.0856)	(0.240)
50.price	0.0773	0.0773
	(0.0930)	(0.104)
Observations	211	211
R-squared	0.014	0.014

Note: This table shows that switching to safe well is not affected by the price paid for test. Clustered bootstrap standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Outcome of test in first phase and the demand in second phase

	Demand in Second Phase				
Share of wells in village tested "red" in first round	0.0384	0.0699	0.0437	0.0933	0.117
	(0.112)	(0.125)	(0.107)	(0.114)	(0.130)
	[0.0301]	[0.0384]	[0.0301]	[0.0326]	[0.0404]
Observations	4,084	3,002	4,084	4,084	3,002
R-squared	0.037	0.060	0.051	0.059	0.082
Price	Yes	Yes	Yes	Yes	Yes
First-round demand	No	No	Linear	Quadratic	Quadratic
Wealth proxies	No	Yes	No	No	Yes
Note: clustered bootstrap standard errors in parentheses, classical standard errors in square brackets.					

Note: This table summarizes the effect of arsenic test outcome in the first phase on the demand in second phase. Cluster bootstrap standard errors in parentheses, classical standard errors in square brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table 9: Selective retaining of the test outcome

Placard color	Fixed on Well			Kept in House			Recall of Placard Color			All three combined		
	Red	Green	Blue	Red	Green	Blue	Red	Green	Blue	Red	Green	Blue
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Dummy for second phase	-0.0942***	0.0584	0.0358	-0.0925**	0.155***	-0.0621	-0.116***	0.0555*	0.0601	-0.0955***	0.118***	-0.0221
	(0.0262)	(0.0419)	(0.0374)	(0.0422)	(0.0513)	(0.0744)	(0.0253)	(0.0312)	(0.0435)	(0.0233)	(0.0317)	(0.0403)
Observations	1,529	1,529	1,529	1,379	1,379	1,379	1,762	1,762	1,762	1,840	1,840	1,840
R-squared	0.010	0.004	0.001	0.006	0.016	0.002	0.020	0.004	0.003	0.014	0.018	0.000

Note: Above table compares the proportion of Red, Green and Blue wells in the recalled and recorded data from phase 1. In phase 1, a placard was fixed to the well after arsenic test, but not all households keep it attached to the well. We also collect evidence if households keep the placard in the house. Overall, the table shows that households proactively take steps to hide the bad news of 'red' (arsenic high) well and prefer to substitute to 'green' (arsenic medium) status. Cluster bootstrap standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.