The Impact of Gavi on Vaccination Rates: Regression Discontinuity Evidence

Sarah Dykstra, Amanda Glassman, Charles Kenny, and Justin Sandefur

Abstract

Since 2001, an aid consortium known as Gavi has accounted for over half of vaccination expenditure in the 75 eligible countries with an initial per capita GNI below $1,000. Regression discontinuity (RD) estimates show aid significantly displaced other immunization efforts and failed to increase vaccination rates for diseases covered by cheap, existing vaccines. For some newer and more expensive vaccines, i.e., Hib and rotavirus, we found large effects on vaccination and limited fungibility, though statistical significance is not robust. These RD estimates apply to middle-income countries near Gavi's eligibility threshold, and cannot rule out differential effects for the poorest countries.

JEL Codes: F35, H51, I15, O11

Keywords: aid, vaccination, immunization, fungibility, regression discontinuity.

Center for Global Development
www.cgdev.org

Working Paper 394
February 2015

The Impact of Gavi on Vaccination Rates: Regression Discontinuity Evidence

Sarah Dykstra
Center for Global Development

Amanda Glassman
Director of Global Health Policy, Center for Global Development

Charles Kenny
Center for Global Development

Justin Sandefur
Center for Global Development

All authors are affiliated with the Center for Global Development (CGD) in Washington, D.C. We are grateful for helpful comments from Elisabetta De Cao, Victoria Fan, Willa Friedman, Owen Ozier, Heather Schofield, and numerous seminar participants. We appreciate comments and clarification offered by individuals from Gavi, the Vaccine Alliance, and the Gates Foundation. CGD does not take institutional positions, and the views expressed here do not necessarily reflect those of CGD's board or its funders. A web appendix containing additional information on data construction and sample countries can be accessed at www.cgdev.org/vaccinesrdappendix.

CGD is grateful for contributions from its funders in support of this work.

Sarah Dykstra, Amanda Glassman, Charles Kenny, and Justin Sandefur. 2015. "The Impact of Gavi on Vaccination Rates: Regression Discontinuity Evidence." CGD Working Paper 394. Washington, DC: Center for Global Development. http://www.cgdev.org/publication/impact-gavi-vaccination-rates-regression-discontinuity-evidence-working-paper-394

Center for Global Development
2055 L Street NW
Fifth Floor
Washington, DC 20036

202.416.4000
(f) 202.416.4050

www.cgdev.org

1 Introduction

Beginning in 2000, a wide array of public and private aid donors pooled funds to increase immunization rates in poor countries through the primary vehicle of bulk purchase of vaccines. This partnership, Gavi, the Vaccine Alliance, has cumulatively disbursed $7 billion worth of vaccines and financing to over seventy developing countries with a combined population of 4.5 billion. In January 2015, donors pledged a new round of $7.5 billion for Gavi's operations over the next five years.

As of 2000, the 75 countries that originally qualified for Gavi assistance had already made considerable progress in rolling out some relatively inexpensive vaccines, vaccinating over 60% of infants against diphtheria, pertussis, and tetanus (DPT) and 20% against hepatitis B (see Figure 1). These are cheap vaccines, costing about $0.10 and $0.30 per child in 2001, respectively. Meanwhile, as of the mid-2000s, an estimated 634,000 children in Africa and Southeast Asia died annually of pneumococcal disease, 433,000 of rotavirus, and 278,000 of Haemophilus influenza type B (Watt et al., 2009; WHO/UNICEF, 2012a; O'Brien et al., 2009). These diseases are preventable with fairly expensive vaccines, with costs ranging from $3 to $7 per child. Those vaccines either did not exist or had not been introduced in almost any developing country as of 2000. Gavi purchased both old, cheaper vaccines and the newer more expensive ones for recipient countries.

This paper attempts to measure the impact of Gavi's distribution of free and highly subsidized vaccines on vaccination rates, i.e., testing whether vaccines delivered have translated into more children vaccinated. Our identification strategy is based on a regression discontinuity at the cut-off line. The Alliance employs an income-based eligibility criteria for countries to receive funding for vaccines, which initially excluded all countries with a per capita gross national income (GNI) in excess of $1,000 in 1998 at market exchange rates. We show that this threshold has been fairly strictly enforced, providing a clean natural experiment in cross-country time-series data.

To address concerns with statistical power, we go beyond the textbook regression discontinuity specification to explore whether stronger modeling assumptions can yield more precise estimates. We estimate a dose-response model, replacing the binary Gavi treatment variable with a continuous measure of aid received per infant. To estimate the model, we require an additional instrumental variable to explain variation in doses. We exploit the fact that larger countries receive far less Gavi funding per capita, implying a larger jump in funding at the income eligibility threshold.

In our benchmark specification, for diseases where a relatively cheap vaccine already existed circa 2000, we find Gavi's provision of free or subsidized vaccines had little or no impact on vaccination rates. The intent-to-treat (ITT) effect of Gavi eligibility on hepatitis B and diphtheria, pertussis, and tetanus (DPT) vaccination rates for countries near the income-eligibility threshold is almost precisely zero. The same is true for measles, which received limited Gavi support, and which we use as a placebo test.

While we cannot reject the null hypothesis of zero impact of Gavi on hepatitis B or DPT coverage, we show the data can reject relevant alternative effect sizes. Specifically, we test and significantly reject the alternative null hypothesis of zero waste and zero fungibility – i.e., that all vaccines delivered by Gavi produced an equiproportional increase in vaccination rates. Given the Alliance often provides close to 100% of the vaccines required to immunize all children and that coverage rates for some of the vaccines Gavi supports were already at high levels in some eligible countries in 2000, this result should be expected. Indeed, the Gavi model 'forced' fungibility on some countries. If they had not reduced their own purchases from 2000 levels, they would have had more doses than children to vaccinate.

In contrast, for diseases covered by newer, less widespread, and more expensive vaccines, there are some signs that Gavi aid had a positive impact, though results are far from conclusive. In the benchmark RD specification we find small and statistically insignificant effects for the three high-priced vaccines promoted by Gavi: Haemophilus influenza type B (Hib); pneumococcal disease; and rotavirus. Using the dose-response model, we find significant positive effects in some specifications, and point estimates across various specifications that are consistent with large impacts and zero fungibility or waste. For the older, cheaper vaccines like hepatitis B and DPT, impact estimates remain very small and statistically insignificant with the dose-response model.

We caveat all our results by noting that we cannot say anything about the impact of Gavi on countries further from the income eligibility cut-off. There has been considerable convergence in vaccine coverage rates between countries above and below the cut-off since 2000, and the average rate of vaccination across all countries has been rising. Hepatitis B, Hib, and DPT vaccination rates on average approach 90% in countries with an active Gavi program by the end of our sample period in 2013. Gavi in-kind support may well have played a large role in raising vaccination rates in the poorest countries above what they would have been absent the program.

A secondary caveat relates to the impact that Gavi may have had on prices of new vaccines in the market in general, that may have allowed countries near the cut-off to access

lower prices for new vaccines than would have been the case in Gavi's absence. Historical vaccine price data for countries in our sample is not available in the public domain.

Our contribution to the literature is threefold. First, we provide an independent measure of the causal impact of Gavi support. Previous evaluations of this $7 billion program have lacked credible counterfactuals. Gavi itself estimates that since 2000, the Alliance has "helped to immunise an additional 440 million children which will save 6 million lives" (Gavi, 2014b). If one interprets that statement as a causal attribution to Gavi aid, vaccination rates would have remained constant since 2000 – an assumption our RD allows us to test and reject for countries just outside the eligibility threshold. A more sophisticated set of independent, academic analyses relies on dynamic panel-data GMM estimators proposed by Blundell and Bond (1998) to estimate Gavi's impact. Using this approach, Lu et al. (2006) find evidence of a positive, significant impact of Gavi funding on DPT vaccination rates for countries with initial coverage rates under 65%. In contrast, Hulls et al. (2010) extend this model to include additional years of data and find no such effect below 65% baseline DPT coverage, but a significant impact for countries between 65% and 80%. These findings are subject to the standard concerns about casual inference using off-the-shelf GMM panel-data estimators (Roodman, 2009). In addition, the justification for partitioning the sample using baseline DPT coverage is not clear. The partitioning highlights and may exacerbate the challenge of causal inference: because almost all countries with very low baseline DPT rates were within Gavi's income eligibility threshold, these models are estimating Gavi's impact by comparing aid recipients (compliers) to countries who failed to negotiate a Gavi aid package for a variety of reasons (non-compliers). Almost all variation in treatment status is driven by unobserved characteristics of recipients.[1]

Our RD estimates provide a more credible measure of Gavi's impacts, relying on exogenous features of the aid program for identification. To repeat an important caveat, the local average treatment effect we estimate is specific to the immediate vicinity of the Gavi eligibility cut-off, i.e., for the wealthiest Gavi recipients. There is *a priori* justification to believe the impact of subsidized vaccines will be smallest for these recipients, both because baseline immunization rates are higher, and because these middle-income countries have the greatest capacity to purchase new vaccines as they emerge. However, focusing on countries near

[1]Note that Lu et al. (2006) and Hulls et al. (2010) also focus on just one of several vaccines funded by Gavi. We expand our focus beyond DPT to include hepatitis B, pneumococcal disease, rotavirus, and Hib, as well as impacts on infant and child mortality. The pneumococcal and rotavirus vaccines were only widely introduced in the late 2000s, and we take advantage of more recent data to estimate Gavi's impact on their diffusion.

the eligibility threshold remains of considerable policy relevance as Gavi seeks a new round of financing from major donors, where negotiations have focused on revising the eligibility rules, particularly 'graduation' rules for rapidly growing economies.

Second, a recurrent debate around aid impact involves fungibility. Donors may be attracted to aid for vaccines not only by their apparent low cost and high efficacy, but by the perception that in-kind distribution mitigates fungibility concerns. Previous work has found significant fungibility of aid generically (Devarajan et al., 1999; Feyzioglu et al., 1998; Pack and Pack, 1993),[2] although with little attention to the endogeneity of aid. Recent findings of high levels of fungibility for health aid specifically (Lu et al., 2010) have been criticized both on identification grounds (Roodman, 2012) and misspecification due to the conflation of on-budget and off-budget aid (de Sijpe, 2013). Our regression discontinuity framework provides a cleaner empirical test of fungibility, and focuses squarely on the type of in-kind, off-budget aid designed to overcome it.

Third, we address a gap in the large cross-country literature on aid effectiveness, which has been overwhelmingly concerned with aid's effects on economic growth, producing mixed results (Burnside and Dollar, 2000; Easterly et al., 2004; Rajan and Subramanian, 2008; Clemens et al., 2012).[3] This growth focus fits awkwardly with the stated goals of health aid, whose growth effects may be limited (Acemoglu and Johnson, 2007; Clemens et al., 2012). Many prominent commentators in the aid debate, including both advocates and skeptics of international aid as a whole, appear optimistic about the effectiveness of health aid (Deaton, 2013; Easterly, 2014). But while the empirical support for this optimism is evident in impact evaluation of specific aid-funded programs, the relationship between large aid programs operating in multiple countries and health impact is not well-established in the literature. A limited number of studies find salutary effects of health aid on child mortality (Arndt et al., 2014; Mishra and Newhouse, 2009), but again causal inference in these dynamic panel regressions is open to debate (Roodman, 2009). This paper seeks to provide additional evidence on the link between aid and health outputs and outcomes.

[2] Pack and Pack (1990) is an exception to the general pattern of results, finding no evidence of fungibility in sector-specific aid to Indonesia. Van de Walle and Mu (2007) is an exception to the otherwise limited attention to causal identification in this literature, finding fungibility within (but not across) broad sector categories in the context of a field experiment.

[3] It is interesting to note that one of the most recent notable contributions to the aid and growth literature is the most methodologically similar to our analysis of vaccines, and finds quite large, positive, growth effects from aid. Galiani et al. (2014) construct a regression discontinuity analysis by exploiting the abrupt income eligibility threshold for International Development Assistance (IDA), the World Bank's most concessionary loan window. Knack et al. (2014) employ the same approach and find that crossing the eligibility threshold results in a significant decrease in aid from bilateral donors, as well as the World Bank.

The remainder of the paper is organized as follows. Section 2 describes the structure of the Gavi aid program and documents our data sources. Section 3 describes our econometric framework to measure treatment effects as well as fungibility and waste. Section 4 describes the results. Section 5 concludes.

2 Program background and data sources

2.1 Eligibility

In Gavi's first board meeting, attendees agreed on a threshold for support. "Countries with small resources and a lack of purchasing power have been considered to be in greatest need of financial support for the new vaccines. It is proposed that this be initially interpreted as those with a GNP/capita equal to or less than 1,000 USD" (Gavi, 1999). No detailed justification was offered for this threshold, other than that it was "IDA-like", i.e., similar to the cut-off for concessional World Bank loans explored by Galiani et al. (2014), an issue we return to below. Since then, Gavi's own financial sustainability models provide some ex post rationalization for the cut-off based on countries' ability to pay for vaccines.

This eligibility rule implies that the appropriate forcing variable in our analysis is the World Bank's measure of gross national income (GNI) per capita (Atlas method, current US$). In phase 1, the $1,000 GNI per capita threshold was measured using 1998 data. Seventy-five countries fell below this cut-off and were eligible to apply for support throughout the first phase.[4] In phase 2, the threshold was maintained at $1,000 based on 2003 GNI data and the list of eligible countries was updated (Gavi, 2009).[5] Beginning in 2011, the threshold was increased to $1,500 and adjusted annually for inflation in subsequent years (Gavi, 2014a). The cut-off was increased to this level because it is roughly equivalent to the

[4]Our sample for phase 1 contains 71 eligible countries. Bolivia had a GNI of $1,010 in 1998 and was not eligible until 2002, when its GNI dropped to $990 per capita. Timor-Leste did not become a country until 2002, at which point it became eligible. No GNI data were available Cuba and Djibouti in 1998, lower-middle income countries.

[5]Gavi's program dates can be confusing. We define phases in terms of eligibility criteria, which differs slightly from Gavi's definition of phases. Gavi's phase 1 operations extended from 2000 until 2005 and phase 2 programming began in 2006, the year in which eligibility was updated based on 2003 GNI data (see Azerbaijan and the Democratic Republic of Congo's applications for funding in 2005). Gavi's website lists phase 2 as beginning in 2007 because the strategic goals governing phase 2 were not approved until that year. Phase 3 began in 2011 for both programming and strategic goals, but because the eligibility threshold was updated annually at this point, we define 2012 as phase 4 and 2013 as phase 5.

original threshold in 2000, adjusting for inflation, and because in a constrained economic climate, it made sense for Gavi to retain fewer eligible countries (Gavi, 2009). The number of eligible countries dropped from 72 in phase 2 to 55 in 2011.

Because some GNI estimates found in the World Bank's databank have been revised in the time since they were used to determine Gavi eligibility, we draw GNI data from original versions of the World Bank's World Development Indicator publications. Section A.7 in the web appendix reports GNI data and vaccination rates for the countries in our sample by phase. Summary statistics appear in Table 1.

As we show below, Gavi remained fairly faithful to its initial, somewhat arbitrary income-eligibility rule. Unfortunately, from a research perspective, it has largely ignored any changes in country eligibility due to changes in the income threshold or economic growth. Once a country has qualified for Gavi aid, subsequent loss of eligibility is unlikely to put an end to aid flows, as we show in Section A.2 and Table 2 in the appendix.[6]

In addition to the income threshold, Gavi instituted coverage thresholds which determined eligibility for participation in the new and under-used vaccines support (NVS) and immunization services support (ISS) program. In phase 1, countries with a DPT vaccination rate of less than 80% were eligible to apply to receive ISS support, which provided performance-based financing for increased DPT coverage. This threshold was eliminated in phase 2, allowing all Gavi-eligible countries to apply. Countries with DPT coverage rates greater than 50% were able to apply for NVS funding, support grants for vaccine purchases. This threshold was increased to 70% in phase 3. In practice, these non-income eligibility thresholds have very little predictive power as to which countries received Gavi assistance, as discussed in Section A.2 in the appendix.

2.2 Aid flows

Most Gavi funding goes to purchasing vaccines, about $4.5 billion of its total $6.1 billion in disbursements from 2000 to 2013. These vaccines can be roughly divided into two groups: older, cheaper vaccines which dominated Gavi's programming for the first several years of its existence, and newer, more expensive vaccines which have rolled out more recently. The

[6]Note that Gavi's graduation policy allows countries to continue to receive funding support for an agreed period of time even if they become ineligible. If a country applied for funding partway through a phase, approved funding would continue to be dispersed, regardless of eligibility during the subsequent phase.

first group includes the vaccines for diphtheria, pertussis, and tetanus (DPT), and hepatitis B, which Gavi has more recently replaced with a single pentavalent vaccine that covers these diseases, with the addition of Hib. The cost per dose for DPT and hepatitis B vaccinations is well under $1, while the pentavalent vaccines costs between $1.30 and $3.50 per dose. Spending on these vaccines has accounted for roughly 45% of Gavi's total budget. Gavi also spent $1.6 billion on broader health systems support, auto-disable syringes and immunization systems support (ISS), a performance-based program designed to incentivize increased DPT coverage.

The second group includes the vaccines for rotavirus and pneumococcal disease, both of which were non-existent in the developing world prior to Gavi roll-out in the late 2000s, and Hib, which was present in only a handful of developing countries at the time of Gavi's launch. These vaccines are more costly, at $3.50 to $5.00 per dose for rotavirus and $3.50 to $7.00 for the pneumococcal vaccine. These vaccines have absorbed roughly a fifth of Gavi's total budget, or $1.2 billion, concentrated in later years.

In recent years, Gavi constitutes the majority of foreign aid for vaccinations and spends more per capita on vaccines in eligible countries than countries spend themselves – by a 6:1 ratio in 2012. The WHO and UNICEF Joint Reporting Form has collected country reports of government and non-government spending on immunizations since 2006. On average from 2006 to 2010 (Gavi phase 2), countries that were eligible for Gavi aid spent about $0.16 per capita per year on immunization programs, and received about $0.36 (mostly in kind) per capita from other sources. By 2012 (Gavi phase 4) government spending had fallen to $0.09 while non-government spending had risen to about $0.45. (See Table 1 and Figure 3.) These numbers are roughly consistent with Gavi's own records, which show that eligible countries received, on average, $0.26 per capita from Gavi in 2006-10 and $0.60 in 2012.

Gavi does not report on the number of doses purchased by country from 2000 to 2012. Using data from Gavi and UNICEF, we construct an estimate of the number of Gavi-purchased doses of each vaccine. UNICEF, which procures vaccines on behalf of Gavi, reports the total number of doses of each vaccine purchased worldwide and the total amount spent on each vaccine, which we use to calculate the average cost per dose across all formulations of a vaccine.[7] While Gavi does not provide information on vaccine presentations purchased, they

[7]These estimates were compared to UNICEF's awarded prices per dose for all presentations of a vaccine, and our estimates are within the range of costs per dose paid by UNICEF, available here. As we use the average cost per dose, our dose estimates may be higher or lower than the actual number of doses purchased, depending on the formulations used within each country. It is also unclear whether disbursement amounts

do publish data on spending by vaccine. We use the average cost per dose to estimate how many doses of each vaccine Gavi purchased by country and year. When using this data to estimate vaccination rates had Gavi not purchased vaccines, we account for buffer stocks and wastage using vaccine-specific rates reported by countries in the Joint Reporting Form, substituting the mean wastage rate by vaccine when an estimate is not available.[8]

2.3 Vaccination outcomes

In the years since the launch of Gavi, there has been considerable progress in vaccination rates in countries that were below the income threshold. Figure 2 displays vaccination rates for Gavi-eligible and ineligible countries in 2000 and 2013. As can be seen, for DPT, hepatitis B and Hib there has been considerable convergence in coverage rates between the two groups, driven by very rapid growth of coverage in Gavi-eligible countries. The coverage gap between the two sets of countries has fallen from 20 percentage points to 10 points for DPT, 38 to 11 points for hepatitis B and 15 to 7 in the case of Hib, all as vaccination rates have increased in both sets of countries. The progress in vaccine roll-out in eligible countries may or may not reflect the impact of Gavi, of course. Progress before Gavi's launch, and progress in ineligible countries both suggest other factors have had a role to play. This paper will examine the relative importance of Gavi support in raising immunization rates for countries near the income threshold.

The immunization coverage indicator we use is the WHO/UNICEF Estimates of National Immunization Coverage (WUENIC) for 2000 to 2013. These coverage estimates are available for all countries on an annual basis and measure the proportion of infants surviving to their first birthday who have completed 6 vaccine series: DPT, hepatitis B, Hib, measles, pneumococcal and rotavirus.

WHO constructs estimates using two sources of coverage data, administrative and survey. Survey data is considered the most reliable measure of coverage, but is only available intermittently and for a limited set of countries (Lim et al., 2008). Administrative data are drawn

are limited to vaccine costs, or also contains associated costs, such as syringes. This calculation of aid flows measured in doses does not affect our main RD treatment effect estimates, but any overestimate (or underestimate) of the number of doses purchased might lead us to overstate (understate) the degree of fungibility and waste in Gavi's aid.

[8]Wastage rates are self-reported through the Joint Reporting Form. As with country estimates of immunization coverage, it is possible that these numbers are biased or inaccurate (Ronveaux et al., 2005). Bosch-Capblanch et al. (2009) find that a majority of countries were able to correctly calculate wastage in their second data quality audit, a standard WHO process used to verify vaccination data quality.

from reports of vaccinations performed by service providers, which are then aggregated to the national level and reported annually to the WHO through the Joint Reporting Form. Administrative coverage figures are subject to both numerator and denominator bias due to erroneous reporting of doses delivered to children older than 12 months, under-reporting of doses provided by private clinics and inaccurate target population size, especially in cases where population projections are based on outdated censuses (Burton et al., 2009).

To account for these issues, WHO and UNICEF review coverage estimates for each country and vaccine on an annual basis to identify and address situations where data may be compromised and are not accurate representations of true coverage (WHO/UNICEF, 2012b). Estimates from any recent household surveys (primarily DHS and MICS) are compared to country estimates.[9] WHO consults with local officials, considers potential biases in each source and attempts to construct a pattern over time which reconciles different estimates.

Administrative data has the potential to over-report coverage levels due to measurement error and incentives from performance-based payment systems (Sandefur and Glassman, 2014; Murray et al., 2003; Lessler et al., 2011). Lim et al. (2008) observe large discrepancies between survey estimates and official country estimates of coverage, particularly among countries receiving ISS funding from Gavi, where over-reporting increased significantly after the introduction of the program. They find that official estimates of DPT coverage show more rapid improvement than survey-based estimates. WHO estimates are a measurable improvement on country estimates and exhibit less over-reporting, but they are still higher than survey estimates on average (Lim et al., 2008).

Because participation in ISS is associated with inflated coverage rates, a potential consequence of using WHO estimates is that we detect an effect on DPT vaccination rates where no true effect exists. While over-reporting due to ISS payments is a concern, the results of Lim et al. (2008) suggest that any misreporting would bias our results in favor of finding a positive effect of Gavi funding on immunization rates. The incentive to misreport created by ISS did not extend to vaccines other than DPT.

Lastly, it is important to note that WHO vaccination rates refer to the target population, which for all of the vaccines examined here is defined as infants under the age of one year. A potential threat to our research design would arise if vaccines were routinely administered

[9]There is often a delay of several years before results from surveys like the DHS are available. If new data from surveys become available which improve upon previous years' estimates, earlier estimates are also corrected. Contextual information, such as stock-outs or other disruptions to service delivery, is also considered. (Mitchell et al., 2012)

to children above this age threshold, a phenomenon referred to as "catch-up vaccination." This would be especially problematic if vaccines delivered by Gavi led to increased catch-up vaccination in the treatment group, leading to an underestimate of aid's impact. Fortunately for the research design, there is strong evidence this is not the case.

While we cannot observe catch-up vaccination in our data, the phenomenon is well-documented in the literature. Clark and Sanderson (2009) draw on Demographic and Health Survey (DHS) data from 45 countries to measure the delays in completing DPT vaccination courses. Delays of several weeks are quite common in many countries, but these delays very rarely push the date of completion beyond the one-year mark. In the full sample of 45 countries, 50% of vaccinated children had received the final dose of DPT by 20 weeks of age, and 95% by 30 weeks.

3 Empirical strategy

The core premise of our empirical strategy is that the relationship between a country's income and its vaccination coverage is relatively smooth and continuous. Richer countries have higher vaccination rates, but on average, countries with similar income will have similar vaccination rates. In contrast, Gavi treats similar countries very differently: if they fall above the income threshold they typically receive nothing, and if they fall below the threshold they typically receive large quantities of vaccines.

As an illustrative example, consider two countries that fell just on either side of Gavi's eligibility threshold in its first funding cycle: Ukraine (eligible at $980 per capita GNI in 1998) and the Philippines (ineligible at $1,050). As one would expect, before Gavi launched they had very similar vaccination rates. In 2000, Ukraine vaccinated 4% of infants against hepatitis B, and the Philippines 7%. But they received very different aid allocations. Adhering to its sharp eligibility rule, in 2004 Gavi gave Ukraine enough hepatitis B doses to vaccinate 81% of its infants, and gave the Philippines nothing.

This starkly different treatment of otherwise similar countries provides a natural experiment that we use to test two null hypotheses. First, we test the null hypothesis of zero impact, i.e., that countries on either side of the threshold look identical despite one group being eligible for Gavi aid and the other not. Rejecting this null implies Gavi had a significant impact. For instance, by 2005, Ukraine was vaccinating 97% of infants against hepatitis

B, compared to just 49% in the Philippines. Ignoring problems of statistical inference to focus on identification for the moment, if this pair of countries was representative of our full sample, we would conclude that aid had a positive impact of 48% on hepatitis B coverage.

Second, we test an alternative null hypothesis which we refer to as zero waste and zero fungibility, i.e., if Ukraine received enough hepatitis B doses to vaccinate 81% of its infants, we should see Ukrainian vaccination rates that are 81 percentage points higher than in the Philippines. Otherwise, either some of Gavi's aid was wasted, or Ukraine reduced its domestic vaccination efforts in response to the aid package. In our example, Gavi's 'impact' on Ukraine was large at 48%, but far less than the amount of aid received. The discrepancy implies waste or displacement equivalent to 33% of infants (81% minus 48%).

Wasting vaccines is very different than crowding out domestic vaccine purchases, and we maintain a high burden of proof before claiming evidence of waste. In extreme cases, though not this example, aid deliveries exceeded the vaccination rate altogether. For instance, from 2001 to 2013, Gavi gave Bangladesh enough hepatitis B vaccine to vaccinate 70% of its infants each year, but the WHO reports only 59% of infants as fully vaccinated on average over that period. Only in these relatively rare instances do we infer actual waste of Gavi vaccines.

Stepping back from this numerical example, the following paragraphs formalize the rationale for using countries just above Gavi's income eligibility threshold as a control group for those just below it. We explain the logic of this regression discontinuity design in the standard potential outcomes framework of the treatment effects literature, following Lee and Lemieux (2009). As we move beyond this two-country example, we discuss the choice of optimal bandwidth – i.e., how many countries how far from the threshold to include in the analysis – and how we adjust for income differences between countries around the cut-off.

3.1 Potential outcomes framework

Suppose that for each observation i the outcome takes a value of $Y_i(1)$ if i is eligible for treatment and $Y_i(0)$ if it is not. Country i is eligible for aid if its income X_i is below the threshold c. For each observation we observe only one of these potential outcomes, and we can estimate $E[Y_i(0)|X > c]$ for the ineligible countries above the Gavi cut-off, and $E[Y_i(1)|X < c]$ for the eligible countries below the cut-off.

The key assumption underlying RD estimation is that potential outcomes are continuous in X, the forcing variable (Hahn et al., 2001). Under this assumption, the treatment effect at the cut-off becomes:

$$E[Y_i(1) - Y_i(0)|X = c] = \lim_{\substack{\varepsilon \to 0 \\ <}} E[Y_i(1)|X = c + \varepsilon] - \lim_{\substack{\varepsilon \to 0 \\ >}} E[Y_i(0)|X = c + \varepsilon] \tag{1}$$

This expression makes clear that RD estimates yield a *local* average treatment effect (LATE) in a very specific sense: the LATE applies to observations in the immediate vicinity of the cut-off, c, as ε approaches zero. This LATE is an intent-to-treat (ITT) effect, as we have defined treatment purely in terms of treatment eligibility. Below we consider the effect of treatment on the treated (TOT), i.e., the LATE applicable to countries near the threshold who actually received Gavi aid.

Now turn to the question of measuring fungibility and waste. Let D_i denote the total vaccine doses delivered by Gavi to country i, measured as a proportion of the target-age population the aid could cover in an ideal scenario. We estimate the number of Gavi-purchased doses using Gavi's disbursement data and the average cost paid per dose annually for each vaccine, calculated using UNICEF data on total doses purchased and total spending -worldwide. Gavi calculated its aid based on estimates of achievable coverage, often $D_i = 100\%$, especially for hepatitis B, Hib, and DPT. As our starting point, we assume that if vaccines are completely non-fungible and there is zero waste, then aid should raise vaccination rates by an amount equivalent to the total vaccines received. This implies that

$$E[\tilde{Y}_i(1) - \tilde{Y}_i(0)|X = c] = 0, \tag{2}$$

where $\tilde{Y}_i(1) \equiv Y_i(1) - D_i(1)$. In words, once we subtract Gavi vaccine doses from the vaccination rate, we should observe no difference between eligible and ineligible countries if aid is not fungible and is not wasted.

To make this concrete, return to our numerical example. Ukraine ($980 per capita GNI in 1998) and the Philippines ($1,050) fell just on either side of the eligibility threshold for Gavi's phase 1. Ukraine recorded a hepatitis B vaccination rate in 2005 of 97%, compared to 49% for the Philippines, while we estimate Ukraine received free vaccines equivalent to 81% coverage in the year before and the Philippines received none.[10] Equation (2) implies that

[10]Using the methodology described in Section 2.2, we estimate that Gavi delivered 1,030,401 doses of hep B to Ukraine in 2004, whereas Ukraine's 2004 Annual Progress Report states that 989,872 doses were delivered, a discrepancy of 4%.

if there is no fungibility or waste, we should expect Ukraine's vaccination rate to be equal to the Philippines' after subtracting the aid to Ukraine – 97% coverage minus 81% aid (i.e., 16%) should equal 49%. The gap between 16% and 49% implies that Gavi aid was either wasted or crowded out domestic expenditure in Ukraine.

3.2 Regression specification

Our basic empirical strategy is a straightforward application of regression discontinuity methods: we regress vaccination and mortality rates within a country on a function of the baseline per capita GNI variable used to determine Gavi eligibility, and test for a discontinuous jump in outcomes at the threshold.[11] We test for balance by showing that pre-treatment trends in vaccination exhibited no discontinuity at the threshold. We estimate our core regression parametrically using a linear regression with polynomial terms on either side of the cut-off, reporting both intent-to-treat (ITT) and average treatment on the treated (TOT) effects. For the latter, we use a fuzzy regression discontinuity design to account for the fact that many Gavi-eligible countries did not receive aid in a given year – though almost all eligible countries received aid in at least some years.

They key challenge in estimation is to provide a reasonable empirical approximation of the (continuous) function linking the forcing variable to the outcome. We follow the advice of Gelman and Imbens (2014) and focus on low-order polynomials on either side of the cut-off, varying the bandwidth of data used in estimation.

To measure the intent-to-treat effect, we regress change in vaccination and mortality rates (Y) in country c and year t on a p^{th}-order polynomial of log GNI per capita in year t, X_{ct}, an eligibility dummy indicating that the country fell below the Gavi eligibility threshold in the given year, E_{ct}, and the interaction of the treatment dummy and the polynomial terms.

$$\begin{aligned} Y_{ct} &= \beta_{00} + \beta_{01}\tilde{X}_{ct} + \beta_{02}\tilde{X}_{ct}^2 + \ldots + \beta_{0p}\tilde{X}_{ct}^p \\ &\quad + \rho E_{ct} + \beta_{11}E_{ct}\tilde{X}_{ct} + \beta_{12}E_{ct}\tilde{X}_{ct}^2 + \ldots + \beta_{1p}E_{ct}\tilde{X}_{ct}^p + \eta_{ct} \end{aligned} \tag{3}$$

In practice, we present linear and quadratic results, $\{p : 1, 2\}$. Note that the tilde over the GNI terms indicates that they have been de-meaned, such that $\tilde{X}_{ct} \equiv X_{ct} - \bar{X}_{ct}$, enabling us to interpret the ρ coefficient as an estimate of the treatment effect at the discontinuity.

[11] A short introduction to regression discontinuity can be found on BetterEvaluation's website. A more extensive guide to the study design can be found in MDRC's A Practical Guide to Regression Discontinuity.

Because estimates may be sensitive to the bandwidth chosen, we employ the method proposed by Imbens and Kalyanaraman (2012) to select the optimal bandwidth to minimize mean-squared error. The optimal bandwidth selection is implemented using the "rd" package in Stata, documented in Nichols (2007) and subsequent updates. Rather than allowing the bandwidth to vary as we estimate effects on, say, hepatitis B and DPT, we use a benchmark bandwidth that is the rough midpoint of the Imbens and Kalyanaraman (2012) estimates across the six vaccines studied. We also report results using half and twice this bandwidth.

In addition to the ITT effect, we are also interested in the effect treatment on the treated (TOT). This is a specific LATE estimate, relevant to eligible countries which may have endogenously selected into treatment and, in this instance, only such countries which are near the eligibility threshold. We can estimate the TOT effect in a two-stage least squares framework commonly referred to as a fuzzy regression discontinuity design.

In the second stage, we regress Y_{ct} on a polynomial of log GNI per capita in year t, X_{ct}, the predicted value of the treatment dummy indicating that the country received Gavi funding in the given year, T_{ct}, and the interaction of the predicted treatment dummy and the polynomial terms.

$$Y_{ct} = \beta_{00} + \beta_{01}\tilde{X}_{ct} + \beta_{02}\tilde{X}_{ct}^2 + \rho\widehat{T}_{ct} + \beta_{11}\widehat{T_{ct}\tilde{X}_{ct}} + \beta_{12}\widehat{T_{ct}\tilde{X}_{ct}^2} + \eta_{ct} \quad (4)$$

To test robustness, we estimate equation (4) using both a linear and quadratic function of X.

In the first stage, we regress the actual Gavi treatment dummy on an indicator of Gavi eligibility, which takes a value of one if the country's GNI was above the cut-off in that phase, E_{ct}. In addition, we control for the same polynomial function of log GNI as used in the second stage in equation 4, as well as the interaction of these polynomial terms and the eligibility dummy.

$$T_{ct} = \gamma_{00} + \gamma_{01}\tilde{X}_{ct} + \gamma_{02}\tilde{X}_{ct}^2 + \pi E_{ct} + \gamma_{11}E_{ct}\tilde{X}_{ct} + \gamma_{12}E_{ct}\tilde{X}_{ct}^2 + \varepsilon_{ct} \quad (5)$$

Note that we rely primarily on a binary indicator as the treatment variable, T, although dollar figures are available from Gavi, because our main instrumental variable (income eligibility) cannot explain any variation in the size of Gavi aid packages across eligible countries.

As a robustness check, we also estimate a specification that is analogous to a dose-response model, where T is defined as a continuous (nonnegative) variable measuring the number of

vaccine doses delivered to the country by Gavi per infant (under one year of age). The Gavi eligibility threshold by itself only predicts a uniform jump in aid for all countries. To be meaningful, this model requires additional instruments to explain variation in the 'dosage', i.e., the size of Gavi aid flows across countries.

There is a clear pattern that larger countries receive less Gavi aid per capita. Thus we propose to use the interaction of a country's infant population and the indicator for eligibility as an additional instrument, as smaller countries should see a bigger jump in aid at the income-eligibility threshold. We control separately for the direct effect of population size on vaccination in the second-stage regression, so that the second-stage equation becomes

$$\begin{aligned} Y_{ct} &= \beta_{00} + \beta_{01}\tilde{X}_{ct} + \beta_{02}\tilde{X}_{ct}^2 + \beta_{0p}P_{ct} \\ &\quad + \rho\widehat{T}_{ct} + \beta_{11}\widehat{T_{ct}\tilde{X}_{ct}} + \beta_{12}\widehat{T_{ct}\tilde{X}_{ct}^2} + \ldots + \beta_{1p}\widehat{T_{ct}\tilde{X}_{ct}^p} + \eta_{ct} \end{aligned} \tag{6}$$

and the first-stage regressions take the following form,

$$\begin{aligned} T_{ct} &= \gamma_{00} + \gamma_{01}\tilde{X}_{ct} + \gamma_{02}\tilde{X}_{ct}^2 + \gamma_{0p}P_{ct} \\ &\quad + \pi E_{ct} + \gamma_{11}E_{ct}\tilde{X}_{ct} + \gamma_{12}E_{ct}\tilde{X}_{ct}^2 + \gamma_{1p}E_{ct}P_{ct} + \varepsilon_{ct}, \end{aligned} \tag{7}$$

where P_{ct} measures the population under age one in country c. Again, the main innovation here is the addition of a second excluded instrument, $E_{ct}P_{ct}$, the interaction of population and eligibility. We achieve greater precision relative to the simple RD specification by making additional identifying assumptions. While we allow for the relationship between population size and vaccination to vary across the income spectrum, the crucial identifying assumption for the dose-response model is that this population-vaccination relationship would not jump discontinuously in the vicinity of the income eligibility threshold except due to the effect of Gavi.

4 Results

4.1 First stage results and balance tests

Before turning to the main results, we address two important preliminaries: (a) ensuring that Gavi aid exhibits a discontinuous jump at the eligibility threshold, and (b) demonstrating that other variables which were pre-determined at the time of Gavi's launch do not.

The first-stage regressions show that Gavi's sharp eligibility threshold has been enforced quite strictly. Results of estimating equation (5) are reported in Table 2 and Figure 4. The probability of receiving Gavi aid falls by about 7% for each additional log point of per capita GNI, and jumps more than 60% at the threshold, with minor variation across specifications. This jump is significant at the 1% level with linear and quadratic controls for log per capita GNI, and is fully robust to allowing the GNI function to vary on either side of the threshold. The jump is significantly less than 100% primarily because of non-compliance rather than defiance: after Gavi was launched in 2000, disbursements began only gradually in many countries, and many country-year cells are coded as zero, particularly in earlier years.

In theory, Gavi eligibility depended not only on the income threshold, but also on separate thresholds defined by baseline DPT vaccination rates, as described in Section 2. In practice, it appears these thresholds were not strictly enforced. A country's classification based on baseline DPT shows some relationship with the probability of receiving Gavi aid, but this relationship is not robust and only marginally significant in the best cases. We explore this issue in detail in the appendix (see Table 3). Because baseline DPT adds little or no predictive power to the first-stage regression, we focus exclusively on the income threshold.

The sharp jump in Gavi aid at the GNI threshold raises the question of whether countries differed previously at this point, and whether Gavi simply chose a pre-existing discontinuity in country performance. For the variables where we have data on pre-treatment trends – including vaccination rates for hepatitis B, Hib, DPT, and measles, as well under-5 and under-1 mortality rates – we fail to reject the null of parallel pre-treatment trends (see Table 3).

More formally, regression discontinuity designs are subject to balance tests, analogous to those reported for randomized experiments. The assumption that counterfactuals are a continuous function of the forcing variable implies, in our setup, that baseline values of the outcome variables (circa 2000) should exhibit no jump at the Gavi eligibility threshold. To test this, we apply equation 3 to baseline values, and report results in Table 3. As seen in the table, sample sizes are relatively small (roughly 10% the size of our main specifications for the treatment effects) and thus power is somewhat limited. Nevertheless, we fail to detect any significant trend break in the relationship between income and vaccination or mortality in 2000 at the cut-off.

4.2 Treatment effects on vaccination and other outcomes

Our treatment effect estimates are somewhat underwhelming. We find no significant impact of heavily subsidized vaccines delivered by Gavi on vaccination rates across any of the six diseases tested in our main specification. Alternative specifications focusing on later rounds of Gavi do, however, suggested positive impacts on newer, more expensive vaccines.

Intent-to-treat (ITT) results from the reduced form regression shown in equation (3) are reported in Table 4. Each column reports results for a different outcome indicator, and the three panels present results using various bandwidths (i.e., ranges of data as defined by the forcing variable). We treat the linear specification with a bandwidth of 1 log point per capita GNI as our preferred specification. Results from this specification are largely consistent with the wider bandwidth and quadratic specifications.

In the benchmark specification, point estimates imply Gavi eligibility led to an increase in vaccination coverage of only about 1 to 2%, with standard errors between 2% and 6%. The one exception among the six vaccines is Hib, with a point estimate of 10%, though this is also insignificant and not robust across specifications.[12] Controlling for baseline vaccination rates yields slightly higher point estimates, but a qualitatively similar overall picture in Table 5, with no statistically significant impacts in our preferred specification.

To appreciate the magnitude of these point estimates, it is useful to examine the treatment-on-the-treated (TOT) estimates. Two-stage least squares results based on the fuzzy regression discontinuity framework of equation (4) are reported in Table 7. The point estimates from the TOT estimates highlight that the confidence interval on Gavi's impact in countries where it is operational spans reasonably large magnitudes from a policy perspective. The point estimate for hepatitis B is just 0.9% in the linear IV specification (bw=1), but the 95% confidence interval includes positive effects upwards of 20%.

We noted earlier that Gavi's initial threshold was designed to be "IDA-like," in reference to the International Development Association, which provides poor countries with access to concessional loans and grants. The operational cut-off for IDA eligibility was $885 in 2000 and has remained slightly lower than Gavi's threshold throughout all phases. It is possible that our results are affected by this other, nearby aid threshold. Another factor which

[12] The quadratic specification with the narrowest bandwidth (0.5 log GNI) produces some anomalous results – in particular, a point estimate implying Gavi reduced Hib vaccination rates 28% – but the standard errors here are very large. It appears this specification involves more parameters than we can reliably estimate with such a thin slice of data around the threshold.

might influence our results is the Pan American Health Organization's (PAHO) Revolving Fund. Through the Fund, PAHO provides member countries, including both Gavi-eligible and -ineligible countries, with access to lower vaccine prices through a bulk purchasing mechanism. We control for both IDA eligibility and PAHO membership and observe no substantive difference in our estimates (results available upon request).

Consistent with these mixed results on vaccination, we find no impact on infant or child mortality. Intent-to-treat results are reported in Table 6. Point estimates vary widely across specifications and there is no evidence of a statistically significant impact on mortality rates.

This should not be surprising. Even optimistic magnitudes of mortality decline would be small and difficult to detect with our sample. As of 2000, about 44% of the roughly 9 million child deaths in developing countries were attributed to vaccine-preventable causes (Black et al., 2003). But diseases covered by Gavi vaccines constitute only 15% of those deaths (Watt et al., 2009; WHO/UNICEF, 2012a; O'Brien et al., 2009; Goldstein et al., 2005; Morris et al., 2008). Furthermore, given the limited roll-out of the newly-developed pneumococcal and rotavirus vaccines up to 2013, it is likely that any impact on child mortality observed during the 2000s would operate through the Hib vaccine, representing no more than 3% of total child deaths in 2000 (Watt et al., 2009). Null results for child mortality neglect Gavi's potential impact on adult morbidity and mortality. Gavi's own estimates predict reduced adult mortality primarily via increased hepatitis B vaccination rates, a channel where we find no impact.

Finding no evidence of an impact on vaccination rates is, of course, not the same as finding evidence of no impact. As noted in the introduction, one might worry that our non-results reflect a lack of power to detect important effect sizes. To address this concern, we pursue two extensions in the following sections: using an additional instrument in the context of a dose-response model to attain more precise treatment effect estimates, and testing whether our data, which fails to reject the null of zero impact, can reject two policy-relevant alternative hypotheses about Gavi's impact.

4.3 Dose-response model

As discussed in Section 3, Gavi allocates far more aid per child to small countries than big countries, creating potentially important heterogeneity in its impacts. Average estimates in

Tables 4 and 7 may understate the impact of the Gavi vaccine distribution in small countries, and acknowledging this heterogeneity may allow us to improve the precision of our results.

To explore this issue, we estimate the dose-response specification from equation (7) which includes the infant population—eligibility interaction term. Results are reported in Table 8. Focusing on our preferred specification (linear, bandwidth = 1), point estimates remain small and insignificant for both hepatitis B and DPT. There is some evidence of a significant impact on Hib coverage however, with a point estimate implying that Gavi delivery of doses sufficient to vaccinate 100% of infants led to an increase in vaccination coverage of 72%, significant at the 10% level. Notably, Hib is the one component of Gavi's preferred pentavalent vaccine which was not already widely covered by other, cheaper vaccines.

Results are also consistently positive and in some specifications – but not all – statistically significant for rotavirus, with the preferred specification showing a greater than one-for-one increase in vaccination rates in response to Gavi aid (a coefficient of 100 in Table 8). There is also some evidence of a statistically significant impact on Hib vaccination rates, but results vary widely across specifications. With a bandwidth of one log point, we cannot reject zero fungibility and waste, but can reject the null of zero impact in the quadratic specification. While these results are far from conclusive, they are suggestive of greater impacts among the more expensive, new generation vaccines relative to the zero impact estimates for DPT and hepatitis B.

Results for pneumococcal vaccine are inconclusive. Given the somewhat positive results for rotavirus and Hib, it is somewhat surprising to see small or negative and insignificant coefficients on aid for the pneumococcal vaccine. Results for measles vaccination rates, the placebo outcome not targeted by Gavi, are not precisely estimated and vary widely, but never differ significantly from zero.

4.4 Fungibility

For older, cheaper vaccines (DPT, hepatitis B), coverage rates around the Gavi threshold are not consistent with the hypothesis that vaccines provided by Gavi raised vaccination rates in equal proportion – that is, the data reject the null hypothesis of no fungibility and no waste.

As explained in Section 3, the regression discontinuity framework provides a way to measure the combination of fungibility and waste. Under the null hypothesis of no fungibility

and no waste, vaccination rates in eligible and non-eligible countries on either side of the eligibility threshold should be indistinguishable after subtracting the Gavi-purchased vaccines from the former, as shown in equation (2). More precisely, we define $Y^* \equiv$(Gavi doses delivered - wastage - buffer stock)/(doses per child $\times$ population under age 1). For ineligible countries, we directly observe their outcomes in the absence of treatment and can estimate $E[Y_i(0)|X > c]$, or more precisely, $\lim_{\varepsilon \underset{>}{\to} 0} E[Y_i(0)|X = c + \varepsilon]$, subject to some functional form assumptions. Then our measure of the counterfactual suggested by Gavi is:

$$Y_{it}(0) = Y_{it}^a = \begin{cases} Y_{it} - Y_{it}^* & \text{if } X_i \leq c; \\ Y_{it} & \text{if } X_{it} > c. \end{cases}$$

For instance, if country i received 3 million doses of hepatitis B vaccine from Gavi, and hepatitis B requires 3 doses per child, we record $Y^* = 3$ million$/3 = 1$ million. If the country's actual vaccination rate is 80% with an infant population of 1.5 million, we deduct the 1 million vaccinated children claimed by Gavi to produce an adjusted (counterfactual) vaccination rate of 33%.

In our preferred specification (linear, bandwidth of 1 log point) in Table 9, there is significant evidence of fungibility or waste for both hepatitis B and DPT. This result is particularly strong and robust for hepatitis B, where vaccination rates in eligible countries are 56% lower than they should be in the absence of fungibility or waste. For DPT, we find that vaccination rates are about 30% below where they would be without fungibility or waste, though this result disappears with the narrower 0.5 log point bandwidth.

Notably, we find no evidence of fungibility or waste for four of the six vaccines, including rotavirus and pneumococcal disease which were introduced later and more slowly, and measles – our 'placebo' outcome, which has received limited Gavi support.

Other counterfactuals are also testable. For instance, the data around the eligibility threshold are also inconsistent with the suggestion that all 440 million additional children vaccinated "with Gavi support" would have remained unvaccinated absent the program. The implicit null hypothesis in this claim is that vaccination rates would have remained at baseline levels without Gavi aid.[13] Recall that the core assumption of regression discontinuity analysis stated in Section 3 is that potential outcomes are a continuous function of the forcing variable. This implies that baseline vaccination rates in Gavi-eligible countries should be

[13]Baseline is defined as the vaccination rate the year before Gavi began funding a vaccine in a country.

equal to current vaccination rates in ineligible countries near the threshold – i.e., we should observe zero increase in vaccination in our control group.

This suggests the following regression. Adding time subscripts to our notation, we define the counterfactual in the data, Y_i^s as follows:

$$Y_{it}(0) = Y_{it}^s = \begin{cases} Y_{i,baseline} & \text{if } X_{it} \leq c; \\ Y_{it} & \text{if } X_{it} > c. \end{cases}$$

The key point is that Y^a and Y^s should be continuous functions of the forcing variable, X, with no jump at the Gavi threshold. If Y^a or Y^s is significantly lower on the eligible side of the cut-off, this implies that impact claims using these benchmarks are biased upward by the use of an incorrectly low counterfactual. The results in Table 4 in the appendix show significant evidence of an artificially low counterfactual for hepatitis B. Rather than remaining stable, our preferred specification suggests coverage rates rose by 42.8% just outside the eligibility threshold. Results for the other vaccines suggest smaller increases in the absence of Gavi, and for none of the other vaccines is the counterfactual trend robustly different from zero.

4.5 Waste

Up to now we have lumped together fungibility and waste, but they have very different implications from a public finance perspective. Fungibility may be optimal from a recipient country's perspective – as subsidized vaccines liberate domestic resources for other legitimate purposes – even if it frustrates the objectives of the aid donor. Waste, in which subsidized vaccines simply go unused, is obviously suboptimal for all parties. Waste is a particular concern with in-kind aid programs such as Gavi. Money doesn't spoil, but vaccines do.

While we cannot observe waste directly, in some extreme cases, we can infer it indirectly when aid donors provided more vaccines than the total number of children vaccinated in a country. Thus we propose a conservative measure of waste, defined as the excess of aid over and above the total vaccination rate in a given country year:

$$\text{Waste}_{it} \equiv \max[-(Y_{it} - D_{it}), 0]. \tag{8}$$

For example, suppose Gavi delivered sufficient doses of pentavalent vaccines to immunize 97% of infants in Djibouti against Hib in 2011, yet Djibouti recorded a Hib vaccination rate of just

87%. We interpret this as an indication that vaccines sufficient to cover 10% of Djiboutian infants were wasted. This estimate is conservative inasmuch as the government Djibouti (or private Djiboutian households) may have also bought some vaccines from domestic resources, pushing the implied wastage rate even higher.

There are three main assumptions required for our estimates of waste. The first is that vaccines are used only for the target population, i.e., infants under one year of age. As discussed in Section 2, there is strong evidence from the Demographic and Health Surveys that this is indeed the case (Clark and Sanderson, 2009). The second is that our constructed measure of Gavi-purchased doses accurately reflects the number of doses provided by Gavi. As prices per dose of vaccine varied slightly across vaccine presentation, waste may be over- or under-estimated depending on the exact price paid in each country.

We must also make some assumption about the time period over which vaccines apply. There is some evidence of lumpiness in Gavi deliveries – i.e., countries receiving in excess of 100% of their needs in one year, and significantly less in the following year. In this case, equation (8) will overestimate waste. To avoid any overestimation of waste, we again take a very conservative stance, and average all aid receipts and vaccination rates within countries over all the years in our data set in which Gavi was active, 2001-2013. In effect, we replace equation (8) with

$$\text{Waste}_{it} \equiv \max[-(\bar{Y}_i - \bar{D}_i), 0]. \tag{9}$$

By averaging over years, this implicitly allows that vaccines never spoil and are saved for future years, and assumes only that they are not reallocated across countries once they are disbursed.

For each of six vaccines, a scatterplot is shown in Figure 6 with the average doses delivered per annum in terms of the population under 1 (D_i) for each country plotted again average vaccination rate.[14] The solid black line is a 45-degree line. Points below the line represent countries which received more doses from Gavi on average than the number of children under 1 vaccinated over the same years. For hepatitis B, for instance, Ghana received enough doses per annum to vaccinate 83% of children, but had an average vaccination rate of 75% during the same time period. Across all countries receiving hepatitis B vaccines from Gavi, the number of doses lost to waste was sufficient to vaccinate approximately 19%

[14]The sample includes only Gavi-eligible countries that received greater than zero doses of the vaccine in question in each year (i.e., compliers).In the figure, the percent of dose series purchased by Gavi is capped at 100%.

of the population – though we would emphasize that in many cases waste occurred where vaccination rates were approaching 100%, so this reflects a global inefficiency, but often a simple arithmetic inevitability within countries. For other diseases, the doses delivered and waste rates are considerably lower, as seen in the figure.

So far we have focused on average waste level across all aid recipients. To compare these numbers to our estimates of impact and fungibility, we need to know the waste levels for a representative country at the Gavi eligibility threshold. To calculate this waste level, we repeat our benchmark RDD regression specification, using waste as the dependent variable. Results are reported in Table 10. In our preferred specification (linear, bandwidth of one log point of GNI), waste rates are statistically indistinguishable from zero for four of the five diseases calculated. (The same would be true for measles, which we do not report because Gavi disbursed so few measles doses that waste is not a significant issue.) The one exception is Hib, where we estimate that 6% of doses were lost to waste for a country near the threshold. Some care is required in interpreting this result, as in most cases, Gavi distributed hepatitis B, Hib, and DPT as a bundle through the pentavalent vaccine. Notably, when we average the vaccination rates across these three vaccines and compare the pentavalent deliveries, we find no evidence of waste at the threshold.

4.6 Summary decomposition

The econometric estimates from the previous sections allow us to divide the vaccines provided by Gavi's aid program between those that immunized children who wouldn't have otherwise have been vaccinated (impact), vaccines that immunized children who would have been vaccinated anyway (fungibility), and vaccines that were apparently never administered to anyone (waste).

$$\text{Gavi aid} \equiv \text{Impact} + \text{Fungibility} + \text{Waste}$$

In theory this is an accounting identity. In practice, we estimate each piece in a separate regression, and the results presented in Table 11 are not strictly guaranteed to add up, but for the most part come fairly close.

To see how the decomposition works, imagine a hypothetical country which just qualifies for Gavi aid – i.e., a country on the regression line at the income threshold. It would have received enough vaccines from Gavi from 2001 to 2013 to vaccinate 66% of infants against hepatitis B each year. Our impact estimates (Table 7) suggest it vaccinated only 0.9% more

infants thanks to aid. Fungibility and waste together (Table 9) accounted for enough vaccines to cover about 59% of infants. Little, if any, of that was pure waste (Table 10). The story is more or less the same for DPT, though aid flows were smaller: no impact, no waste, but lots of crowding out.

For Hib, the same country would have received enough aid to vaccinate about 35% of infants each year. There is still no significant evidence of a positive impact, but the point estimates are much larger – representing a TOT effect of about 16 percentage points. The point estimate for fungibility and waste is zero, so clearly the total does not add up to 35 percentage points. In essence, the data is too noisy to provide a firm indication of where the Gavi-funded vaccines ended up, but point towards a non-trivial improvement in coverage. This is consistent with the fact that baseline coverage in 2000 for Hib was much lower than for hepatitis B and DPT. Gavi's decision to prioritize the pentavalent vaccine in lieu of all three of those vaccines might simultaneously have had no impact on hepatitis B or DPT coverage while pulling up Hib rates.

Because it is based on the simple RD specification (and not the dose-response model) the decomposition is relatively uninformative for the other newer vaccines introduced in later Gavi phases, pneumococcal disease and rotavirus. The data cannot reliably distinguish potential positive impacts of 1 to 3 percentage points from zero, or from outright waste. Finally, as a validation of the data overall, the column corresponding to measles – which Gavi essentially did not fund, and we use as a placebo test – shows no impacts, no fungibility, and no waste.

5 Conclusion

To summarize, we have attempted to measure the impact on vaccination coverage rates of a foreign aid program, known as Gavi, which has provided large quantities of free and subsidized vaccines to developing countries since 2001. We distinguish between diseases covered by older, cheaper vaccines – diphtheria, pertussis, tetanus, and hepatitis B – and diseases covered by newer, more expensive vaccines, such as Haemophilus influenza type B, rotavirus, and pneumococcal disease.

For the older, cheaper vaccines, we find no statistically significant impact on vaccination rates. This is true despite the fact that Gavi often delivered sufficient vaccines to cover

100% of the infant population in eligible countries. The lack of impact can be explained by the fairly high and rising rates of coverage for these vaccines in putatively comparable ineligible countries. Furthermore, the lack of detectable impact does not appear to be solely an artefact of noisy data. There is statistically significant evidence of fungibility, in the sense that subsidized vaccines appear to have displaced domestic vaccination spending in countries near the threshold. There is relatively little direct evidence, however, of outright waste of Gavi-purchased vaccines.

For the newer, more expensive vaccines, results are more promising. In our preferred specification, we find a positive, statistically significant impact of Gavi aid on Hib vaccination rates. For both Hib and rotavirus, point estimates imply Gavi aid raised vaccination rates one-for-one, and offer no evidence of fungibility or waste. For pneumococcal disease, point estimates are much more modest, possibly attributable to the still modest roll-out of the vaccine to date.

Methodologically, our estimates based on a regression discontinuity design are broadly consistent with some previous panel data analyses (Lu et al., 2006), inasmuch as both find no significant impact of Gavi on DPT vaccination rates for countries which are wealthier or had more developed vaccination systems at baseline. But our results contradict other previous panel data analyses (Chee et al., 2008) which have found positive effects for countries with higher baseline vaccination rates. In addition to a clearer identification strategy than these previous studies which relied on internal instruments in a GMM framework, we also have also significantly expanded the analysis by looking beyond DPT to five other vaccines, including those which have been areas of major focus for Gavi in more recent years (pneumococcal disease and rotavirus).

It is important to emphasize, yet again, that all of our econometric estimates are applicable to countries near the aid eligibility threshold. There are *a priori* reasons to suspect the impact of heavily subsidized vaccines will be smallest near the threshold where our regression discontinuity estimates apply. Free vaccines will be more likely to displace domestic vaccination programs in these middle-income countries, compared to low-income countries which may have weaker vaccination programs in the absence of aid. We also note that we have not estimated Gavi's impact on morbidity or other health outcomes via, for instance, the dissemination of safer vaccination techniques.

Our estimates have focused on the impact of Gavi's direct aid to individual countries – its private transfers, as opposed to its public goods provision and activities that may have

had substantial international spillovers. Gavi supported the development of a new vaccine, and it is not evident that the pneumococcal vaccine including serotypes prevalent in low-income countries would have existed by 2013 absent Gavi's advance market commitment. Additionally, Gavi's efforts may have had indirect impacts that extended beyond the strict eligibility threshold and are not captured here. Large aid purchases grew the market for vaccines and may have contributed to a drop in vaccine prices over time. The Alliance has a specific strategy to bring new entrants into the vaccine market and drive down prices. This may have had an effect on prices faced by non-Gavi countries. Between 2000 and 2010 the price of the hepatitis B vaccine dropped from US$ 0.55 to US$ 0.19 per dose, for example. Gavi also brought enormous international attention to the issue of vaccine coverage. This may have altered political calculations regarding vaccine spending and coverage above the threshold.

Nevertheless, our finding that Gavi funding for older, cheaper vaccines largely displaced funding for domestic vaccination programs in threshold countries may have broader relevance to the aid literature and policy discussions. This finding casts doubt on the aid strategy of using in-kind transfers to bypass concerns about fungibility, or of targeting only low-cost, easy-to-deliver interventions where the potential for duplication may be high. The design of Gavi's intervention may have essentially forced fungibility. If the Alliance provided enough pentavalent vaccines to fully immunize every child in the country, clearly countries would (and should) not purchase additional pentavalent vaccines even if they would have done so absent Gavi's intervention.

Turning to implications for Gavi's aid program in the future, our findings are open to a variety of interpretations. An obvious response would be to lower the income eligibility threshold for subsidized vaccines that are cheap to buy, in order to avoid duplicating national efforts, since aid appears to have little or no marginal impact on vaccination rates near this threshold. Alternatively, Gavi might consider *raising* the GNI threshold to give even more countries access to the bulk purchasing discounts it receives through UNICEF, while phasing out the actual aid transfer of subsidized vaccines. The wisdom of these or other approaches hinges on feasibility considerations beyond the scope of this study.

References

Acemoglu, D. and S. Johnson (2007). Diseases and development: The effect of life expectancy on economic growth. Journal of Political Economy 115(6), 925.

Arndt, C., S. Jones, and F. Tarp (2014). Assessing foreign aid's long run contribution to growth and development. World Development.

Black, R. E., S. S. Morris, and J. Bryce (2003). Where and why are 10 million children dying every year? The Lancet 361(9376), 2226–2234.

Blundell, R. and S. Bond (1998). Initial conditions and moment restrictions in dynamic panel data models. Journal of econometrics 87(1), 115–143.

Bosch-Capblanch, X., O. Ronveaux, V. Doyle, V. Remedios, and A. Bchir (2009). Accuracy and quality of immunization information systems in forty-one low income countries. Tropical Medicine & International Health 14(1), 2–10.

Burnside, C. and D. Dollar (2000). Aid, policies, and growth. American Economic Review 90(4), 847–868.

Burton, A., R. Monasch, B. Lautenbach, M. Gacic-Dobo, M. Neill, R. Karimov, L. Wolfson, G. Jones, and M. Birmingham (2009). WHO and UNICEF estimates of national infant immunization coverage: methods and processes. Bulletin of the World Health Organization 87(7), 535–541.

Chee, G., V. Molldrem, N. His, and S. Chankova (2008). Evaluation of GAVI phase 1 performance. Bethesda, MD: Abt Associates, Inc.

Clark, A. and C. Sanderson (2009). Timing of children's vaccinations in 45 low-income and middle-income countries: An analysis of survey data. The Lancet 373(9674), 1543–1549.

Clemens, M. A., S. Radelet, R. R. Bhavnani, and S. Bazzi (2012). Counting chickens when they hatch: Timing and the effects of aid on growth. The Economic Journal 122(561), 590–617.

de Sijpe, N. V. (2013, December). The fungibility of health aid reconsidered. Journal of Development Studies 49(12), 1746–1754.

Deaton, A. (2013). The great escape: Health, wealth, and the origins of inequality. Princeton: Princeton University Press.

Devarajan, S., A. S. Rajkumar, and V. Swaroop (1999). What does aid to Africa finance? Mimeograph, Development Research Group, World Bank, Washington D.C.

Easterly, W. (2014). The tyranny of experts: Economists, dictators, and the forgotten rights of the poor. New York: Basic Books.

Easterly, W., R. Levine, and D. Roodman (2004). Aid, policies, and growth: Comment. American Economic Review, 774–780.

Feyzioglu, T., V. Swaroop, and M. Zhu (1998). A panel data analysis of the fungibility of foreign aid. The World Bank Economic Review 12(1), 29–58.

Galiani, S., S. Knack, L. C. Xu, and B. Zou (2014). The effect of aid on growth: Evidence from a quasi-experiment. World Bank Policy Research Working Paper 6865.

Gavi (1999). First board meeting minutes. http://libdoc.who.int/hq/1999/GAVI_99.02.pdf. Online; accessed 2014-08-06.

Gavi (2009). Review of the eligibility policy. http://www.gavialliance.org/About/Governance/GAVI-Board/Minutes/2011/16-Nov/Minutes/Review-of-the-Eligibility-Threshold/. Online; accessed 2014-06-21.

Gavi (2014a). Countries eligible for support. http://www.gavialliance.org/support/apply/countries-eligible-for-support/. Online; accessed 2014-06-21.

Gavi (2014b). Healthy children with a healthy future. http://www.gavialliance.org/Library/Publications/GAVI/GAVI-Alliance-replenishment-brochure/. Online; accessed 2014-08-06.

Gelman, A. and G. Imbens (2014). Why high-order polynomials should not be used in regression discontinuity designs. www.nber.org/papers/w20405. NBER Working Paper 20405.

Goldstein, S. T., F. Zhou, S. C. Hadler, B. P. Bell, E. E. Mast, and H. S. Margolis (2005). A mathematical model to estimate global hepatitis b disease burden and vaccination impact. International Journal of Epidemiology 34(6), 1329–1339.

Hahn, J., P. Todd, and W. Van der Klaauw (2001). Identification and estimation of treatment effects with a regression-discontinuity design. Econometrica 69(1), 201–209.

Hulls, D., P. Venkatachalam, K. Kumar, T. Cochrane, H. Kaur, N. Gulati, and D. Jones (2010). Second GAVI evaluation report. http://www.gavialliance.org/library/gavi-documents/evaluations/second-gavi-evaluation-2006-2010/. Online; accessed 2014-06-05.

Imbens, G. and K. Kalyanaraman (2012). Optimal bandwidth choice for the regression discontinuity estimator. The Review of Economic Studies 79(3), 933–959.

Knack, S., L. C. Xu, and B. Zou (2014). Interactions among donors aid allocations: Evidence from an exogenous world bank income threshold. World Bank Policy Research Working Paper 7039.

Lee, D. S. and T. Lemieux (2009). Regression discontinuity designs in economics. NBER Working Paper No. 14723.

Lessler, J., C. J. E. Metcalf, R. F. Grais, F. J. Luquero, D. A. Cummings, and B. T. Grenfell (2011). Measuring the performance of vaccination programs using cross-sectional surveys: A likelihood framework and retrospective analysis. PLoS medicine 8(10), e1001110.

Lim, S. S., D. B. Stein, A. Charrow, and C. J. Murray (2008). Tracking progress towards universal childhood immunisation and the impact of global initiatives: A systematic analysis of three-dose diphtheria, tetanus, and pertussis immunisation coverage. The Lancet 372(9655), 2031–2046.

Lu, C., C. M. Michaud, E. Gakidou, K. Khan, and C. J. Murray (2006). Effect of the Global Alliance for Vaccines and Immunisation on diphtheria, tetanus, and pertussis vaccine coverage: An independent assessment. The Lancet 368(9541), 1088–1095.

Lu, C., M. T. Schneider, P. Gubbins, K. Leach-Kemon, D. Jamison, and C. J. Murray (2010). Public financing of health in developing countries: A cross-national systematic analysis. The Lancet 375(9723), 1375–1387.

Mishra, P. and D. Newhouse (2009). Does health aid matter? Journal of Health Economics 28(4), 855–872.

Mitchell, V., V. J. Dietz, J. M. Okwo-Bele, and F. T. Cutts (2012). Immunization in developing countries. In S. Plotkin, W. Orenstein, and P. Offit (Eds.), Vaccines (Sixth edition ed.). Saunders.

Morris, S. K., W. J. Moss, and N. Halsey (2008). Haemophilus influenzae type b conjugate vaccine use and effectiveness. The Lancet infectious diseases 8(7), 435–443.

Murray, C. J., B. Shengelia, N. Gupta, S. Moussavi, A. Tandon, and M. Thieren (2003). Validity of reported vaccination coverage in 45 countries. The Lancet 362(9389), 1022–1027.

Nichols, A. (2007). Causal inference with observational data. Stata Journal 7(4), 507.

O'Brien, K. L., L. J. Wolfson, J. P. Watt, E. Henkle, M. Deloria-Knoll, N. McCall, E. Lee, K. Mulholland, O. S. Levine, and T. Cherian (2009). Burden of disease caused by streptococcus pneumoniae in children younger than 5 years: Global estimates. The Lancet 374(9693), 893–902.

Pack, H. and J. R. Pack (1990). Is foreign aid fungible? The case of Indonesia. The Economic Journal, 188–194.

Pack, H. and J. R. Pack (1993). Foreign aid and the question of fungibility. The Review of Economics and Statistics, 258–265.

Rajan, R. G. and A. Subramanian (2008). Aid and growth: What does the cross-country evidence really show? The Review of Economics and Statistics 90(4), 643–665.

Ronveaux, O., D. Rickert, S. Hadler, H. Groom, J. Lloyd, A. Bchir, and M. Birmingham (2005). The immunization data quality audit: Verifying the quality and consistency of immunization monitoring systems. Bulletin of the World Health Organization 83(7), 503–510.

Roodman, D. (2009). A note on the theme of too many instruments*. Oxford Bulletin of Economics and Statistics 71(1), 135–158.

Roodman, D. (2012). Doubts about the evidence that foreign aid for health is displaced into non-health uses. The Lancet 380(9846), 972–973.

Sandefur, J. and A. Glassman (2014). The political economy of bad data: Evidence from african survey and administrative statistics. CGD Working Paper 373.

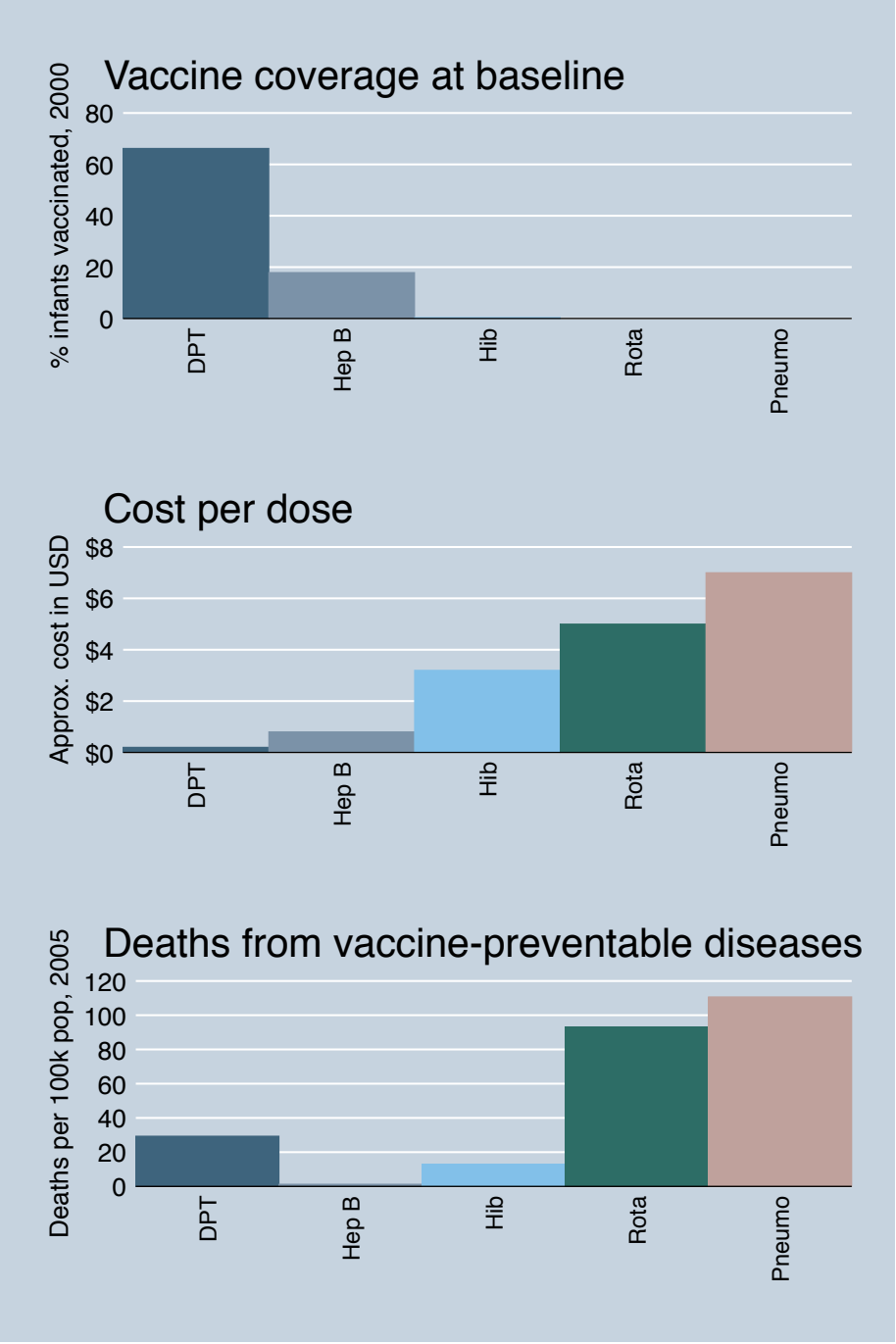
Van de Walle, D. and R. Mu (2007). Fungibility and the flypaper effect of project aid: Micro-evidence for vietnam. Journal of Development Economics 84(2), 667–685.

Watt, J. P., L. J. Wolfson, K. L. O'Brien, E. Henkle, M. Deloria-Knoll, N. McCall, E. Lee, O. S. Levine, R. Hajjeh, K. Mulholland, et al. (2009). Burden of disease caused by haemophilus influenzae type b in children younger than 5 years: Global estimates. The Lancet 374(9693), 903–911.

WHO/UNICEF (2012a). Estimated rotavirus deaths for children under 5 years of age. http://www.who.int/immunization/monitoring_surveillance/burden/estimates/rotavirus/en/. Online; accessed 2014-08-08.

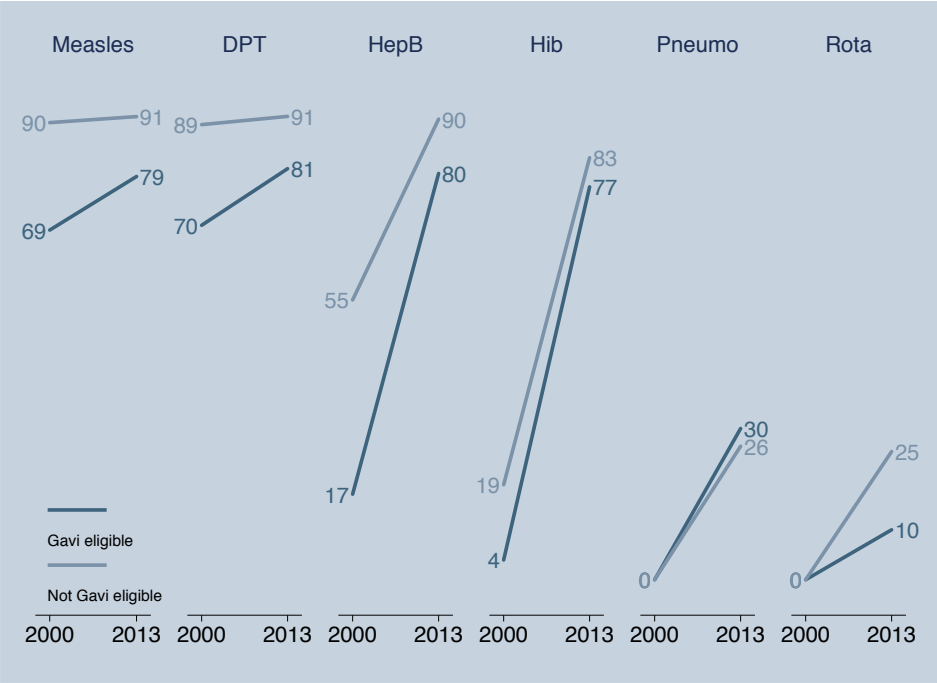
WHO/UNICEF (2012b). User's reference to country reports of WHO and UNICEF estimates of national infant immunization coverage. http://www.childinfo.org/files/Immunization_WUENIC_guide_and_mark-up.pdf. Online; accessed 2014-06-21.

Figure 1: Coverage, cost, and mortality by vaccine



Coverage and mortality numbers are average values for the 71 Gavi-eligible countries with data available at baseline, weighted by their infant population. Mortality figures are taken from the Global Burden of Disease database (IHME 2013). For rotavirus, death rates are from 2008. For pneumococcal vaccine, all meningitis deaths and 50% of deaths from lower-respiratory infections are treated as vaccine preventable, based on Riley (1986). Note that death rates are influenced by vaccination coverage rates, and would be higher in the absence of high coverage rates for DPT and hepatitis B.

Figure 2: Immunization rates by Gavi eligibility, 2000 to 2013



The figure shows simple averages of country vaccination rates across the Gavi-eligible and ineligible countries in each year. Note that the pool of eligible countries changes over time. Eligible countries include many which received no Gavi aid in a given year, and vice versa.

Figure 3: Government and non-government vaccine financing, 2006 and 2012

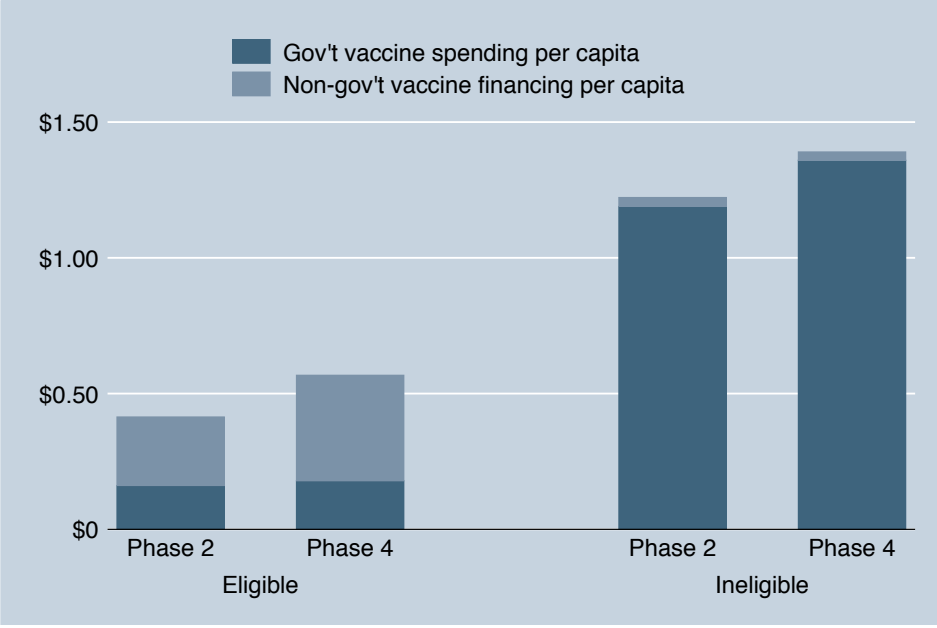
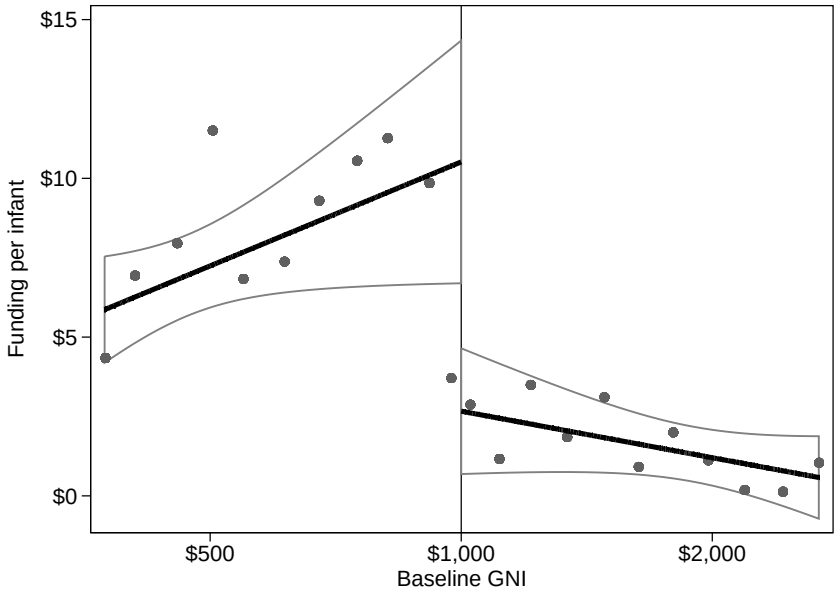
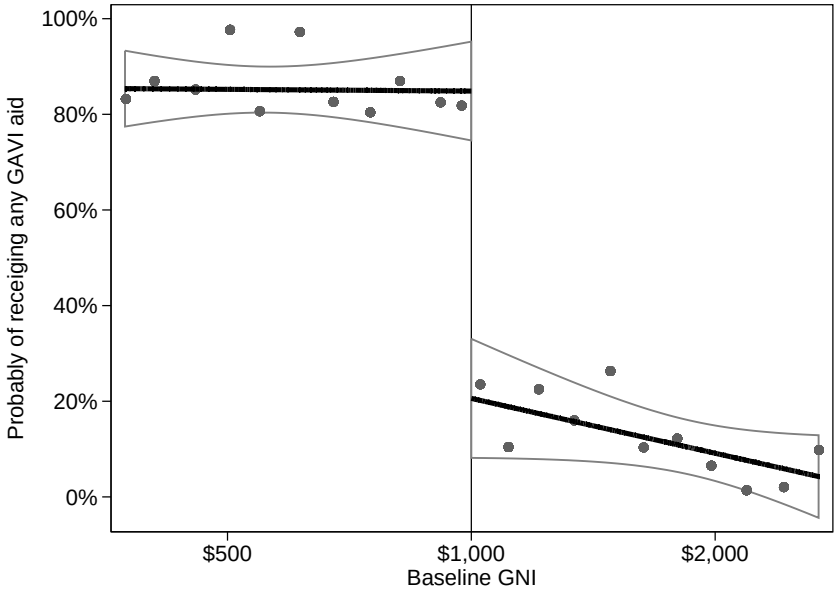


Figure 4: Gavi's adherence to its income eligibility threshold

(a) Funding amount per infant

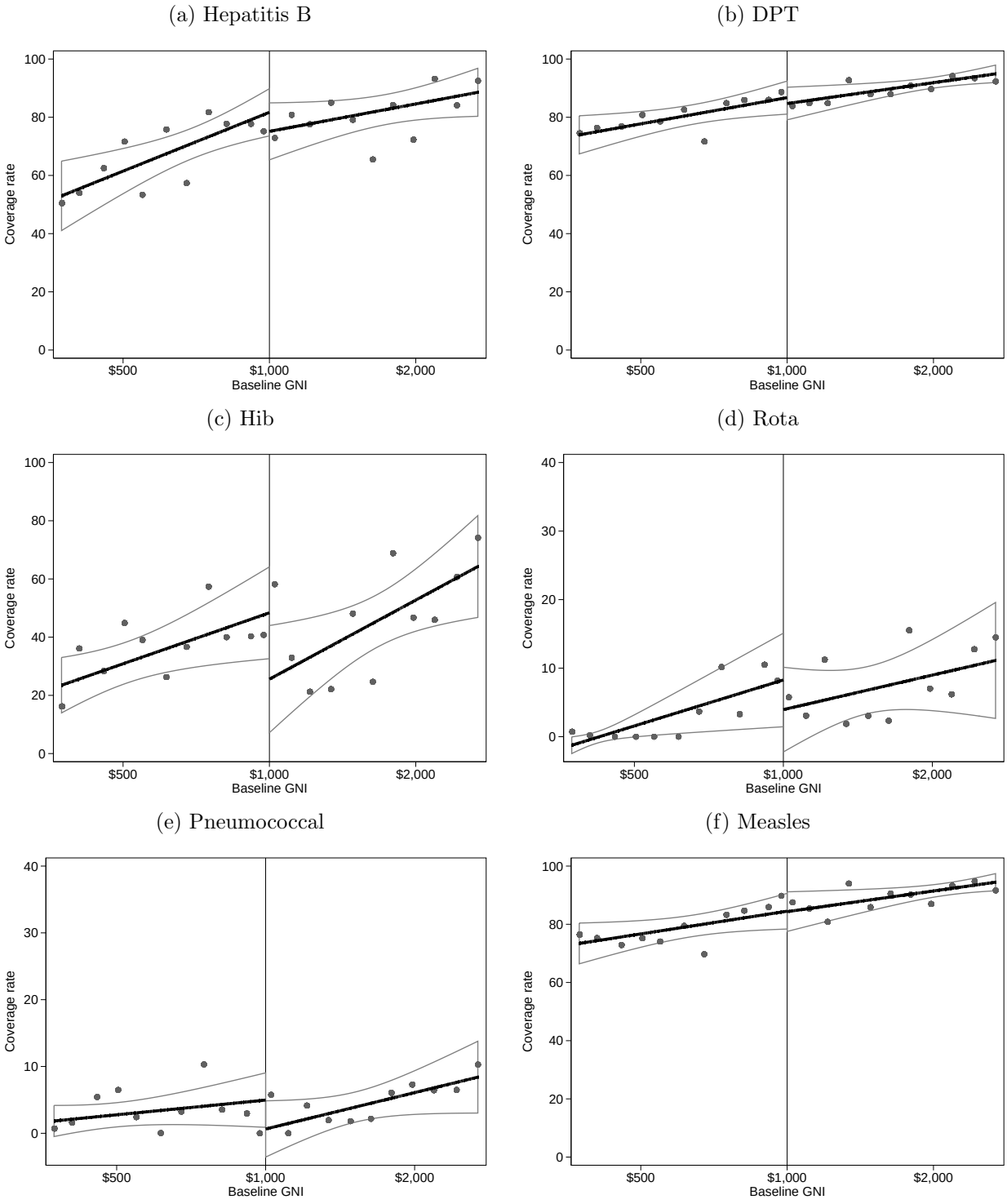


(b) Probability of receiving funding



The figures measure Gavi's adherence to its own eligibility rules. Each graph presents a linear prediction plot of Gavi funding (dollars per child in top graph, and a binary indicator for receipt of funding in bottom graph) on countries' baseline log per capita GNI, run separately on each side of the administrative cut-off. Gray lines denote the 95% confidence interval around the regression line. Scatterplots show the average value within bins 0.1 log points wide.

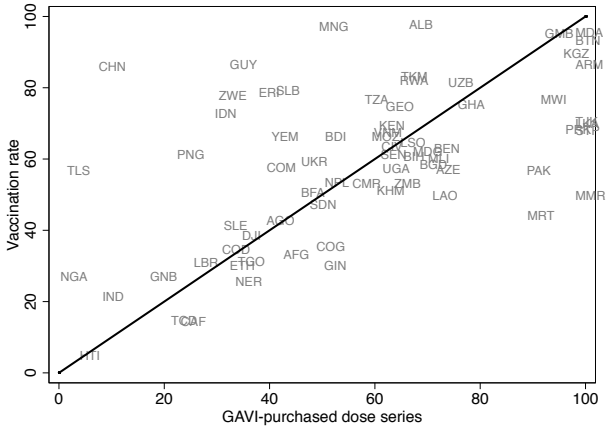
Figure 5: Vaccination rates, 2000-2013, by GNI



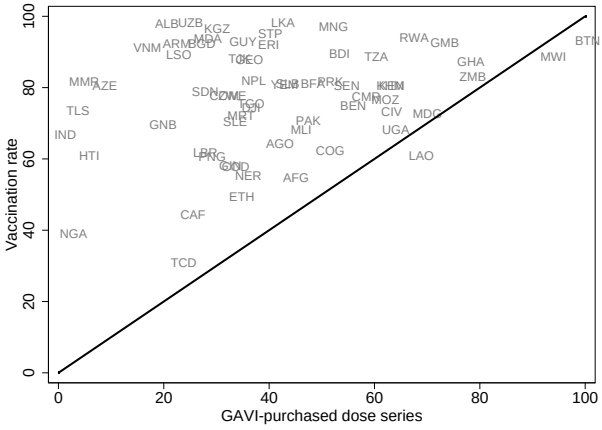
Figures show a linear prediction plot of vaccine coverage on countries' baseline log per capita GNI, run separately on each side of the administrative cut-off. The solid black lines use the actual data series. Gray lines denote the 95% confidence interval around the regression line. Scatterplots show the average value within bins 0.1 log points wide.

Figure 6: Gavi-purchased dose series and vaccination rates, 2001-2013

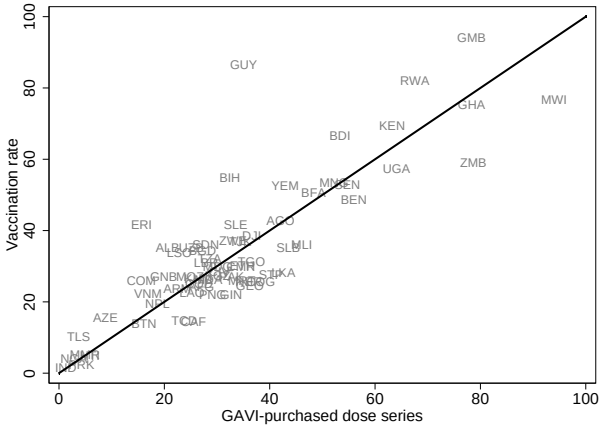
(a) Hepatitis B



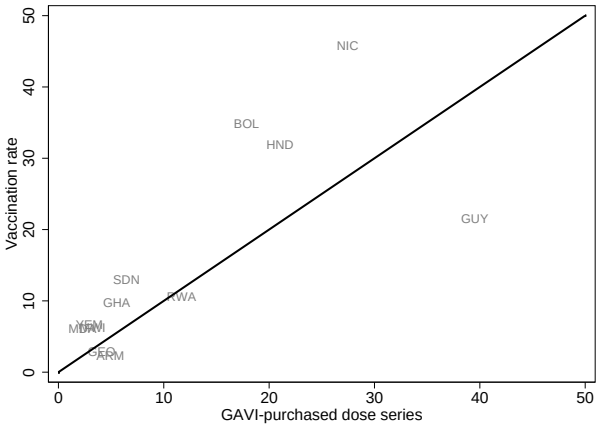
(b) DPT



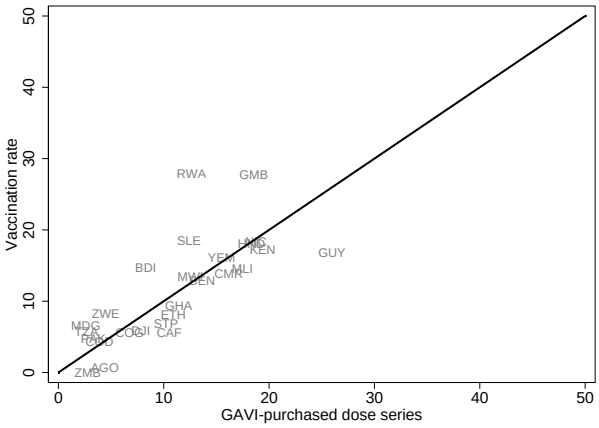
(c) Hib



(d) Rota



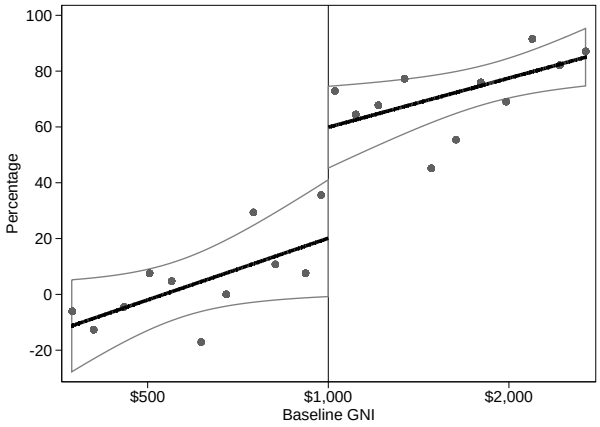
(e) Pneumo



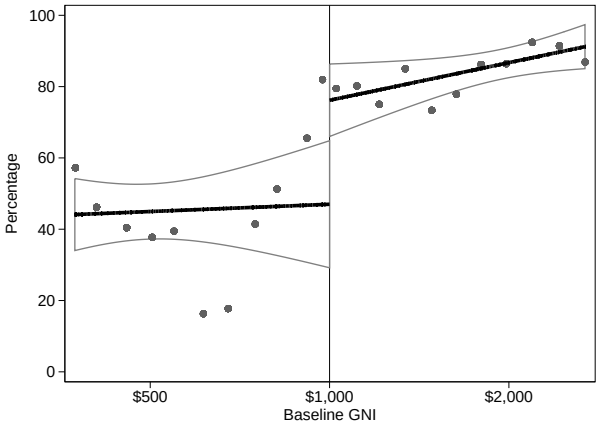
Scatterplots show the average estimated percent of infants who could be fully immunized using vaccines bought by Gavi, and the average vaccination rate by country in 2001-2013 for countries receiving Gavi aid for at least one year. The solid black line is a 45-degree line. Points below the line represent countries which received more dose series from Gavi than the number of children vaccinated, on average. The percent of infants receiving dose series purchased by Gavi is capped at 100%.

Figure 7: Testing levels of fungibility, 2000-2013

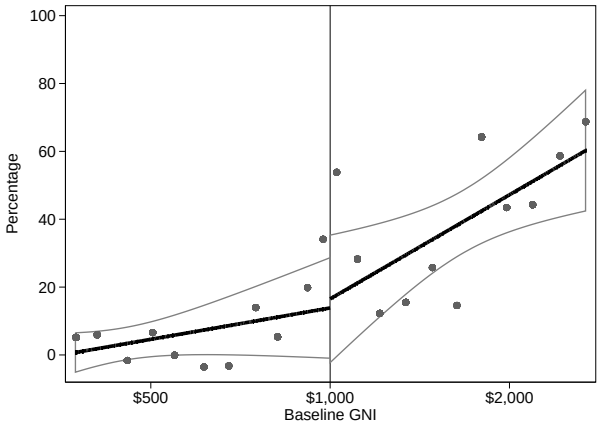
(a) Hepatitis B



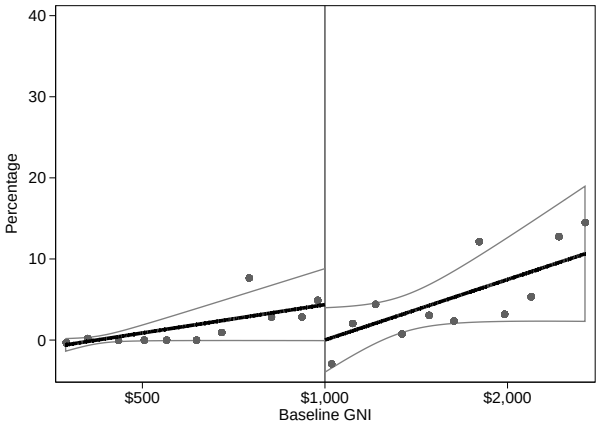
(b) DPT



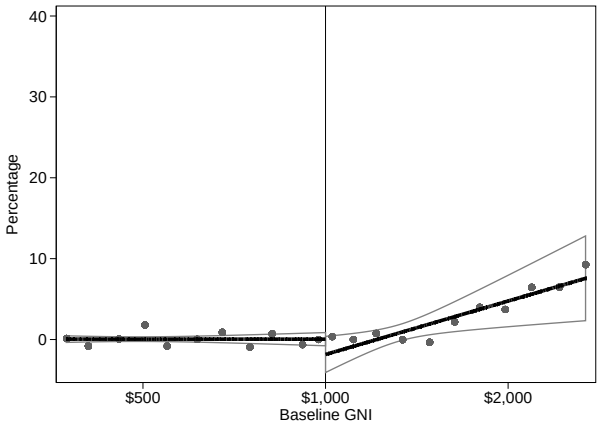
(c) Hib



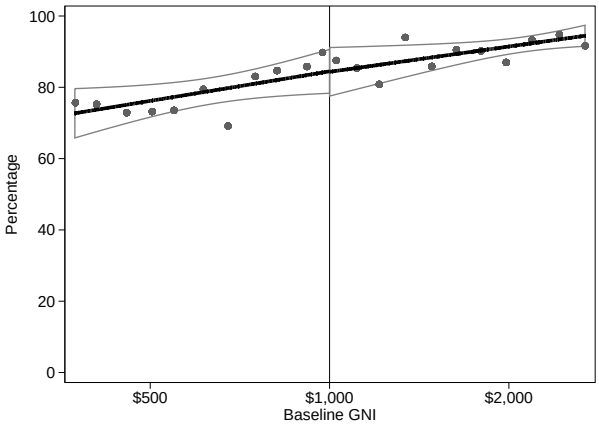
(d) Rota



(e) Pneumococcal

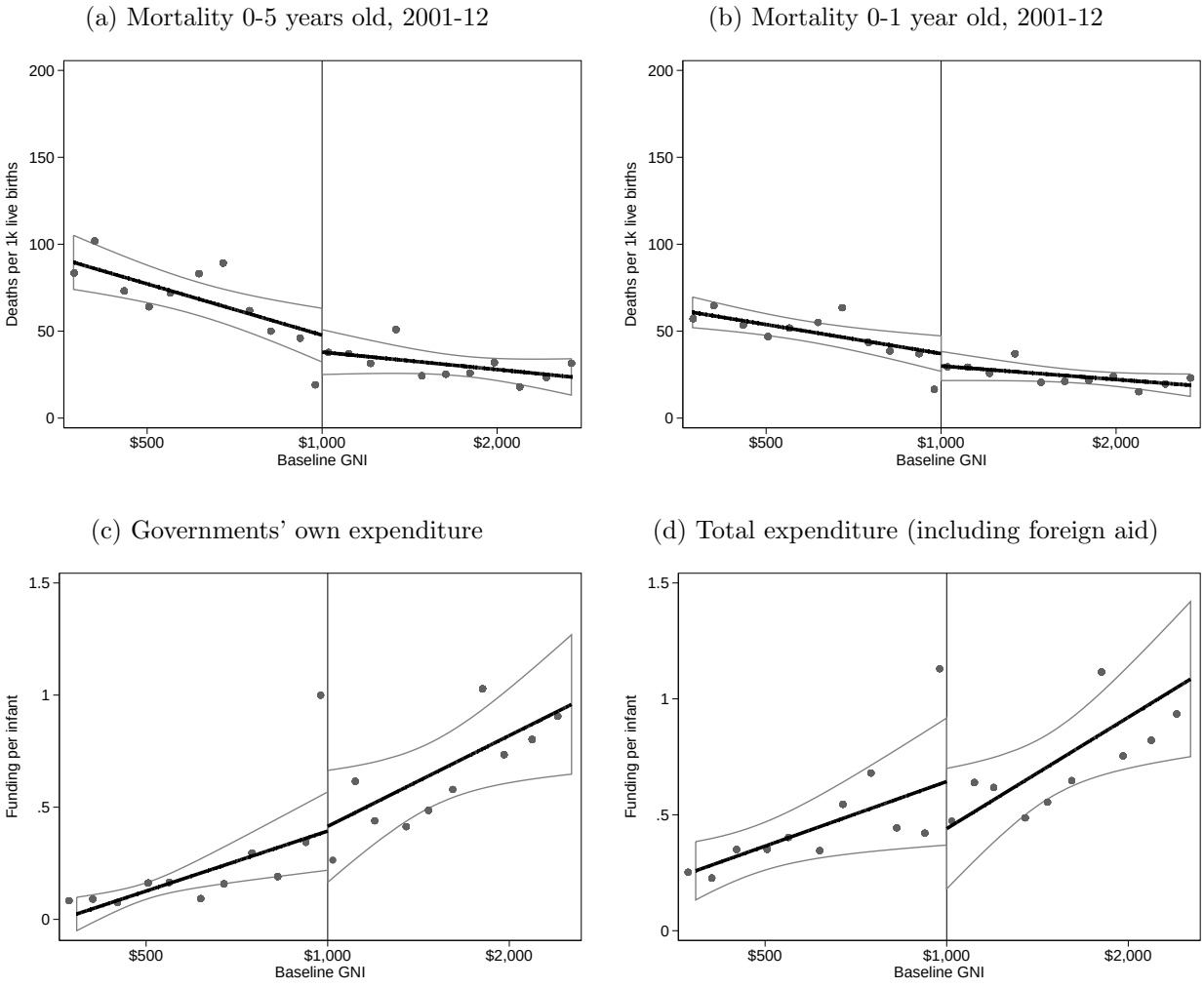


(f) Measles



Figures show a linear prediction plot of vaccine coverage on countries' baseline log per capita GNI, run separately on each side of the administrative cut-off. The solid black lines represent the coverage rate, subtracting the percent of infants who could be fully immunized using our estimate of dose series bought by Gavi: i.e., the counterfactual to Gavi if all Gavi aid represents a marginal increase in vaccination rates. Gray lines denote the 95% confidence interval around the regression line. Scatterplots show the average value within bins 0.1 log points wide.

Figure 8: Child mortality and vaccination funding



Panel (a) shows a linear prediction plot of under-five mortality (deaths per 1,000 live births) on countries' baseline log per capita GNI, run separately on each side of the administrative cut-off. Gray lines denote the 95% confidence interval around the regression line. Scatterplots show the average value within vertical bins 0.1 log points wide. Panel (b) repeats this exercise using mortality rate on the y-axis. Panel (c) shows results for the log of per capita expenditure on vaccines by national governments. Fungibility would imply a negative Gavi treatment effect. Panel (d) shows results for the log of total per capital expenditure on vaccines, including domestic spending and foreign aid. Anything short of a positive treatment effect equal to the amount of Gavi aid could be interpreted as evidence of fungibility.

Table 1: Summary statistics

	Baseline		Phase 1		Phase 2		Phase 3		Phase 4		Phase 5	
	Eligible	Ineligible	Eligible	Ineligible	Eligible	Ineligible	Eligible	Ineligible	Eligible	Ineligible	Eligible	Ineligible
DPT (coverage rate)	69.5	89.3	74.1	89.6	80.0	92.3	80.9	90.7	81.6	91.3	80.6	90.9
	(21.0)	(12.1)	(19.0)	(12.3)	(15.9)	(9.28)	(16.6)	(9.25)	(15.9)	(9.33)	(16.0)	(11.0)
Hepatitis B (coverage rate)	16.8	54.9	41.9	70.6	72.2	89.0	77.6	90.1	79.3	90.7	79.7	90.4
	(32.5)	(43.8)	(40.2)	(37.0)	(29.5)	(18.6)	(23.2)	(11.6)	(21.3)	(11.4)	(19.2)	(11.9)
Hib (coverage rate)	3.87	18.6	11.3	41.1	40.5	64.0	73.6	74.9	75.0	79.6	77.1	82.8
	(18.3)	(34.3)	(28.5)	(44.1)	(41.9)	(41.9)	(30.0)	(33.0)	(27.9)	(30.3)	(23.4)	(25.7)
Rotavirus (coverage rate)	0	0	0	0	2.29	7.33	2.40	17.7	5.81	19.9	9.79	25.1
	(0)	(0)	(0)	(0)	(13.6)	(21.9)	(14.0)	(34.2)	(19.3)	(35.7)	(26.9)	(37.7)
Pneumococcal (coverage rate)	0	0	0	0	0.58	3.84	11.6	16.2	19.8	23.4	29.7	26.2
	(0)	(0)	(0)	(0)	(7.51)	(16.1)	(26.7)	(31.3)	(35.5)	(37.6)	(39.9)	(39.3)
Measles (coverage rate)	68.6	89.7	73.5	89.6	78.2	91.7	78.9	90.8	80	90.9	79.1	90.8
	(20.3)	(10.5)	(18.7)	(12.4)	(16.0)	(10.8)	(15.2)	(9.69)	(14.9)	(9.83)	(16.3)	(10.6)
Under-5 mortality rate	108.6	33.1	89.6	27.7	77.2	20.8	.	.	77.9	25.2	.	.
	(56.6)	(25.7)	(49.3)	(24.1)	(42.6)	(15.7)	(.)	(.)	(34.4)	(22.9)	(.)	(.)
Infant mortality rate	70.6	26.0	59.8	21.8	52.7	16.9	.	.	54.1	20.1	.	.
	(30.3)	(17.1)	(27.5)	(15.8)	(24.0)	(11.0)	(.)	(.)	(20.2)	(15.0)	(.)	(.)
GNI per capita ($)	430.8	2682.7	430.8	2682.7	479.4	2959.4	729.1	4652.7	778.1	4899.4	798.2	5126.9
	(212.0)	(1479.5)	(210.7)	(1470.6)	(230.5)	(1539.9)	(312.2)	(2317.7)	(328.7)	(2366.7)	(364.0)	(2471.5)
Gov't vaccine spending ($/capita)	.	.	.	.	0.17	1.14	0.093	1.14	0.096	1.09	.	.
	(.)	(.)	(.)	(.)	(0.23)	(1.33)	(0.097)	(1.19)	(0.094)	(1.03)	(.)	(.)
Total vaccine spending ($/capita)	.	.	.	.	0.42	1.46	0.56	1.55	0.53	1.53	.	.
	(.)	(.)	(.)	(.)	(0.50)	(3.04)	(0.74)	(2.99)	(0.46)	(2.95)	(.)	(.)
GAVI funding ($/capita)	0	0	0.10	0.00010	0.26	0.0026	0.44	0.057	0.65	0.074	0.72	0.076
	(0)	(0)	(0.16)	(0.0015)	(0.23)	(0.017)	(0.36)	(0.17)	(0.48)	(0.21)	(0.53)	(0.19)
GAVI funding ($/infant)	0	0	3.21	0.0038	8.45	0.24	13.5	3.35	21.1	3.33	23.5	3.49
	(0)	(0)	(4.45)	(0.054)	(7.04)	(1.56)	(10.7)	(9.22)	(15.7)	(8.70)	(16.5)	(7.58)
Probability of receiving GAVI (%)	0	0	82.0	0.60	91.2	3.53	98.1	19.5	100	19.8	98.1	24.4
	(0)	(0)	(38.5)	(7.72)	(28.4)	(18.5)	(13.9)	(39.9)	(0)	(40.1)	(13.7)	(43.2)
N	69	67	69	67	68	68	52	82	52	81	53	78

Note: The table reports the mean and standard deviation across countries and years for all variables, except GNI per capita, which is represented by the GNI used to determine eligibility (1998 GNI per capita for phase 1, for example). Our sample is restricted to countries within 2 log points of the eligibility threshold. Baseline is defined as the year 2000, phase 1: 2001-2005, phase 2: 2006-2010, phase 3: 2011, phase 4: 2012 and phase 5: 2013. More information on all variables is available in the web appendix.

Table 2: Aid and income eligibility: first-stage regressions

	Bandwidth = 0.5		Bandwidth = 1		Bandwidth = 2	
	(1)	(2)	(3)	(4)	(5)	(6)
E	0.645***	0.744***	0.643***	0.588***	0.642***	0.602***
	(0.0999)	(0.158)	(0.0789)	(0.108)	(0.0563)	(0.0798)
$\bar{Y}$	0.0916	0.143	-0.164*	-0.114	-0.116***	-0.223*
	(0.322)	(1.048)	(0.0901)	(0.353)	(0.0313)	(0.119)
$\bar{Y}^2$		-0.103		-0.0492		0.0558
		(2.008)		(0.328)		(0.0517)
$E \times \bar{Y}$	-0.269	0.791	0.159	-0.247	0.0416	0.133
	(0.384)	(1.247)	(0.121)	(0.447)	(0.0475)	(0.173)
$E \times \bar{Y}^2$		2.296		-0.275		-0.0643
		(2.834)		(0.467)		(0.0755)
Constant	0.156*	0.151	0.206***	0.197**	0.180***	0.216***
	(0.0852)	(0.120)	(0.0635)	(0.0857)	(0.0449)	(0.0655)
Optimal b.w.						
Observations	435	435	1014	1014	1758	1758
No. of countries	64	64	102	102	150	150
R-squared	0.45	0.45	0.54	0.54	0.68	0.68

Note: See equation 5 in the text. The table reports estimates of the first-stage regression for the 2SLS fuzzy RD model. Each column presents a separate OLS regression, the dependent variable is the Gavi treatment dummy. Standard errors are clustered at the country level.

Table 3: Balance test: applying the RD to baseline data

	Vaccination Rate					Mortality Rate
	(1) Hep B	(2) DPT	(3) Hib	(4) Measles	(5) $\leq$ 5 yr	(6) $\leq$ 1 yr
Bandwidth = 2						
Linear	-16.2	-3.0	5.6	-4.9	-8.5	-2.5
	(13.5)	(5.8)	(8.0)	(5.5)	(12.5)	(7.6)
Quadratic	-25.8	2.7	4.5	0.2	-6.1	-4.6
	(19.0)	(8.2)	(11.9)	(7.6)	(16.6)	(10.8)
Obs.	136	136	136	136	136	136
No. of countries	136	136	136	136	136	136
Bandwidth = 1						
Linear	-18.9	1.3	5.7	-2.0	-3.2	-3.7
	(17.7)	(7.6)	(10.8)	(7.2)	(16.0)	(10.1)
Quadratic	-33.3	12.8	-7.5	13.9	-34.7	-18.7
	(24.5)	(10.5)	(14.5)	(9.9)	(20.9)	(13.9)
Obs.	81	81	81	81	81	81
No. of countries	81	81	81	81	81	81
Bandwidth = 0.5						
Linear	-43.0*	12.5	-9.8	12.6	-31.9	-16.4
	(23.9)	(10.3)	(13.0)	(9.7)	(20.5)	(13.7)
Quadratic	-32.7	16.0	-33.4	7.7	-11.0	-6.2
	(30.5)	(11.9)	(23.1)	(12.1)	(26.3)	(18.5)
Obs.	35	35	35	35	35	35
No. of countries	35	35	35	35	35	35

Note: See equation 3 in the text. The sample includes data points from 2000, before Gavi began disbursement in 2001. Each coefficient in the table (i.e., each row and each column) represents a separate regression. Each column uses a different dependent variable, listed in the column heading. Each row uses a different set of controls (i.e., polynomials of GNI and interactions with the Gavi treatment dummy). Standard errors are clustered at the country level.

Table 4: Impact estimates: intent-to-treat effect (ITT)

	Older Vaccines		Newer Vaccines			Placebo
	(1) Hep B	(2) DPT	(3) Hib	(4) Rota	(5) Pneumo	(6) Measles
Bandwidth = 2						
Linear	-8.4	-4.3	4.0	-5.3	2.4	-5.3
	(5.5)	(3.0)	(8.2)	(6.1)	(6.3)	(3.5)
Quadratic	-4.4	0.8	10.9	-0.09	-3.2	0.3
	(6.4)	(3.9)	(12.2)	(6.8)	(8.6)	(4.7)
Obs.	1755	1755	1755	803	667	1755
No. of countries	149	149	149	147	147	149
Bandwidth = 1						
Linear	-0.4	1.4	10.2	1.7	4.3	-0.05
	(5.6)	(3.6)	(11.4)	(6.4)	(7.4)	(4.5)
Quadratic	-7.7	3.4	0.7	8.8	-6.5	6.2
	(8.8)	(5.0)	(16.8)	(9.8)	(11.5)	(5.3)
Obs.	1011	1011	1011	444	363	1011
No. of countries	101	101	101	96	96	101
Bandwidth = 0.5						
Linear	-5.3	4.9	-2.9	-4.3	-8.5	7.1
	(9.0)	(5.5)	(16.7)	(8.8)	(10.3)	(5.7)
Quadratic	5.5	11.3	-28.5	-1.2	-21.1	8.4
	(13.4)	(7.1)	(22.7)	(27.7)	(18.4)	(6.9)
Obs.	435	435	435	194	161	435
No. of countries	64	64	64	53	53	64

Note: See equation 3 in the text. Each coefficient in the table (i.e., each row and each column) represents the ITT treatment effect from a separate regression. Our sample contains data from the year Gavi began disbursing financing for each vaccine up to 2013. Each column uses a different dependent variable, listed in the column heading. Each row uses a different set of controls (i.e., polynomials of GNI and interactions with the Gavi treatment dummy). Standard errors are clustered at the country level.

Table 5: Impact estimates: intent-to-treat effect (ITT), controlling for baseline vaccination rate

	Older Vaccines		Newer Vaccines			Placebo
	(1) Hep B	(2) DPT	(3) Hib	(4) Rota	(5) Pneumo	(6) Measles
Bandwidth = 2						
Linear	-0.5	-0.1	1.4	-5.3	2.4	-0.8
	(5.2)	(2.1)	(8.0)	(6.1)	(6.3)	(2.6)
Quadratic	5.8	4.0	9.8	-0.09	-3.2	4.1
	(6.5)	(3.0)	(11.9)	(6.8)	(8.6)	(3.6)
Obs.	1738	1738	1738	803	667	1738
No. of countries	144	144	144	147	147	144
Bandwidth = 1						
Linear	7.2	4.0	4.5	1.7	4.3	3.5
	(5.9)	(2.6)	(10.8)	(6.4)	(7.4)	(3.4)
Quadratic	2.0	2.5	-0.7	8.8	-6.5	5.4
	(8.1)	(4.0)	(16.0)	(9.8)	(11.5)	(4.5)
Obs.	1002	1002	1002	444	363	1002
No. of countries	99	99	99	96	96	99
Bandwidth = 0.5						
Linear	3.7	3.7	-0.5	-4.3	-8.5	5.9
	(9.3)	(4.1)	(15.6)	(8.8)	(10.3)	(4.4)
Quadratic	9.7	3.5	-20.5	-1.2	-21.1	3.3
	(12.7)	(5.7)	(20.7)	(27.7)	(18.4)	(6.3)
Obs.	432	432	432	194	161	432
No. of countries	62	62	62	53	53	62

Note: See equation 3 in the text. Each coefficient in the table (i.e., each row and each column) represents the ITT treatment effect from a separate regression, with baseline vaccination rate added as a control. Our sample contains data from the year Gavi began disbursing financing for each vaccine up to 2013. Each column uses a different dependent variable, listed in the column heading. Each row uses a different set of controls (i.e., polynomials of GNI and interactions with the Gavi treatment dummy). Standard errors are clustered at the country level.

Table 6: Impact estimates: mortality rates and spending

	Mortality Rate		Spending ($)	
	(1) Under-5	(2) Infant	(3) Total	(4) Gov't
Bandwidth = 2				
Linear	9.7	8.0	-0.06	0.2
	(8.2)	(5.3)	(0.4)	(0.2)
Quadratic	14.7	9.8	0.5	-0.1
	(10.9)	(7.2)	(0.5)	(0.2)
Obs.	404	404	795	823
No. of countries	147	147	136	138
Bandwidth = 1				
Linear	8.5	6.4	1.2	-0.04
	(9.6)	(6.2)	(1.0)	(0.2)
Quadratic	-9.3	-5.6	-1.4	0.2
	(12.1)	(8.6)	(1.3)	(0.2)
Obs.	228	228	455	476
No. of countries	98	98	87	90
Bandwidth = 0.5				
Linear	-16.0	-9.2	0.05	0.02
	(10.8)	(8.0)	(0.2)	(0.2)
Quadratic	-5.2	-3.2	0.1	0.03
	(19.2)	(13.3)	(0.5)	(0.4)
Obs.	101	101	211	213
No. of countries	63	63	50	50

Note: See equation 3 in the text. Each coefficient in the table represents the ITT treatment effect from a separate regression. Mortality estimates are available for 2005, 2010 and 2012 and vaccine finance information is available annually from 2006 to 2013. Each column uses a different dependent variable, listed in the column heading. Each row uses a different set of controls (i.e., polynomials of GNI and interactions with the Gavi treatment dummy). Standard errors are clustered at the country level.

Table 7: Impact estimates: effect of treatment on the treated (TOT)

	Older Vaccines		Newer Vaccines			Placebo
	(1) Hep B	(2) DPT	(3) Hib	(4) Rota	(5) Pneumo	(6) Measles
Bandwidth = 2						
Linear	-13.2	-6.9	5.7	-8.4	0.6	-8.6
	(9.2)	(5.0)	(13.4)	(11.4)	(13.1)	(5.9)
Quadratic	-7.0	2.0	19.6	7.4	-6.8	1.2
	(12.1)	(7.1)	(21.7)	(14.2)	(21.8)	(8.4)
Obs.	1755	1755	1755	803	667	1755
No. of countries	149	149	149	147	147	149
Bandwidth = 1						
Linear	0.9	2.7	16.0	7.3	6.6	0.02
	(9.9)	(6.2)	(19.1)	(13.0)	(18.9)	(7.7)
Quadratic	-23.7	7.2	-6.7	45.5	-209.7	12.8
	(26.3)	(11.1)	(40.9)	(66.2)	(1323.2)	(12.0)
Obs.	1011	1011	1011	444	363	1011
No. of countries	101	101	101	96	96	101
Bandwidth = 0.5						
Linear	-11.2	8.2	-7.3	-4.5	-28.0	12.0
	(17.4)	(9.4)	(31.3)	(21.5)	(40.6)	(9.6)
Quadratic	5.9	14.9**	-35.6	15.5	-46.3	11.7
	(17.0)	(7.2)	(42.5)	(63.5)	(50.7)	(8.0)
Obs.	435	435	435	194	161	435
No. of countries	64	64	64	53	53	64

Note: See equation 4 in the text. Each coefficient in the table (i.e., each row and each column) represents the "treatment on the treated" (TOT) effect from a separate 2SLS regression, aka' fuzzy RD model. Our sample contains data from the year Gavi began disbursing financing for each vaccine up to 2013. Each column uses a different dependent variable, listed in the column heading. Each row uses a different set of controls (i.e., polynomials of GNI and interactions with the Gavi treatment dummy). The Gavi treatment dummy and its interactions with log baseline GNI are instrumented with the Gavi *eligibility* dummy and its equivalent interactions. Standard errors are clustered at the country level.

Table 8: Impact estimates: dose-response model

	Older Vaccines		Newer Vaccines			Placebo
	(1) Hep B	(2) DPT	(3) Hib	(4) Rota	(5) Pneumo	(6) Measles
Bandwidth = 2						
Linear	-5.5	-15.9	51.2	73.6	-8.9	45.1
	(12.4)	(14.2)	(53.6)	(63.2)	(118.4)	(681.1)
Quadratic	4.1	25.3	110.1	126.3	47.5	12470.7
	(34.5)	(38.9)	(147.1)	(117.6)	(114.0)	(14353.0)
Obs.	1708	1708	1708	763	632	1708
No. of countries	145	145	145	137	137	145
Bandwidth = 1						
Linear	3.6	10.9	60.7	122.2	-38.2	105.8
	(15.5)	(19.5)	(63.5)	(75.7)	(87.4)	(1132.8)
Quadratic	-7.7	-23.8	296.4*	99.1**	-151.0	1057.7
	(29.2)	(44.5)	(165.6)	(48.7)	(785.3)	(2572.5)
Obs.	996	996	996	431	352	996
No. of countries	100	100	100	93	93	100
Bandwidth = 0.5						
Linear	-19.5	-330.9	-193.7	127.4***	67.9	4270.3
	(28.3)	(3259.7)	(430.0)	(34.0)	(51.3)	(6066.6)
Quadratic	1.1	52.0	-193.0	-71.2	-430.3	91204.0
	(21.1)	(59.4)	(285.0)	(404.1)	(496.9)	(106201.7)
Obs.	427	427	427	187	155	427
No. of countries	63	63	63	51	51	63

Note: See equation 6 in the text. Each coefficient in the table (i.e., each row and each column) represents the "treatment on the treated" (TOT) effect from a separate 2SLS regression, aka fuzzy RD model. Our sample contains data from the year Gavi began disbursing financing for each vaccine up to 2013. Each column uses a different dependent variable, listed in the column heading. Each row uses a different set of controls (i.e., polynomials of GNI and interactions with the Gavi treatment dummy). Standard errors are clustered at the country level.

Table 9: Testing for fungibility and waste

	Older Vaccines		Newer Vaccines			Placebo
	(1) Hep B	(2) DPT	(3) Hib	(4) Rota	(5) Pneumo	(6) Measles
Bandwidth = 2						
Linear	-90.7***	-38.0***	-17.2	-5.2	-3.2	-8.9
	(12.8)	(11.2)	(12.3)	(10.5)	(8.6)	(5.9)
Quadratic	-65.3***	-20.5	-1.0	8.7	-6.3	1.4
	(21.0)	(15.6)	(19.8)	(13.2)	(9.1)	(8.4)
Obs.	1755	1755	1755	803	667	1737
No. of countries	149	149	149	147	147	149
Bandwidth = 1						
Linear	-59.0***	-31.0**	-9.7	10.4	5.0	0.3
	(18.2)	(14.2)	(16.8)	(12.6)	(8.0)	(7.7)
Quadratic	-93.9**	33.9	3.6	46.8	-35.8	13.2
	(42.8)	(31.0)	(40.4)	(69.4)	(252.7)	(11.9)
Obs.	1011	1011	1011	444	363	998
No. of countries	101	101	101	96	96	101
Bandwidth = 0.5						
Linear	-75.8**	22.7	-17.4	13.2	-6.0	11.9
	(32.5)	(18.6)	(29.4)	(15.8)	(9.3)	(9.6)
Quadratic	-62.7*	19.2	-33.4	-13.5	-1.7	11.7
	(37.2)	(16.6)	(50.3)	(87.9)	(11.5)	(8.0)
Obs.	435	435	435	194	161	431
No. of countries	64	64	64	53	53	64

Note: See equation 4 in the text. This table tests the claim that Gavi-delivered doses were additional and did not offset the number of doses supplied by other sources. Coefficients less than zero imply fungibility and/or waste. Each coefficient in the table (i.e., each row and each column) represents the "treatment on the treated" effect from a separate 2SLS regression. Each column uses a different dependent variable, listed in the column heading. Each row uses a different set of controls (i.e., polynomials of GNI and interactions with the Gavi treatment dummy). The Gavi treatment dummy and its interactions with log baseline GNI are instrumented with the Gavi *eligibility* dummy and its equivalent interactions. Standard errors are clustered at the country level.

Table 10: Testing for waste

	Older Vaccines		Newer Vaccines		
	(1) Hep B	(2) DPT	(3) Hib	(4) Rota	(5) Pneumo
Bandwidth = 2					
Linear	-13.6***	-0.005	-3.9**	-0.6	-0.7**
	(4.6)	(0.7)	(1.8)	(0.5)	(0.4)
Quadratic	-3.8	-0.05	-4.8*	-2.2	-1.7
	(7.7)	(0.5)	(2.6)	(1.9)	(1.1)
Obs.	1755	1755	1755	803	667
No. of countries	149	149	149	147	147
Bandwidth = 1					
Linear	-3.0	-0.2	-5.3**	-1.8	-2.0**
	(7.6)	(0.3)	(2.3)	(1.5)	(1.0)
Quadratic	-18.7	1.3	-5.1	-9.2	-21.7
	(13.4)	(1.4)	(4.1)	(11.6)	(122.9)
Obs.	1011	1011	1011	444	363
No. of countries	101	101	101	96	96
Bandwidth = 0.5					
Linear	-11.7	1.4	-5.1	-3.9	-3.6
	(9.9)	(1.3)	(3.6)	(3.3)	(2.9)
Quadratic	-13.4	-0.2	-4.6	-3.3	-0.4
	(10.5)	(0.6)	(3.9)	(3.8)	(2.8)
Obs.	435	435	435	194	161
No. of countries	64	64	64	53	53

Note: See equation 4 in the text. This table tests the null hypothesis that all Gavi aid translated directly into a marginal increase in vaccination rates. Coefficients less than zero imply waste. Each coefficient in the table (i.e., each row and each column) represents the "treatment on the treated" effect from a separate regression. Each column uses a different dependent variable, listed in the column heading. Each row uses a different set of controls (i.e., polynomials of GNI and interactions with the Gavi treatment dummy). Standard errors are clustered at the country level.

Table 11: Decomposition: Where did subsidized vaccines go?

	Older Vaccines		Newer Vaccines			Placebo
	HepB	DPT	Hib	Rota	Pneumo	Measles
(A) Gavi deliveries	66.2	40.4	34.5	4.4	4.9	0.43
	(11.21)	(8.60)	(6.94)	(1.62)	(2.20)	(1.34)
≈ (B) Gavi impact	0.9	2.7	16.0	1.7	1.3	0.02
	(9.9)	(6.2)	(19.1)	(4.7)	(4.7)	(7.7)
+ (C) Fungibility	56.0	30.8	4.4	0	0	0
	(18.2)	(14.2)	(16.8)	(.)	(.)	(.)
+ (D) Waste	3.0	0.2	5.3	2.1	1.2	0
	(7.6)	(0.3)	(23)	(1.8)	(0.8)	(.)

Note: This table presents estimates of vaccine deliveries, impact, waste and fungibility at the eligibility threshold. Gavi deliveries are an estimate of the percent of the population under 1 provided with a full series of doses by Gavi for a country at the eligibility threshold. Estimates of impact attributable to Gavi, fungibility and waste are drawn from the linear models, bandwidth 1, in Table 7, Table 9 and Table 10, respectively. Estimates of fungibility and waste are capped at 0.