Barcelona School of Economics

Master's in International Trade, Finance, and Development

Crowd In or Crowd Out? PEPFAR's Impact on Domestic HIV/AIDS Expenditures

Authors: Carley Clontz, Arielle Lang, Yavuz Milar, Kevyn O'Connell
Supervising Professors: Fernando Broner and Geert Mesters

7 June 2025

Abstract: This paper examines how PEPFAR foreign aid influences domestic HIV/AIDS expenditure in recipient countries. Using annual data from 188 countries between 2000 and 2015, it applies two-way fixed effects (TWFE) and nonparametric difference-in-differences models to assess fiscal responses to aid intensity. TWFE estimates suggest significant crowd out effects, but further analysis reveals that this effect is neither universal nor linear. Crowd out is concentrated in focus countries based in Sub-Saharan Africa, and primarily occurs at moderate aid intensities. Results show that cutting PEPFAR aid jeopardizes the program's 2030 transition to country-led responses. Methodologically, they reveal TWFE's limitations under heterogeneous treatment effects and stress regional and dose-level differences in aid impact.

Resumen: Este documento analiza cómo la ayuda externa de PEPFAR influye en el gasto nacional en VIH/SIDA en los países receptores. Utilizando datos anuales de 188 países entre 2000 y 2015, se aplican modelos de efectos fijos bidireccionales (TWFE) y diferencias en diferencias no paramétricas para evaluar las respuestas fiscales según la intensidad de la ayuda. Las estimaciones TWFE sugieren efectos significativos de desplazamiento (crowd-out), pero un análisis más detallado revela que dicho efecto no es ni universal ni lineal. El crowd-out se concentra en los países focales de África subsahariana y ocurre principalmente a intensidades moderadas de ayuda. Los resultados muestran que reducir la ayuda de PEPFAR pone en riesgo la transición del programa hacia un liderazgo nacional para 2030. Metodológicamente, estos hallazgos ponen de manifiesto las limitaciones del enfoque TWFE ante efectos de tratamiento heterogéneos y subrayan las diferencias regionales y de "dosis" en el impacto de la ayuda.

Keywords: *PEPFAR, foreign aid, crowd in / crowd out*

Keywords: *PEPFAR, ayuda extranjera, atracción de inversión / desplazamiento en el gasto doméstico*

Contents

1 Introduction and Motivation

The U.S. President's Emergency Plan for AIDS Relief (PEPFAR) is the largest commitment by any nation to combat a single disease in history. Since its launch in 2004, it has disbursed over $110 billion globally, with funding administered primarily through the U.S. Agency for International Development (USAID) and the Centers for Disease Control and Prevention (CDC). It is focused on HIV/AIDS prevention, treatment, and health-systems strengthening. Evaluations have consistently shown that PEPFAR has produced substantial health benefits and declines in HIV prevalence: from 2004 to 2018, all-cause mortality declined by 20% in supported countries (KFF 2021), while between 2010 and 2023, new HIV infections fell by 52% and AIDS-related deaths by 59% (HIV.gov 2024). Recent research also links PEPFAR to improved economic and educational outcomes (Crown 2023).

Despite this progress, little is known about how PEPFAR affects the broader HIV/AIDS financing landscape, and specifically, whether its funding displaces or complements other forms of HIV/AIDS spending. This question has gained urgency amid growing uncertainty about U.S. foreign assistance. In 2023, PEPFAR's five-year reauthorization was delayed for the first time since its inception (KFF 2025), and then in early 2025, the Trump administration began a comprehensive restructuring of U.S. foreign aid by dismantling USAID. This move coincided with Executive Order 14169, which imposed a 90-day freeze on all foreign assistance. Although a limited waiver allowed some critical PEPFAR services to resume temporarily, many prevention programs remain suspended, exacerbating uncertainty regarding the program's future. These developments come as PEPFAR pledged to enter a new strategic phase to help end the global HIV/AIDS epidemic by 2030 and to transition to sustainable, country-led responses (U.S. Department of State 2024; HIV.gov 2024). Whether that transition is viable depends in part on how recipient governments have historically responded to PEPFAR funding. If PEPFAR has catalyzed domestic investment, countries may be positioned to lead, but if it has displaced domestic and private spending, future funding cuts could leave major gaps. Understanding the extent to which PEPFAR crowds in or crowds out other HIV/AIDS expenditures is therefore key to assessing the risks posed by donor withdrawal and to informing sustainability planning and the trajectory of the global HIV/AIDS epidemic holistically.

This study contributes to this goal by answering the key question: ***To what extent do PEPFAR aid flows crowd in or crowd out other types of expenditures on HIV/AIDS in recipient countries?*** To answer this question, this report examines how changes in PEPFAR aid flows (as a percent of GDP) affect domestic government and private HIV/AIDS spending (also as a percent of GDP) across countries over time. This analysis focuses on government expenditure, as it is the largest contributor to domestic spend on HIV/AIDS, as shown in the subsequent Section 4. *(Data).* Private spending will be considered as a secondary metric. Using panel data on U.S. HIV/AIDS assistance from ForeignAssistance.gov and domestic HIV/AIDS spending from the Institute for Health Metrics and Evaluation (IHME), this study applies the continuous differences-in-differences (DiD) framework proposed by Callaway, Goodman-Bacon, and Sant'Anna (2024), and exploits annual variation in aid intensity to estimate recipient governments' fiscal responses. In the study period (2000-2003 pre-period, 2004-2015 post), two-way fixed effects (TWFE) estimates suggest significant crowd out effects, but further analysis reveals that this effect is neither universal nor linear. Crowd out is concentrated in focus countries based in Sub-Saharan Africa, and primarily occurs at moderate aid intensities. The results underscore that reducing or halting PEPFAR funding could undermine the program's transition to country-led ownership and jeopardize progress against the HIV/AIDS epidemic. Econometrically, they reveal TWFE's limitations under heterogeneous treatment effects and stress the nuance of regional and dose-level differences in the impact of aid on recipient countries' budget decisions.

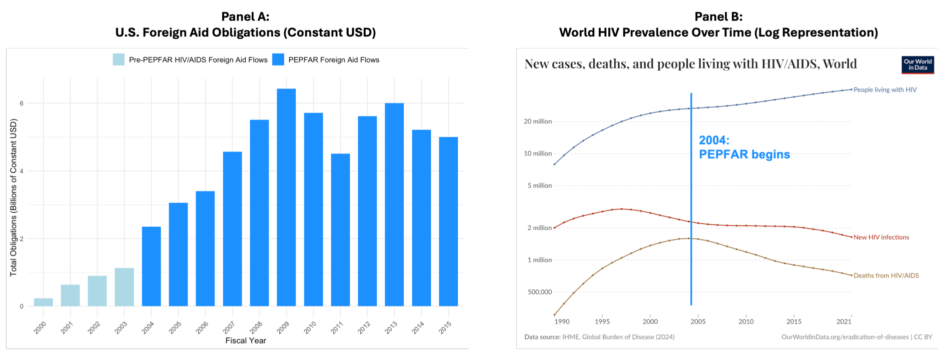
The document proceeds as follows. First, it provides background on PEPFAR (2) and a review of the relevant literature (3). Next, it describes the data and how the key independent variable, outcome measures, and controls are constructed (4). It then outlines the empirical strategy (5) and presents the results (6). After that, it explores the policy implications of our findings (7) and discuss limitations and opportunities for future research (8). The report closes with a brief conclusion (9).

2 PEPFAR: Program Design and Policy Context

Launched in 2004 by President George W. Bush, PEPFAR has evolved into the world's largest bilateral health initiative, operating in almost 100 countries and investing over $110 billion to date (HIV.gov 2024; U.S. Department of State 2024). Its early years were focused on emergency relief, rapidly scaling up antiretroviral therapy (ART), prevention of mother-to-child transmission, and infrastructure for HIV/AIDS testing and care. Notably, PEPFAR's launch was driven by political and ideological priorities within the U.S. administration, rather than by a contemporaneous global spike in HIV prevalence (AAAS Science & Diplomacy, 2013) (Figure 1). This makes its initial roll-out particularly well-suited to causal identification, as the scale-up represents a large, exogenous increase in HIV/AIDS-specific aid.

This report exploits the introduction of PEPFAR as an exogenous spending shock: as shown in Figure 1, U.S. foreign aid for HIV/AIDS more than doubled when the program launched in 2004, according to the United States Foreign Assistance database. This rapid increase created a natural experiment across countries, with large variations in exposure to new HIV/AIDS funding.

Figure 1: PEPFAR Program Exogenous Shock



PEPFAR funds are disbursed primarily through Country Operational Plans (COPs), which are developed annually in partnership with recipient governments. These plans outline national priorities, expected targets, and budgets for U.S. government implementing agencies such as USAID and the CDC (U.S. Department of State 2024). COPs are particularly concentrated in a set of 15 "focus countries", most of which are located in Sub-Saharan Africa (Table 1), which together receive over 90% of absolute global PEPFAR direct-to-country obligated funding per Foreign Assistance data. As noted by Over, Mead, and Glassman (2013), while there are some criteria for selecting focus countries such as HIV prevalence, health infrastructure, government capacity, and prior aid relationship with the U.S. government, the methodology to select focus countries was inconsistent, individualistic, and notably required significant judgment ("semi-random" selection). This further confirms the program's utility for empirical evaluation as it supports Assumption 1 (random sampling) in Section 5. (*Empirical Strategy*).

Table 1: Focus Countries and Regions

Focus Country	**Region**	**Focus Country**	**Region**
Botswana	Sub-Saharan Africa	Kenya	Sub-Saharan Africa
Côte d'Ivoire	Sub-Saharan Africa	Mozambique	Sub-Saharan Africa
Ethiopia	Sub-Saharan Africa	Namibia	Sub-Saharan Africa
Guyana	Latin America & Caribbean	Nigeria	Sub-Saharan Africa
Haiti	Latin America & Caribbean	Rwanda	Sub-Saharan Africa
South Africa	Sub-Saharan Africa	Tanzania	Sub-Saharan Africa
Uganda	Sub-Saharan Africa	Vietnam	East Asia & Pacific
Zambia	Sub-Saharan Africa		

Funding peaked in the late 2000s and has since stabilized around \$6-7 billion annually (Figure 1) with modest fluctuations driven by Congressional appropriations (ForeignAssistance.gov 2024). More recently, PEPFAR has integrated goals around country ownership, emphasizing the need for recipient governments to transition toward financing and managing their own HIV/AIDS responses (U.S. Department of State 2024). However, the political and fiscal feasibility of this shift remains uncertain, especially given the growing domestic pressures in donor countries and the fragility of health financing systems in many recipient contexts.

3 Literature Review

3.1 PEPFAR Outcomes and Empirical Evaluations

The positive health outcomes of PEPFAR are well understood and well documented. For example, a comprehensive evaluation by the Institute of Medicine (2013) found that PEPFAR had a transformative impact on health systems in supported countries, significantly expanding access to antiretroviral therapy, reducing mortality, and improving HIV prevention. By 2011, over 3.9 million people were receiving ART through PEPFAR, and HIV testing, counseling, and prevention services had reached tens of millions. The report also highlights improvements in health infrastructure, workforce training, and supply chain systems, emphasizing PEPFAR's contribution to broader health system strengthening. In addition to this comprehensive report, annual impact reports posted by the U.S. Department of State document extensive program monitoring and evaluation across countries and time. Extending beyond health outcomes, Crown et al. (2023) use country-level panel data from 1990–2018, and finds that PEPFAR is associated with increases in the GDP per capita growth rate and lower school disengagement.

However, few econometric studies examine how PEPFAR influences fiscal behavior; that is, whether recipient governments increase, reduce, or maintain their own HIV/AIDS spending in response. This is a critical gap given PEPFAR's strategic goal of financial transition and co-financing. The current study adds to the literature by explicitly focusing on fiscal substitution and complementarity within the HIV/AIDS budget envelope.

3.2 Aid and Fiscal Behavior: Theoretical and Cross-Sector Evidence

The standard fiscal response model, first articulated by Heller (1975) and later refined by Franco-Rodriguez et al. (1998), conceptualizes external aid as an additional revenue stream that recipient governments treat much like any other source of public funds when making utility-maximizing budget decisions. In this framework, aid effectively lowers the marginal cost of financing public investments: once a donor provides resources earmarked for a particular sector (e.g., health), the government's own budgetary contribution to that sector can be scaled back, freeing up domestic funds for other uses. In practice, such reallocation may reflect a combination of political incentives, administrative or capacity constraints, or a deliberate strategy to shift budget priorities in response to evolving constituency demands. This is referred to as aid fungibility: although donors target a specific sector, recipient governments can reduce their own spending in that area and redeploy the freed-up funds elsewhere.

Early empirical investigations provide stark evidence of this displacement. Using cross-country data on sector-specific grants, Feyzioglu et al. (1998) estimates that only about two cents of every targeted aid dollar to health, education, or agriculture actually translate into additional domestic outlays in the same sector. Subsequent work by Lu et al. (2010) focuses specifically on development assistance for health (DAH) and found that when DAH is channeled through government budgets, domestic health expenditures fall by an average of forty-three cents for every dollar disbursed. These findings underscore the risk that well-intentioned, sector-targeted aid might simply substitute for domestic resources rather than add to them.

Beyond sectoral displacement, an important behavioral indicator in this framework is tax effort, or the ratio of tax revenue to GDP, which captures how vigorously a government mobilizes its own resources. Remmer (2004) finds that, in a panel of developing countries, higher aid inflows coincide with larger government size but lower tax effort, suggesting that aid functions like intergovernmental transfers by spurring spending growth without prompting additional domestic revenue mobilization. However, more recent work complicates this picture. Carter (2013), using panel data and instrumental-variable methods, detects no systematic reduction in tax effort following aid inflows, while Morrissey and Torrance (2014) show that aid can actually boost tax effort in countries with stronger institutions, particularly when aid is predictable and aligned with recipient priorities. These results imply that the fiscal response to aid depends critically on governance quality, aid modality, and donor–recipient coordination. As such, program-specific analyses are key to understand the nuanced impacts of aid on domestic fiscal capacity.

3.3 Sector-Specific and Health Aid Effects

In the health sector, evidence on aid's fiscal effects is mixed and context-dependent. For example, Liang and Mirelman (2014) employ a two-stage least squares (2SLS) panel-data approach with country fixed effects and to identify the impact of DAH on domestically financed government health expenditure (DGHE) across 120 countries from 1995–2010. They find that DAH is partially fungible: a one percent increase in DAH is associated with a 0.03–0.04 percent decrease in DGHE. Moreover, fungibility is more pronounced in countries with higher corruption or ethnic tensions.

On the contrary, Ouattara (2006) uses a panel data approach across aid-recipient countries and shows that foreign aid, particularly in developmental sectors like health and education, is positively associated with public investment, without reducing tax effort. Their findings support a crowd in dynamic, particularly when aid is targeted to infrastructure and services. This suggests that, under certain conditions, aid can enhance rather than replace domestic budgetary commitments.

At the sub-national level, Lu, Cook, and Desmond (2017) conduct a facility-based analysis in Rwanda and provide further support for crowding in. They find that aid provided to rural health centers was positively associated with increased government co-financing and expanded service delivery. Their findings highlight the importance of aid design and delivery mechanisms, particularly when funds are channeled directly to providers and aligned with local planning processes.

Despite these contributions, there remains insufficient evidence on how disease-specific or program-targeted aid, such as PEPFAR, affects corresponding spending from governments and the private sector. Much of the health aid literature aggregates across conditions or relies on sector-wide expenditure data, making it difficult to assess whether large vertical programs reinforce or displace other funding. By isolating PEPFAR's effects within the HIV/AIDS financing landscape, this study helps fill an important gap and provides evidence relevant to ongoing debates about aid sustainability and transition planning.

4 Data

This analysis uses a novel country-year panel drawing on three primary data sources: (1) all obligated U.S. foreign assistance for HIV/AIDS programming; (2) domestic HIV/AIDS expenditure estimates, disaggregated by funding source; and (3) an array of macroeconomic and health sector control variables. To facilitate meaningful comparisons both across countries and over time, spending variables are normalized by each country's GDP and values are converted into 2017 U.S. Dollars (USD) that are adjusted for purchasing-power parity (PPP). This adjustment is crucial: expressing expenditures in constant 2017 PPP USD removes distortions from domestic price-level differences and inflation, allowing a comparison between the real resource commitment to HIV/AIDS where a "dollar" buys the same bundle of goods and services everywhere. The final panel spans 188 countries over 16 years (2000–2015), yielding 2,928 total observations. Of these, 98 countries received PEPFAR funding in at least one year from 2004 onward.

4.1 U.S. Foreign Aid: HIV/AIDS and PEPFAR Flows

This report's primary measure of U.S. HIV/AIDS assistance is drawn from the transaction-level U.S. Foreign Assistance database (ForeignAssistance.gov), which records every foreign aid obligation and disbursement from 1947 until early 2025. Each transaction includes key information such as fiscal year and quarter, agency and funding account (e.g., USAID), activity name/description (with "PEPFAR" and/or "HIV/AIDS" identifiers), recipient country and implementing partner, transaction type (obligation versus disbursement), and transaction amount in both current and constant U.S. dollars. To treat the data, the analysis (1) filters transactions by activity description to isolate PEPFAR HIV/AIDS funding (obligations in current USD); (2) restricts the sample to 2000–2015 to match the domestic expenditure data (Section 4.2.); (3) groups all transactions for the same recipient country-year combination; (4) converts each country-year total into 2017 PPP-adjusted USD; (5) normalizes by recipient GDP, sourced from the World Bank (2017 PPP USD) and multiplies by 100 to yield a percentage-point measure of aid intensity. This harmonized aid flow variable is the main treatment in all subsequent analyses. The descriptive statistics in Figure A.1 in the Appendix highlights the large variation in aid flows by country, which directly informs this report's empirical strategy discussed in Section 5. *(Empirical Strategy)*. Additionally, Table 2 highlights that most of the funding is going to Sub-Saharan Africa each year, a critical aspect to this report's findings.

Table 2: Percentage of PEPFAR Funds Received by Each Region (2004–2015)

Year	Asia	Europe	MENA	South & Central Asia	Sub-Saharan Africa	LAC
2004	8.2	3.2	1.0	3.9	75.6	8.0
2005	7.6	2.4	0.5	2.2	79.2	8.1
2006	6.2	2.0	0.8	1.8	81.6	7.7
2007	4.6	1.9	0.4	2.0	80.9	10.2
2008	5.9	1.6	0.4	2.5	82.0	7.7
2009	4.2	1.3	0.5	1.5	84.6	7.8
2010	4.4	0.9	0.4	2.9	85.1	6.2
2011	4.4	0.6	0.0	2.0	88.8	4.2
2012	3.7	0.7	0.6	2.3	88.9	3.8
2013	3.3	0.5	0.2	1.4	89.6	5.0
2014	2.3	0.5	0.4	2.0	90.9	4.0
2015	2.2	0.6	0.1	1.1	91.5	4.6

4.2 Domestic HIV/AIDS Expenditure

To understand spending behavior on HIV/AIDS in recipient countries, this analysis uses the dataset "Global HIV/AIDS Spending 2000-2015" constructed by the Institute for Health Metrics and Evaluation (IHME), a population health research organization based at the University of Washington School of Medicine in the United States. The dataset was released in 2018 and contains country-year domestic expenditures on HIV/AIDS for 188 countries between 2000-2015. To estimate domestic expenditures, the team collected published data from 1995 to 2015 – including from online databases, country reports, and proposals submitted to multilateral organizations – and applied spatiotemporal Gaussian process regression to generate estimates.[1] The included data estimates of interest to this study are (1) government spending on HIV/AIDS, which is self explanatory; (2) prepaid private spending on HIV/AIDS, which refers to healthcare payments made in advance to private insurers or community health schemes (typically as insurance premiums or contributions); and (3) out-of-pocket spending on HIV/AIDS, or direct payments made by individuals at the time of receiving health care without reimbursement (such as co-payments, deductibles, and full fees). As shown in Table 3, the large majority of domestic spending is from the government each year. All data are recorded in 2017 PPP-adjusted $US, and as such, the only treatment that is applied in this analysis is normalizing these expenditures by recipient country GDP from the World Bank. See Table 4 for descriptive statistics by region, which highlights that, on average, Sub-Saharan Africa spends substantially more on HIV/AIDS as a percent of GDP than any other region.

[1]For a full description of the estimation methods, see Dieleman et al. (2018) in the References.

Table 3: Percent of Total Domestic HIV/AIDS Expenditure Paid by Each Source by Year

Year	Out-of-Pocket (%)	Prepaid Private (%)	Public (%)
2000	16.6	6.1	77.4
2001	16.3	6.1	77.6
2002	16.1	6.1	77.7
2003	15.7	6.0	78.3
2004	15.1	5.8	79.1
2005	14.3	5.6	80.1
2006	13.3	5.4	81.3
2007	12.6	5.3	82.0
2008	11.9	5.1	83.0
2009	11.4	4.9	83.7
2010	10.8	4.7	84.5
2011	10.3	4.7	85.0
2012	9.8	4.6	85.5
2013	9.3	4.4	86.2
2014	8.9	4.5	86.6
2015	8.7	4.6	86.7

Table 4: Government Expenditure on HIV/AIDS (as a Percent of GDP) by Region (2004–2015 aggregated)

Region	Mean_Percent	Median_Percent	Min_Percent	Max_Percent
East Asia & Pacific	0.04	0.02	0.00	0.21
Europe & Central Asia	0.07	0.04	0.01	0.46
Latin America & Caribbean	0.10	0.09	0.01	0.40
Middle East & North Africa	0.02	0.01	0.00	0.10
South Asia	0.02	0.01	0.01	0.04
Sub-Saharan Africa	0.41	0.13	0.00	4.33

4.3 Control Variables

In addition to the GDP data described in Sections 4.1 and 4.2, the analysis incorporates annual country-level controls from the World Bank's Development (WDI) and Governance (WGI) Indicators: life expectancy at birth; HIV prevalence (percent of population); private and public health expenditure (percent of GDP); and control of corruption (WGI percentile rank, where higher values indicate stronger control). For one additional control, the analysis uses IHME's DAH Database to source non-U.S. development assistance for HIV/AIDS, and also normalize this by GDP and convert to 2017 PPP-adjusted USD. Country names were harmonized across sources, and missing values were handled using a combination of forward- and bi-directional filling (within units), log-linear extrapolation (e.g., for South Sudan's GDP pre-2008), and regional-year imputation where appropriate (e.g., for health spending in South Sudan and Venezuela). Finally, countries lacking sufficient WDI or IHME coverage, most notably North Korea and Taiwan, were dropped from the analysis entirely.

5 Empirical Strategy

5.1 Base Model: Two-Way Fixed Effects (TWFE)

To assess whether PEPFAR aid crowds in or crowds out domestic HIV/AIDS spending, the baseline model uses PEPFAR aid as a percent of GDP as the key independent variable and domestic HIV/AIDS expenditure as a percent of GDP as the outcome. The model leverages 2000–2003 as the pre-PEPFAR period and 2004–2015 as the post period.

The TWFE estimation strategy begins with a continuous DiD methodology, which exploits variation in aid intensity across countries and over time. This method matches the needs of this empirical environment since it has considerable variation in treatment intensity (see Figure A.1 in the Appendix). This method also allows for time-invariant country characteristics and common shocks to be absorbed. The model is specified as:

$$Y_{it} = \alpha_i + \lambda_t + \beta^{twfe}(D_{it} * POST_t) + \gamma^T X_{it} + \epsilon_{it} \tag{1}$$

The interaction of $D_{it} * POST_t$ captures how changes in aid intensity after the introduction of PEPFAR affect domestic spending on HIV/AIDS. The coefficient $\hat{\beta}^{twfe}$ is the parameter of interest, representing the marginal effect of an increase in PEPFAR aid as a percent of GDP on the domestic expenditure outcome during the post-PEPFAR period.

Variables are defined as:

- Y_{it} = Country-year domestic government (or private) expenditure on HIV/AIDS (as a percent of GDP)
- α_i = Country fixed effects
- λ_t = Year fixed effects
- D_{it} = PEPFAR obligations (as a percent of GDP)
- POST$_t$ = 1 if $\{t \geq 2004\}$
- X_{it} = Matrix of control variables:
  - D_{it}*:* Separate from interaction
  - $POST_t$: Separate from interaction (dropped due to collinearity)
  - *Life expectancy:* Life expectancy at birth (years)
  - *HIV prevalence:* (as a percent of population)
  - *Corruption:* Perception of control of corruption (percentile rank; closer to 100 = less corrupt)
  - *Current health expenditure:* Private health expenditure (as a percent of GDP)
  - *Government health expenditure:* Public health expenditure (as a percent of GDP)
  - *Non-US DAH:* (as a percent of GDP)
- ε_{it} = Error term, clustered by country

5.1.1 Assumptions

This report relies on the following assumptions to use this methodology, which are outlined in the reference paper on continuous DiD, Callaway et al. (2024):

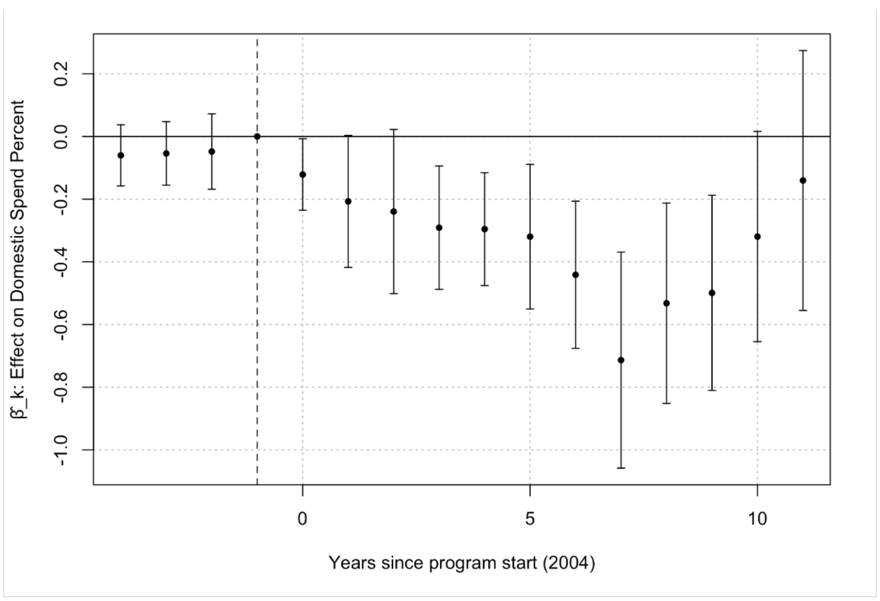
1. **Random sampling and exogeneity:** Quasi-random selection of PEPFAR focus countries and the program's exogenous roll-out support these assumptions (see Section 2. *(PEPFAR)*). To mitigate that PEPFAR country selection was not strictly randomized, the model includes time-varying controls (above) that capture observable determinants of both aid allocation and outcomes, thereby isolating the exogenous variation in foreign HIV assistance. This model also relies on the asssumption of country-level independence (no spillovers), meaning that aid to one country does not affect another, so observations can be treated as independent and identically distributed. As such, these assumptions are considered to be satisfied.

2. **Continuous treatment**: PEPFAR aid as a percent of GDP varies among countries and over time, as shown in Section 4. *(Data)*. While no country receives PEPFAR aid in the pre-PEPFAR period (i.e., $t = 1$), all countries begin receiving varying doses of treatment in the post-period ($t = 2$). The treatment support is continuous and strictly positive, with all values within a bounded range [3×10^{-9} and 2.01] (see Figure A.1). This satisfies the condition for a continuous or multi-valued discrete treatment, as required for identification under the continuous DiD framework.

3. **No anticipation and observed outcomes**: This assumption requires that countries do not alter their HIV/AIDS spending before PEPFAR takes effect. To account for this, this analysis strategically uses PEPFAR obligations (official funding announcements) as the treatment variable instead of disbursements, since obligations mark when governments could start adjusting budgets. Additionally, per Sridhar and Batniji (2008), a rapid scale-up of PEPFAR largely bypassed recipient budget-planning procedures. Consequently, anticipation effects can be assumed to be minimal, and pre-treatment outcomes can be interpreted as an accurate reflection of the counterfactual in the absence of PEPFAR.

4. **Parallel trends**: This assumption states that, in the years preceding PEPFAR's rollout, countries later assigned to different aid intensities follow parallel trajectories in their domestic HIV/AIDS spending. This means that, had PEPFAR never been introduced, each group would have continued on the same spending path. The event-study plot (Figure 2) supports this: pre-2004 TWFE estimates fluctuate narrowly around zero and their confidence intervals overlap, demonstrating no systematic differences in spending trends across eventual treatment groups. Additionally, because these pre-treatment coefficients are not statistically significant, there is evidence that any divergence in HIV/AIDS spending observed after 2004 can be attributed to PEPFAR rather than to pre-existing trends.

Figure 2: TWFE Event Study



5.1.2 Robustness Check: Binary Treatment

To test the robustness of this continuous DiD model, this analysis uses a binary treatment variable to re-estimate the outcome of equation (1). Specifically, the continuous PEPFAR intensity measure D_{it} is replaced with a dummy variable $Treated_{it}$, which equals 1 if a country received *any* PEPFAR aid in year t, and 0 otherwise. This yields the following specification:

$$Y_{it} = \alpha_i + \lambda_t + \beta^{binary}(Treated_{it} \times POST_t) + \gamma^\top X_{it} + \epsilon_{it} \tag{2}$$

Although this binary model masks variation in aid intensity, it provides a useful robustness check. The results are considered robust if the signs of $\hat{\beta}^{binary}$ are directionally consistent with those of $\hat{\beta}^{twfe}$, and the results are statistically significant. That is, even without information about how much aid was received, simply being classified as treated should move the outcome in the same direction. This binary framework offers a simpler test of whether *any* exposure to PEPFAR yields detectable changes in HIV/AIDS spending.

5.2 Non-Parametric Modeling

While the TWFE estimator remains a workhorse in policy evaluation, a growing strand of modern DiD literature, including Roth, Sant'Anna, Bilinski, and Poe (2023); de Chaisemartin and d'Haultfœuille (2023); Callaway (2023); and others, has highlighted various shortcomings of TWFE, particularly under treatment-effect heterogeneity or continuous treatments.

In line with these broader critiques, Callaway, Goodman-Bacon, and Sant'Anna (2024) show that TWFE can produce misleading or uninterpretable causal estimates when treatment intensity varies. The key limitations they identify are discussed below.

5.2.1 Limitation 1: Coefficient weighting

In much of the traditional empirical literature, including influential work by Acemoglu and Finkelstein (2008), the coefficient from a TWFE regression with continuous treatment is interpreted as a marginal effect. That is, it is taken to represent the change in the outcome associated with a one-unit increase in treatment intensity, conditional on fixed effects.

Recent work by Callaway et al. (2024), however, challenges this interpretation. They show that in a DiD context with continuous treatment, the TWFE estimator does not recover a true marginal effect. Instead, it captures a level effect: a weighted average of dose-specific treatment effects, $ATT(d \mid d)$, where the weights can be negative and do not necessarily sum to one.

This issue becomes clear for this case in Table 5. TWFE assigns negative weights to low-dose treated units and positive weights to high-dose units, which distorts the interpretation of the estimated coefficient. These patterns reinforce the concern that TWFE conflates comparisons across differently treated groups in ways that do not reflect the average causal effect of the true population. This provides motivation for adopting a more flexible, nonparametric approach that directly estimates $ATT(d \mid d)$ without imposing linearity or relying on misleading weighting schemes.

Table 5: Distribution of TWFE Weights by Dose Bin

Dose Bin of Aid Shares	**Total Weight**	**Avg. Weight**	**N**
(−0.00201, 0.201]	-1.260	-0.00120	1052
(0.201, 0.402]	0.149	0.00178	84
(0.402, 0.603]	0.235	0.00405	58
(0.603, 0.804]	0.305	0.00635	48
(0.804, 1.01]	0.252	0.00869	29
(1.01, 1.21]	0.121	0.0110	11
(1.21, 1.41]	0.107	0.0134	8
(1.41, 1.61]	0.0460	0.0153	3
(1.81, 2.01]	0.0424	0.0212	2

5.2.2 Limitation 2: Parallel trends is insufficient

Callaway et al. (2024) point out that while standard parallel trends are sufficient to identify level effects, or the average impact of moving from zero to some dose d, they do not allow us to infer how outcomes respond to marginal increases, such as going from d to $d + \Delta$. To recover the slope of the dose-response relationship, this methodology requires a stronger assumption: strong parallel trends (SPT).

SPT requires that, for each aid level d, the observed change in domestic HIV expenditure for countries that actually received d is exactly what every country would have experienced if it, too, had received d. In other words, there must be no correlation between a country's unobserved responsiveness to aid and the amount of aid it receives. This rules out "selection on gains", or the idea that countries receiving more aid may also be the most likely to respond strongly to any aid, regardless of dose.

In the context of this report, this assumption is likely violated. The allocation of PEPFAR aid intensity reflects a complex mix of factors, including epidemiological need, governance capacity, health system infrastructure, and geopolitical considerations. Many of these are plausibly correlated with a country's ability to translate external funding into domestic spending. Additionally, institutional effectiveness and absorptive capacity, both unobservable, can vary widely across countries, meaning two countries with similar aid levels might still respond differently. Since these traits cannot be observed or counterfactual spending paths constructed for each dose, SPT cannot be directly tested and may not hold.

Despite this, the analysis proceeds under the assumption that SPT is satisfied. Doing so allows the TWFE estimator to be interpreted at each aid level d as the marginal effect of an incremental increase in PEPFAR aid, that is, the slope of the dose-response curve. While this assumption is acknowledged to be strong, it enables the generation of policy-relevant insights regarding how small changes in aid intensity may influence domestic HIV spending.

5.3 Application of Non-Parametric Methods using B-Spline Transformations

A spline-based continuous DiD approach relaxes linearity assumptions, accommodates treatment effect heterogeneity across the aid distribution, and yields both level and marginal causal estimates. This provides a richer and more actionable understanding of how PEPFAR influences domestic HIV/AIDS expenditures across the full spectrum of aid intensity.

For context, a spline is a piecewise polynomial function that bends at a set of chosen knots (specific values of PEPFAR aid as a percent of GDP) where one polynomial segment transitions smoothly into the next. By placing these knots at meaningful thresholds, the spline flexibly captures two dynamics: (1) how overall changes in domestic HIV spending vary when PEPFAR aid is low, medium, or high, and (2) how the marginal effect of each additional percentage point of aid depends on a country's existing aid level.

To address the linear TWFE model's inability to uncover these nuanced, incremental patterns discussed in Section 5. *(Non-Parametric Modeling)*, this analysis adopts a cubic B-spline basis model. This approach treats aid intensity as a continuous dose and allows its impact on domestic HIV expenditure to vary nonlinearly across countries. Unlike natural splines, which constrain the tails of the curve to be straight beyond the boundary knots, B-splines retain full flexibility across the entire support of aid intensity. This is especially important for our use case, where we have a large portion of our data at the leftmost tail.

From the fitted spline, two key quantities are recovered:

- **Average Treatment Effect of the Treated,** $ATT(d \mid d)$**:** The average effect of receiving aid at dose d for countries that actually received that level of treatment. It reflects the difference in domestic HIV spending between the observed outcome and the counterfactual scenario in which the same countries received no aid. Under SPT, this is equivalent to the overall average treatment effect, or the $ATE(d)$.

- **Average Causal Response,** $ACR(d)$**:** The derivative of the $ATT(d \mid d)$ function with respect to d. It captures the marginal effect of a one percentage point increase in PEPFAR aid on domestic HIV spending, evaluated at aid level d. Plotting ACR(d) reveals whether the returns to additional aid are increasing, constant, or diminishing across the aid distribution.

In this sample of 98 PEPFAR recipient countries over 2000–2015, the spline-based method offers two practical advantages. First, it flexibly accommodates nonlinear responses in domestic HIV spending, allowing countries that receive very little aid to exhibit different spending patterns from those receiving substantial support.

Second, by estimating $ACR(d)$ directly, concrete policy questions can be addressed such as: *if one country's PEPFAR funding increases from one to two percent of GDP, how much additional domestic HIV spending should be expected?* This stands in contrast to a standard TWFE model, which, as recent critiques suggest, would only report the average effect of moving from zero to roughly one percent, for example, without offering insight into finer increments.

5.3.1 Nonparametric Estimation Strategy

To estimate level-specific effects, the "transformed outcome" in domestic HIV spending is regressed on a flexible spline in the PEPFAR aid share, D_i. To do this, this analysis adapts the two-period approach outlined in Callaway et al. (2024) and extends it to a multi-time period setting using the complementary set-up as outlined in Callaway et al. (2024). The general specification is described as follows[2]:

$$y_i = \psi^K(D_i)^T \beta_K + X_i^T \gamma + \epsilon_i \tag{3}$$

where

$$y_i = \Delta Y_i - E[\Delta Y_i \mid D_i = 0]$$

computes the change in the outcome relative to the pre-treatment period and subtracts the average change among untreated units. Here, $\psi^K(d) = (\psi_{K1}(d), \psi_{K2}(d), \ldots, \psi_{KK}(d)$ is a K-dimensional vector of B-spline basis functions in the treatment, $\beta_K = (\beta_{K1}, \beta_{K2}, \ldots, \beta_{KK})$ is a vector of finite-dimensional (unknown) parameters, X_i is a vector of control variables (e.g., life expectancy, HIV prevalence, corruption rank), and ϵ_i is an idiosyncratic error term. Each $\psi_{Kk}(D_i)$ is a smooth, piecewise polynomial supported over a localized interval of aid intensity; together, they form a continuous, differentiable system that bends wherever the data suggests nonlinearity. The number of basis functions, K, is selected using 5-fold cross-validation, where the sample is split into five parts, the model is trained on four and tested on the fifth in turn, and the configuration that yields the lowest average out-of-sample mean squared error is chosen. R's default placement of internal knots (roughly at the 25th, 50th, and 75th percentiles of D when $K = 6$[3]) ensures the spline adapts curvature where observations of D are dense, without imposing linear tails.

The coefficient vector $\hat{\beta}_{\hat{K}}$ will subsequently be the OLS estimated coefficient of Equation (3), which regresses the "transformed outcome" onto the K-dimensional B-spline $\psi^K(d)$ in the sub-sample of units that have received positive treatment dosage. Once $\hat{\beta}_{\hat{K}}$ is estimated, the fitted curve provides a smooth estimate of the expected change in domestic HIV spending at a given level of aid, or the average treatment effect among units that actually received dose level d, ATT($d \mid d$). Under strong parallel trends, this also estimates the average treatment effect, which, in the notation of Callaway et al. (2024), can be represented as:

$$\widehat{ATE}_{\hat{K}}(d) = (\psi^{\hat{K}}(d))^T \hat{\beta}_{\hat{K}} \tag{4}$$

The average causal response, $ACR(d)$, can be obtained by differentiating that curve. In practice, the $ACR(d)$ is approximated as:

$$\widehat{ACR}_{\hat{K}}(d) = (\delta\psi^{\hat{K}}(d))^T \hat{\beta}_{\hat{K}} \tag{5}$$

Plotting this over the range of d then shows whether each extra percentage-point of aid yields rising, flat, or diminishing returns on domestic HIV spending. Hence, under strong parallel trends, this derivative has a clear causal interpretation as "the marginal effect of one more point of aid at level d". This report will use both outcome measures to draw policy implications. Finally, this analysis computes point-wise 90% confidence intervals using the standard errors from the spline regression.

[2]The t subscript is dropped in alignment with the aforementioned general form set-up.
[3]$K = 6$ was the optimal number of basis functions selected by this cross-validation procedure.

6 Results

6.1 TWFE

This report's base analysis utilizes a continuous treatment DiD model with two-way fixed effects to estimate how marginal changes in PEPFAR aid (normalized by GDP) affect various HIV spending outcomes (also normalized by GDP). This interpretation can be made by adopting SPT, as mentioned in Section 5. *(Empirical Strategy)*.

The TWFE results are summarized in Table 6 below. They reveal a statistically significant negative relationship between PEPFAR aid intensity and domestic government HIV spending: a one percentage-point increase in PEPFAR aid corresponds to a 0.28 percentage-point reduction in government-financed HIV expenditures ($p < 0.05$). In economic terms, if a country's PEPFAR support rises from, for example, one percent of GDP to two percent of GDP, this analysis predicts an approximate 0.28 percentage-point drop in that country's own HIV budget as a share of GDP. This finding implies a crowding out effect, whereby higher levels of external assistance lead recipient governments to reduce their own budgetary commitments to HIV programs.

Table 6: TWFE Estimates: Continuous and Binary DiD Models

	Govt HIV Spending		Private HIV Spending		Total HIV Spending	
	(Continuous)	(Binary)	(Continuous)	(Binary)	(Continuous)	(Binary)
PEPFAR Aid (D) $\times$ Post	-0.279*	-0.041**	-0.220*	-0.021**	-0.499**	-0.061**
	(0.013)	(0.006)	(0.021)	(0.003)	(0.006)	(0.001)
Life Expectancy	-0.008	-0.007	-0.012	-0.012	-0.020*	-0.020*
	(0.094)	(0.101)	(0.080)	(0.083)	(0.035)	(0.038)
HIV Prevalence	0.118**	0.116**	0.043*	0.042*	0.162**	0.158**
	(0.007)	(0.007)	(0.026)	(0.032)	(0.001)	(0.001)
Corruption Rank	0.001	0.001	0.000	0.000	0.001	0.002
	(0.188)	(0.153)	(0.369)	(0.305)	(0.140)	(0.103)
Non-U.S. HIV DAH	0.031	0.014	0.051	0.037	0.082	0.052
	(0.330)	(0.668)	(0.358)	(0.467)	(0.261)	(0.460)
Health GDP	0.005	0.004	-0.000	-0.001	0.005	0.003
	(0.458)	(0.590)	(0.989)	(0.863)	(0.635)	(0.813)
Govt Health GDP	0.019	0.023	0.012	0.015	0.031	0.038
	(0.172)	(0.121)	(0.366)	(0.305)	(0.156)	(0.113)
Dose	0.069	-0.134	-0.002	-0.165	0.068	-0.299*
	(0.423)	(0.065)	(0.973)	(0.073)	(0.550)	(0.022)
No. Obs.	2,928	2,928	2,928	2,928	2,928	2,928
R^2	0.919	0.919	0.767	0.763	0.897	0.896
RMSE	**0.13**	**0.13**	**0.09**	**0.09**	**0.17**	**0.18**

Notes: a) $\cdot p < 0.1^{*}, p < 0.05^{**}$, $p < 0.001^{***}$, *b) SEs are in parenthesis and are clustered at the country level, c) the Post indicator was omitted due to multicollinearity with the interaction term, d) all models include year and country FEs.*

There is a similar pattern observed in private spending behavior: an incremental one percentage-point increase in PEPFAR aid is associated with a 0.22 percentage-point decrease in private HIV spending (which here combines out-of-pocket and prepaid private expenditures) ($p < 0.05$). The interpretation is consistent with the above: as aid intensity rises, households and insurers also scale back their contributions to HIV treatment and prevention. When public and private spending are aggregated into a total HIV spending measure, the estimated effect yields a 0.50 percentage-point decline in overall HIV expenditures ($p < 0.01$). This total-spending coefficient essentially sums the negative effects on public and private budgets, indicating that each additional percent of aid induces roughly half a percentage-point less spending on HIV in recipient countries (relative to GDP).

The binary specification robustness check corroborates these results. Table 6 shows how collapsing all aid levels into a single dummy yields a more muted response, illustrating that the full elasticity of crowding out is better captured when exploiting variation in aid intensity. Nonetheless, the binary results are directionally similar and statistically significant, indicating robustness in the results of this analysis.

These results suggest that the classical fiscal response mechanisms discussed in Section 3. *(Literature Review)*, such as budget fungibility, are at play. When aid lowers the marginal cost of public investment, governments may reallocate their own funds elsewhere due to political incentives, capacity constraints, or various other factors that lead to crowding out. Section 7. *(Policy Implications)* will explore where displaced funds may be reallocated and discuss the broader consequences of this crowding out effect.

6.1.1 Regional Effects

Table 7 presents the differences in the impact of PEPFAR aid across regions, using interactions between the continuous treatment variable and regional dummies. The estimates are compared to the overall TWFE coefficient for government HIV/AIDS spending to highlight differential effects by region.

Table 7: Heterogeneous Effects by Region

Variable	**Estimate**	**p-value**
$D \times POST$ (East Asia & Pacific)	$-0.231^{\cdot}$	0.076
$D \times POST$ (South Asia)	0.249^{*}	0.011
$D \times POST$ (Europe & Central Asia)	-0.331	0.263
$D \times POST$ (MENA)	0.220	0.364
$D \times POST$ (SSA)	-0.333^{**}	0.002
$D \times POST$ (LAC)	$-0.106^{\cdot}$	0.084

Notes: a) Standard errors are clustered at the country level,
b) $^{\cdot}p < 0.1^{*}, p < 0.05^{**}$, $p < 0.001^{***}$

We find substantial variation: in particular, Sub-Saharan Africa shows a large and statistically significant negative coefficient, which can be interpreted as Sub-Saharan Africa experiencing a -0.333 percentage point difference ($p < 0.01$) versus the average population result (-0.279, $p < 0.05$). South Asia displays a positive and statistically significant effect, meriting future exploration. All other regions have insignificant results.

These results indicate that regional differences play a central role in shaping the average treatment effect and underscore the importance of disaggregating DiD models where regional variances are suspected.

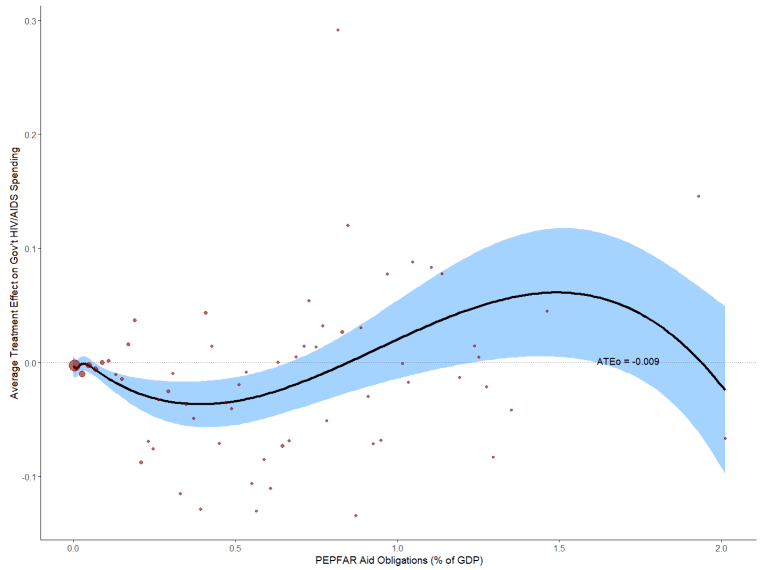
6.2 Nonparametric

6.2.1 Full Sample

Figure 3 below shows the estimated dose-response function for PEPFAR aid using the nonparametric approach. The solid black line represents $ATE(d)$, the average treatment effect for each observed dose of aid (as a percent of GDP). 90-percent point-wise confidence intervals are shown in blue, and red dots reflect binned means of the transformed outcome.

The curve reveals a clear nonlinear relationship between aid intensity and domestic HIV spending. At very low aid shares, changes in government HIV budgets are minimal. At moderate aid shares (roughly 0.05 to 0.7% of GDP), a statistically significant negative effect is observed, which is consistent with crowd out behavior: countries receiving modest aid reduce their domestic HIV budgets relative to untreated units. This is in alignment with the TWFE results of the analysis. In contrast, at higher doses (0.8 to 1.8%), countries appear to increase their domestic spending, suggesting crowd in effects. At very high doses, however, (particularly above 1.5%), wider confidence intervals reflect sparse data in this range, limiting the strength of any policy inference.

Figure 3: Nonparametric Estimates of $ATE(d)$ for PEPFAR Aid Flows



Notes: This figure plots nonparametric estimates of $ATE(d)$ using the methods proposed in Section 5.3.1. The blue-shaded region is the 90-percent point-wise confidence interval. The red dots represent the binned average change in government HIV spending from the pre-treatment period.

Averaging across all dose levels, the estimated overall effect is $ATE^o \approx -0.009$, indicating that negative responses at moderate doses nearly cancel out the higher-range increases. This is a critical observation with respect to the markedly more negative TWFE estimate (-0.279). The smaller average effect that is recovered suggests that TWFE may exaggerate the extent of crowd out by overweighting outcomes in dose ranges where the effect is most negative.

Nonetheless, this result underscores that the effect of PEPFAR varies meaningfully with aid intensity and the characterization of a single crowd out narrative should be made with caution. At mid-range doses, aid displaces domestic effort, but at higher intensities, PEPFAR may catalyze increased public investment (though this crowd in pattern could be strengthened with more high aid observations).

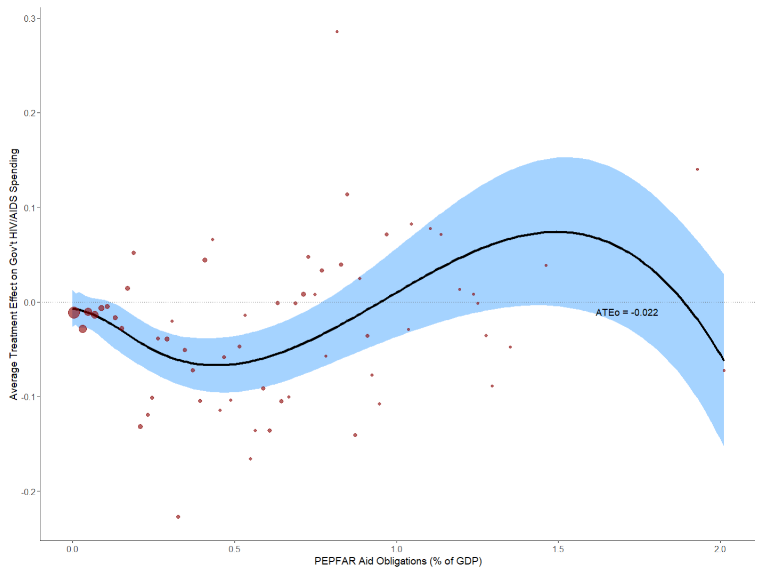
6.2.2 Sub-Saharan Africa (SSA) Sample

Figure 4 shows that the overall shape and magnitude of the estimated full sample dose-response curve are largely driven by countries in SSA, which receive the overwhelming share of PEPFAR funding as seen in Section 2. *(Data)*. When restricting the sample to this region, the shape of the curve remains similar to the full sample: crowd out at moderate doses and crowd in at higher levels, though the effects are somewhat more pronounced versus the full sample.

The downward dip between 0.2% and 0.7% of PEPFAR aid obligations in SSA is double that of the full sample, and the average across all doses, $ATE^o = -0.022$, is substantially more negative than the full sample correlate (-0.009). This more negative average reflects that, while the shape of the dose-response curve is similar, the distribution of observed doses in SSA is more concentrated in the range where effects are negative, which pulls the average effect downward. In short, the negative net effect in the full sample is not evenly spread across all regions, but rather it emerges primarily from the dose-response dynamics in SSA[4].

[4]See Figure A.2 for other region-specific nonparametric estimates, showing mainly flat-line or muted responses.

Figure 4: Nonparametric Estimates of $ATE(d)$ for PEPFAR Aid Flows in Sub-Saharan Africa

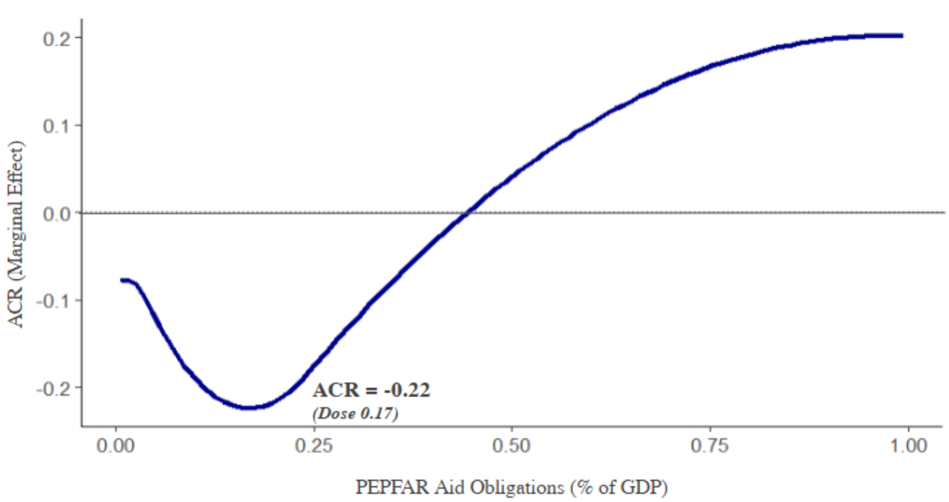


Notes: This figure plots nonparametric estimates of $ATE(d)$ using the methods proposed in Section 5.3.1., restricting to SSA only. As before, the blue-shaded region is the 90-percent point-wise confidence interval.

Figure 5 presents an estimate of the slope of the function presented in Figure 4 above, which under strong parallel trends, equals the $ACR(d)$. This analysis focuses on the range of PEPFAR aid obligations between 0.0% and 0.5% of GDP, which covers over 95% of the sample.

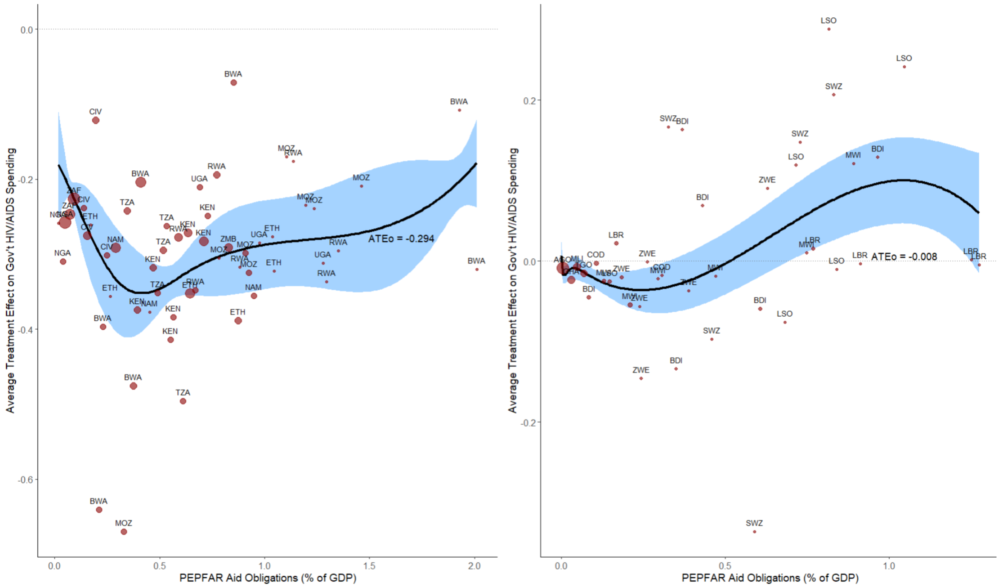
In this range, the results show that the $ACR(d)$ reaches a minimum value of -0.22 at 0.17% of GDP. Additionally, in this range the $ACR(d)$ is below zero and negative, reinforcing that marginal increases in aid are associated with reductions in domestic government spending over the majority of observations in this sample.

Figure 5: Nonparametric Estimates of $ACR(d)$ for PEPFAR Aid Flows in Sub-Saharan Africa



6.2.3 SSA Focus and Non-Focus Countries Sample

Figure 6: Nonparametric Estimates of $ATE(d)$ for PEPFAR Aid Flows in SSA Focus and Non-Focus Countries



Notes: This figure plots nonparametric estimates of $ATE(d)$ using the methods proposed in Section 5.3.1., restricting to SSA focus and non-focus countries only. As before, the blue-shaded region is the 90-percent point-wise confidence interval.

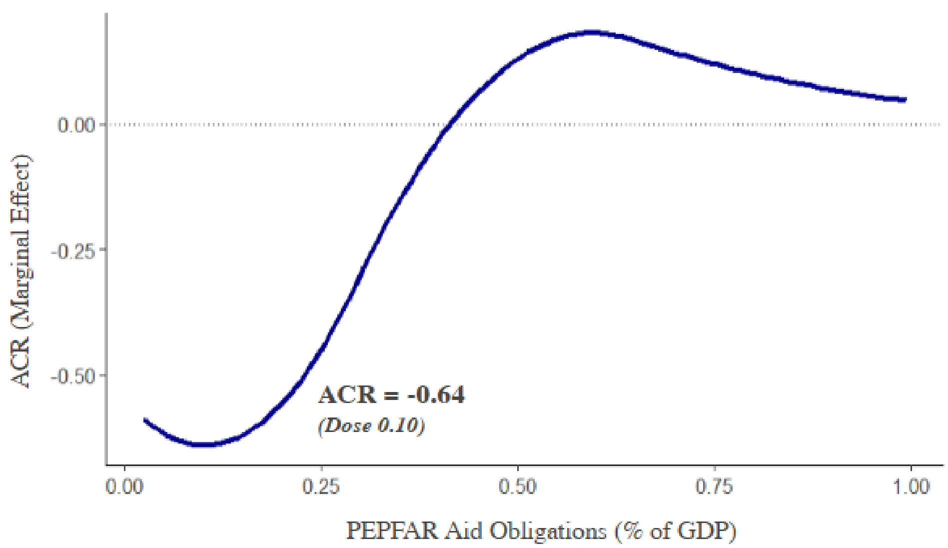
Figure 6 compares dose-response functions for focus versus non-focus countries within Sub-Saharan Africa. The left panel shows results for PEPFAR's original focus countries, and specifically, those prioritized for intensive funding and programmatic support. Across the entire aid distribution, the treatment effect curve for focus countries remains consistently negative, with no portion of the 90% confidence band rising above zero. This indicates persistent crowd out, even at relatively high levels of aid. The average treatment effect, $ATE^o = -0.284$, is both economically and statistically significant, and closely mirrors the TWFE estimate from earlier models (-0.28). This alignment suggests that the TWFE estimate is largely driven by the focus group countries, which both received more aid and exhibited more consistently negative responses.

By contrast, the right panel shows that non-focus countries display a more heterogeneous response. While the curve dips slightly below zero at low doses, it becomes positive and statistically distinguishable from zero between approximately 0.6% and 1.5% of GDP, suggesting a crowd in effect at moderate aid levels. Overall, the average treatment effect in non-focus countries is much smaller at $ATE(d) = -0.008$, where variation is primarily driven by a handful of countries (like Lesotho, Swaziland, and Burundi).

Together, these patterns further reinforce that PEPFAR's impact varied systematically not only by dose, but also by country context. In focus countries, where aid was largest and most sustained, the consistently negative treatment effects may reflect strategic resource allocation, as governments substituted external funds for their own. This response appears particularly strong in countries like Botswana and Kenya, both of which received high-intensity aid and show sharply negative estimates, consistent with deliberate fiscal displacement. By contrast, non-focus countries received more inconsistent support, and in some cases increased domestic spending in response. The crowd in observed at moderate doses suggests that aid acted as a fiscal complement where governments were less able or willing to substitute away from HIV budgets.

Lastly, the $ACR(d)$ for SSA focus countries further strengthens the theory that the crowd out effect is driven by this sub-sample. As shown below in Figure 7, the $ACR(d)$ confirms that there is a crowding out effect between 0.0% and about 0.4% of PEPFAR aid, with the $ACR(d)$ reaching its minimum at -0.64 at a 0.1% aid level. Hence, not only does this $ACR(d)$ further justify the conclusion of SSA focus countries being the primary driver of the crowding out effect, but further supports the argument that the crowding out effect is most prevalent in low to moderate aid doses.

Figure 7: Nonparametric Estimates of $ACR(d)$ for PEPFAR Aid Flows in Focus Countries (SSA Only)



6.2.4 Explaining Divergence Between $\hat{\beta}^{\mathbf{TWFE}}$ and ATE^o

The substantial gap between the baseline estimate from the TWFE model (-0.279, $p < 0.05$) and the spline-based average treatment effect most likely reflects differences in weighting strategy, as discussed in Callaway et al. (2024). While the latter averages dose-specific treatment effects across the empirical distribution of aid exposure, β^{TWFE} applies a variance-based weighting scheme that places disproportionate emphasis on units with extreme treatment values, particularly those receiving high doses. This leads β^{TWFE} to be heavily influenced by a small set of high-intensity treated units (in this report's case, SSA focus countries).

This can generate misleading conclusions when treatment effects vary. In this setting, the nonlinear spline regression reveals a pattern where moderate doses produce negative effects, while higher-range aid levels generate positive responses. Fitting a single linear slope to such a curve—as β^{TWFE} does—obscures this structure and effectively collapses the variation into a contrast between low and high doses.

In short, β^{TWFE} captures a model-driven weighted average that overrepresents high-dose observations, whereas ATE^o reflects a transparent average across the actual dose distribution. This underscores the risks of relying solely on β^{TWFE} in settings with strong treatment effect heterogeneity.

7 Policy Implications

Heterogeneous Fiscal Responses and the Importance of Aid Calibration

The findings of this paper point to an overall crowd out effect of aid, alongside substantial heterogeneity depending on the level of support received. This pattern holds across all methods employed, including TWFE estimates (for both continuous and binary treatments) and the non-parametric analysis. As the intensity of aid increases, particularly at moderate levels in the focus countries, we observe clear signs of a crowd out effect, where increased external funding is associated with reductions in domestic HIV-related budget allocations.

This pattern carries important implications for the design of future aid programs, particularly single-disease initiatives. While each program has its own structure and dynamics, PEPFAR, as the largest disease-specific aid initiative in history, offers valuable lessons. Its fiscal effects highlight the need for careful consideration of how aid is structured and delivered, informing both future program design and potential reforms to existing aid frameworks. This underscores the need for aid programs to be wary of making a singluar ”crowd out” or ”crowd in” conclusion, and instead take a more nuanced approach of understanding how fiscal responses vary by aid intensity. Tracking marginal returns to aid and maintaining transparency about government resource allocation are essential steps toward optimizing the “dose” of support.

Reallocation of Crowded-Out Funds

However, crowding out can sometimes yield positive outcomes: when displaced funds are redirected to other high priority sectors such as primary care, maternal health, or education, the overall welfare impact of donor aid may remain beneficial. By covering HIV-related costs, aid can also free up domestic resources to address other urgent fiscal priorities. Batniji and Bendavid (2013) argue that the core question is not whether displacement occurs, but what becomes of the displaced funds.

The results provide some reassurance on this point (see Table A.5 Although the precision of the magnitude of our estimates is limited by the lack of data on relevant control variables, the results suggest a directionally positive relationship: increases in PEPFAR aid are associated with higher spending on broader non-HIV health expenditures.

Supporting this view, a joint study by amfAR and the World Bank (2025) found that PEPFAR-recipient countries in Sub-Saharan Africa experienced a 150 percent greater increase in per capita health spending compared to non-recipient countries over the same period. This may help explain the observed decline in all-cause mortality (KFF 2021).

Future research could investigate whether these reallocations improve long term health outcomes. For now, the findings of this report point to a form of sectoral continuity rather than erosion, which is consistent with PEPFAR's stated goal of strengthening health systems alongside disease-specific programming.

Fiscal Vulnerability and the Risks of Abrupt Aid Withdrawal

The focus countries in Sub-Sarahan Africa that exhibit crowd out effects are also those most vulnerable to fiscal disruption if external support is reduced or withdrawn. These countries depend heavily on donor funding to sustain their HIV response. With growing uncertainty surrounding the program's long-term funding, countries in this region may be forced to reallocate domestic resources to fill emerging gaps. This reallocation could create significant fiscal pressure, as governments may have to divert funds from other pressing needs. As a result, they may face difficult trade-offs between maintaining HIV service coverage and sustaining other essential services such as education, infrastructure, or primary healthcare.

Accounting for fiscal displacement in treatment evaluations

These findings point to an important caveat for aid impact evaluations: ignoring fiscal displacement can distort estimates of effectiveness. When external funding reduces domestic effort, the actual increase in available resources is smaller than the nominal value of aid. Evaluations assuming full additionality risk understating treatment effects.

This issue affects both retrospective impact assessments and forward-looking cost-effectiveness analyses. Programs may appear more efficient or scalable than they are if impacts are calculated by assuming that all funds are additive. On the other hand, programs that trigger beneficial reallocations might be undervalued.

More accurate evaluations should model both direct outcomes and induced fiscal spillovers. Future research should consider quantifying net resource changes by combining aid flows with domestic responses. This may involve disaggregated budget data, sectoral reallocation tracking, or estimating crowd out elasticities to integrate into impact models.

8 Limitations and Possible Extensions

8.1 Limitations

The primary limitations of this analysis are (1) structure of PEPFAR data and (2) feasibility of adopting the SPT assumption. On (1), the right-skewness of PEPFAR funding data as a percent of GDP, as shown in Section 4. *(Data)*, implies sparse observations at high aid levels. As such, conclusive policy implications cannot be made for countries receiving the highest amount of aid as a percent of GDP. Regarding (2), as noted in Section 5. *Empirical Strategy*, the causal interpretation of findings relies on the SPT assumption. However, as SPT cannot be directly tested, it is possible it may not hold in this empirical setting.

8.2 Possible Extensions

As noted in Section 7. *(Policy Implications)*, further research is needed to understand the nuance of where governments are allocating funds that were previously used on domestic HIV/AIDS response, as well as the associated outcomes to health and well-being. With detailed budget data, it would be useful to build a more granular treatment-response framework that can track shifts in domestic expenditure in response to foreign aid flows.

9 Conclusion

This report provides new empirical evidence on how PEPFAR aid affects domestic HIV/AIDS spending. Across all specifications, we find robust evidence that aid partially displaces domestic government spending on HIV. Overall, the effect is strongest in Sub-Saharan Africa and in PEPFAR focus countries. Similar patterns emerge in private spending as well.

While the TWFE estimator shows consistent crowd out, the nonparametric B-spline modeling reveals a more nuanced picture. The estimated effect varies across dose levels, resulting in an overall average treatment effect much smaller than TWFE. Disaggregation shows that SSA focus countries experience consistent crowd out, while non-focus countries display much weaker crowd out, with some showing positive responses, suggesting heterogeneity and overall limited displacement.

Aforementioned countries in Sub-Saharan Africa are at significant fiscal risk if PEPFAR support is reduced or withdrawn. They not only rely heavily on external funding to sustain their HIV response, but have also reallocated domestic resources over time, creating structural dependence on PEPFAR aid. The dual exposure, high aid inflows and domestic fiscal reallocation, raises concerns about abrupt service disruptions and difficult trade-offs across sectors. More broadly, the analysis reveals that the fiscal response to aid is both heterogeneous and dose-dependent, with partial crowd out effects at moderate aid levels and signs of positive reallocation in some cases. These findings point to the need for tailored transition planning, closer tracking of budgetary shifts, and evaluation methods that account for variation in treatment intensity and endogenous fiscal behavior.

Econometrically, these findings reinforce growing concerns in the literature about the limitations of TWFE in settings with continuous treatments and heterogeneous effects (Callaway et al. 2024). When treatment intensity varies (as with PEPFAR aid) and responses are nonlinear, TWFE can produce biased estimates by averaging over dose-specific effects and disproportionately weighting high-variance groups or time periods. By contrast, the spline-based approach provides a more flexible and policy-relevant characterization of the dose-response relationship. This yields a more accurate and nuanced understanding of how external aid shapes domestic fiscal behavior, an essential step for designing more effective and accountable global health financing strategies.

References

[1] Acemoglu, D.; and Finkelstein, A. (2008). Input and Technology Choices in Regulated Industries: Evidence from the Health Care Sector. *Journal of Political Economy*, 116(5), 837–880. Retrieved from `https://doi.org/10.1086/595014`.

[2] Callaway, Brantly; Goodman-Bacon, Andrew; and Sant'Anna, Pedro H. C. (2024). *Difference-in-Differences with a Continuous Treatment*. arXiv preprint (first draft: July 6, 2021; updated January 26, 2024). Retrieved from `https://psantanna.com/files/CGBS.pdf`.

[3] Callaway, Brantly, Andrew Goodman-Bacon, and Pedro H. C. Sant'Anna. 2024. "Event Studies with a Continuous Treatment." AEA Papers and Proceedings 114: 601–05. Retreived from `https://www.aeaweb.org/articles?id=10.1257/pandp.20241047`.

[4] Carter, Patrick. (2013). Does Foreign Aid Displace Domestic Taxation? *Journal of Globalization and Development*, 4(1), Article 4. Available at: `https://doi.org/10.1515/joeg-2013-0004`.

[5] Crown, W.; Boden, L. I.; Higuera, I. A.; Loughlin, A. M.; Ochalek, J.; and Torres-Rueda, S. (2023). PEPFAR's impact on economic development, health system strengthening, and health outcomes. *PLOS ONE*, 18(10), e0289909. Retrieved from `https://doi.org/10.1371/journal.pone.0289909`

[6] Dieleman, Josephine L.; Haakenstad, Annie; Micah, Ariel; Moses, Matthew; Abbafati, C.; Acharya, P.; ... Murray, Christopher J. L. (2018). Spending on health and HIV/AIDS: Domestic health spending and development assistance in 188 countries, 1995–2015. *The Lancet*, 391(10132), 1799–1829. Retrieved from `https://doi.org/10.1016/S0140-6736(18)30698-6`

[7] Feyzioglu, Tarhan; Swaroop, Vinaya; and Zhu, Min. (1998). A Panel Data Analysis of the Fungibility of Foreign Aid. *The World Bank Economic Review*, 12(1), 29–58. Retrieved from `https://www.jstor.org/stable/3990208`.

[8] Franco-Rodriguez, Sonia; McGillivray, Mark; and Morrissey, Oliver. (1998). Aid and the public sector in Pakistan: Evidence with endogenous aid. *World Development*, 26(7), 1241–1250. Retrieved from `https://doi.org/10.1016/S0305-750X(98)00048-5`

[9] Global Burden of Disease Health Financing Collaborator Network. (2018). *Global HIV/AIDS Spending 2000–2015*. Seattle, WA: Institute for Health Metrics and Evaluation (IHME). Retrieved from `http://ghdx.healthdata.org/record/ihme-data/ghdx-global-hiv-aids-spending-2000-2015`.

[10] Heller, Peter S. (1975). A model of public fiscal behavior in developing countries: Aid, investment and taxation. *American Economic Review*, 65(3), 429–445. JSTOR Stable URL: `https://www.jstor.org/stable/1804844`

[11] HIV.gov. (2024). The U.S. President's Emergency Plan for AIDS Relief (PEPFAR). Retrieved April 10, 2025, from `https://www.hiv.gov/federal-response/pepfar-global-aids/pepfar`.

[12] Institute for Health Metrics and Evaluation. (2024). Development Assistance for Health Database 1990–2023. Global Health Data Exchange. Retrieved May 25, 2025, from `https://ghdx.healthdata.org/record/ihme-data/development-assistance-health-database-1990-2023`.

[13] Institute of Medicine. (2013). *Evaluation of PEPFAR: Phase 2*. Washington, DC: The National Academies Press. Retrieved from `https://www.nap.edu/catalog/13545/evaluation-of-pepfar-phase-2`.

[14] Kaiser Family Foundation. (2021). Assessing PEPFAR's impact: Analysis of mortality in PEPFAR countries. Retrieved April 10, 2025, from `https://www.kff.org/global-health-policy/issue-brief/assessing-pepfars-impact-analysis-of-mortality-in-pepfar-countries/`.

[15] Kaiser Family Foundation. (2025). The Trump administration's foreign aid review & the status of PEPFAR. Retrieved April 10, 2025, from `https://www.kff.org/global-health-policy/fact-sheet/the-trump-administrations-foreign-aid-review-status-of-pepfar/`.

[16] Liang, Lihui-Lou; and Mirelman, Ana J. (2014). Why do some countries spend more for health? An assessment of sociopolitical determinants and international aid for government health expenditures. *Social Science & Medicine*, 114, 161–168. Retrieved from https://doi.org/10.1016/j.socscimed.2014.05.044

[17] Lu, Christopher; Cook, Benjamin; and Desmond, Chris. (2017). Does foreign aid crowd out government investments? Evidence from rural health centres in Rwanda. *BMJ Global Health*, 2(1), e000364. Retrieved from https://doi.org/10.1136/bmjgh-2017-000364

[18] Morrissey, Oliver; and Torrance, Susan. (2014). Aid and taxation: Exploring the relationship using new data (ICTD Working Paper 21). Brighton, UK: Institute of Development Studies. Retrieved from https://assets.publishing.service.gov.uk/media/57a089eded915d622c000475/ICTD-WP21-FINAL.pdf.

[19] Ouattara, Boubacar. (2006). Foreign aid and government fiscal behaviour in developing countries: Panel data evidence. *Economic Modelling*, 23(3), 506–511. Retrieved from https://doi.org/10.1016/j.econmod.2006.02.001

[20] Our World in Data. (n.d.). HIV/AIDS. Retrieved June 4, 2025, from https://ourworldindata.org/hiv-aids.

[21] Over, Mead; and Glassman, Amanda. (2013). *Financial Flows and Public Health: The Impact of PEPFAR Investments on Health Systems Strengthening*. Washington, DC: Center for Global Development. Retrieved from https://www.cgdev.org/sites/default/files/financial-flows-pepfar_.pdf.

[22] Remmer, Karen L. (2004). Does foreign aid promote the expansion of government? *American Journal of Political Science*, 48(1), 77–92. Retrieved from https://doi.org/10.1111/j.0092-5853.2004.00057.x

[23] Science & Diplomacy. (2013, December 1). Making PEPFAR. Retrieved April 10, 2025, from https://www.sciencediplomacy.org/article/2013/making-pepfar.

[24] Sridhar, Devi; and Batniji, Rajeev. (2008). Misfinancing global health: A case for transparency in disbursements and decision making. *The Lancet*, 372(9644), 1185–1191. Retrieved from https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(08)61485-3/fulltext.

[25] U.S. Department of State. (2024). PEPFAR: Latest global results fact sheet (December 2024). Retrieved April 10, 2025, from https://www.state.gov/pepfar-latest-global-results-factsheet-dec-2024/.

[26] U.S. Department of State. (n.d.). Data – ForeignAssistance.gov. Retrieved April 10, 2025, from https://foreignassistance.gov/data.

[27] World Bank. (2025). World Development Indicators. World Bank Open Data. Retrieved May 25, 2025, from https://databank.worldbank.org/source/world-development-indicators.

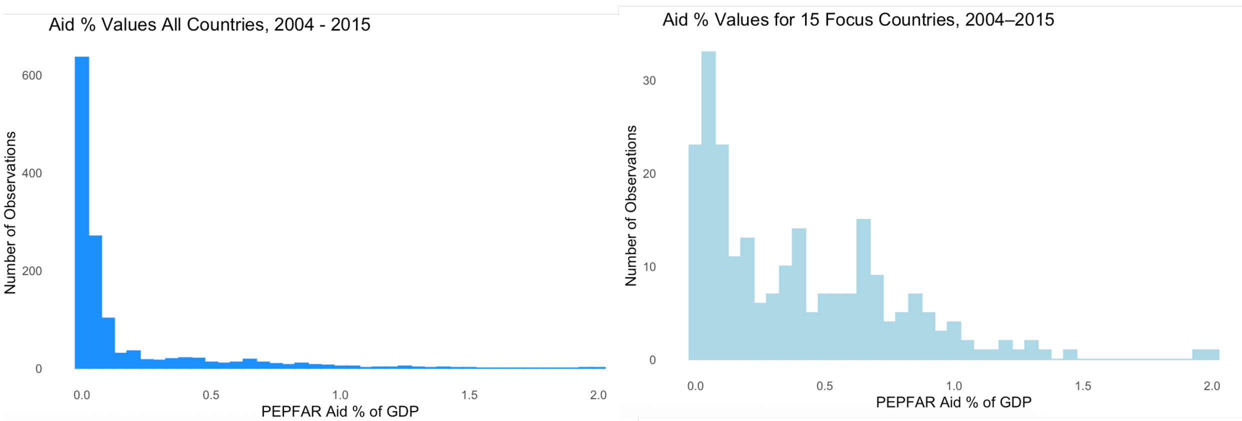
[28] World Bank. (2025). Worldwide Governance Indicators. World Bank Open Data. Retrieved May 25, 2025, from https://databank.worldbank.org/source/worldwide-governance-indicators.

[29] amfAR and The World Bank. (2025). *Progress in PEPFAR-Supported Countries: Domestic Resource Mobilization and Health System Strengthening*. Retrieved June 5, 2025, from https://www.amfar.org/wp-content/uploads/2025/04/IB-World-Bank-Country-Analysis.pdf

A Additional Figures

Figure A.1: Histograms of PEPFAR obligations observations, normalized by GDP, plus summary statistics.

Panel A: All Recipient Countries | **Panel B: Focus Countries**



Summary Statistics of Aid Share (All Countries)

Min Aid %	Max Aid %	Mean Aid %	Median Aid %
3e-09	2.010194	0.1353531	0.02608911

Summary Statistics of Aid Share (Focus Countries)

Min Aid %	Max Aid %	Mean Aid %	Median Aid %
0.0016	2.0102	0.4001	0.297

Figure A.2: Nonparametric Estimates of $ATE(d)$ in All Regions

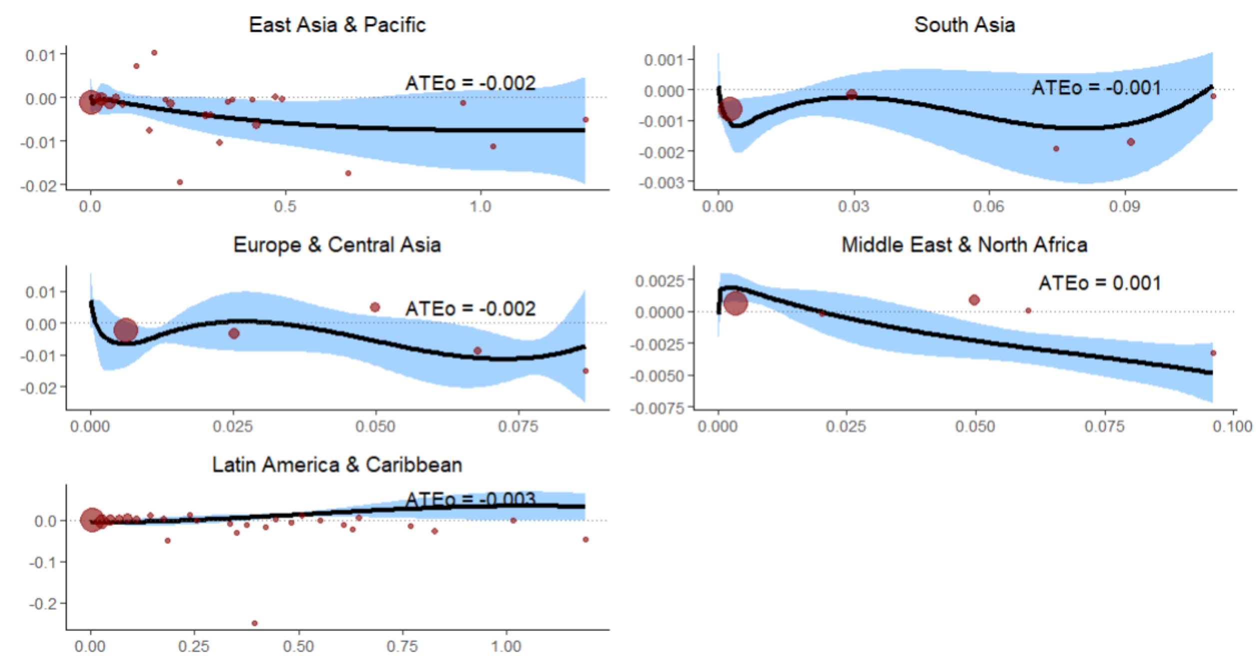


Table A.1: Nonparametric Regression on PEPFAR Aid (Full Sample)

Variable	**Estimate**	**Std. Error**	***t*-value**
Intercept	-0.0843	0.0358	-2.356
bs(D, df = 5)[,1]	-0.0069	0.0049	-1.410
bs(D, df = 5)[,2]	0.0069	0.0052	1.337
bs(D, df = 5)[,3]	-0.1467	0.0529	-2.771**
bs(D, df = 5)[,4]	0.2194	0.0894	2.455**
bs(D, df = 5)[,5]	-0.0201	0.0450	-0.446
Life expectancy	0.0014	0.0006	2.229**
HIV prevalence	0.0031	0.0021	1.466
Corruption rank	-0.0003	0.0002	-1.565
Non-U.S. HIV DAH (% GDP)	-0.0092	0.0133	-0.689
Health spending (% GDP)	-0.0005	0.0014	-0.357
Gov't health spending (% GDP)	0.0007	0.0037	0.181
RMSE		0.0763	
Adjusted R^2		0.0314	

Table A.2: Nonparametric Regression on PEPFAR Aid (Sub-Saharan Africa)

Variable	**Estimate**	**Std. Error**	***t*-value**
Intercept	-0.1641	0.0933	-1.760·
bs(D, df = 5)[,1]	-0.0010	0.0173	-0.060
bs(D, df = 5)[,2]	-0.0062	0.0172	-0.362
bs(D, df = 5)[,3]	-0.2030	0.0661	-3.066**
bs(D, df = 5)[,4]	0.2936	0.1091	2.690*
bs(D, df = 5)[,5]	-0.0522	0.0628	-0.874
Life expectancy	0.0028	0.0017	1.659
HIV prevalence	0.0036	0.0025	1.444
Corruption rank	-0.0008	0.0006	-1.251
Non-U.S. HIV DAH (% GDP)	-0.0045	0.0165	-0.274
Health spending (% GDP)	-0.0012	0.0030	-0.385
Gov't health spending (% GDP)	0.0055	0.0214	0.256
RMSE		0.1129	
Adjusted R^2		0.0279	

Notes: Outcome is transformed Δ in government HIV/AIDS spending as % of GDP. Splines use 6 degrees of freedom. Standard errors clustered by country (location_id).
** $p < 0.01$, * $p < 0.05$, $^{\cdot}$ $p < 0.1$

Table A.3: Nonparametric Regression on PEPFAR Aid (Focus SSA Countries)

Variable	**Estimate**	**Std. Error**	***t*-value**
Intercept	-0.7973	0.2542	-3.136**
bs(D, df = 5)[,1]	-0.0530	0.0886	-0.598
bs(D, df = 5)[,2]	-0.2150	0.0949	-2.265*
bs(D, df = 5)[,3]	-0.0438	0.1253	-0.349
bs(D, df = 5)[,4]	-0.1471	0.1591	-0.925
bs(D, df = 5)[,5]	0.0066	0.0533	0.124
Life expectancy	0.0118	0.0038	3.062*
HIV prevalence	0.0021	0.0025	0.586
Corruption rank	-0.0016	0.0010	-1.472
Non-U.S. HIV DAH (% GDP)	0.0020	0.0190	0.158
Health spending (% GDP)	-0.0197	0.0432	-0.456
Gov't health spending (% GDP)	0.0197	0.0542	0.364
RMSE		0.1530	
Adjusted R^2		0.0427	

Table A.4: Nonparametric Regression on PEPFAR Aid (Non-Focus SSA Countries)

Variable	**Estimate**	**Std. Error**	***t*-value**
Intercept	-0.0906	0.0592	-1.529
bs(D, df = 5)[,1]	-0.0361	0.0158	-1.933·
bs(D, df = 5)[,2]	-0.0080	0.0099	-0.808
bs(D, df = 5)[,3]	-0.1441	0.0680	-2.118*
bs(D, df = 5)[,4]	0.2055	0.0772	2.658**
bs(D, df = 5)[,5]	0.0533	0.0475	1.111
Life expectancy	0.0016	0.0010	1.657
HIV prevalence	0.0044	0.0024	1.843
Corruption rank	0.0000	0.0004	0.085
Non-U.S. HIV DAH (% GDP)	-0.0371	0.0169	-2.196*
Health spending (% GDP)	0.0003	0.0018	0.184
Gov't health spending (% GDP)	0.0005	0.0180	0.225
RMSE		0.0728	
Adjusted R^2		0.1638	

Notes: Outcome is transformed Δ in government HIV/AIDS spending as % of GDP. Splines use 6 degrees of freedom. Standard errors clustered by country (location_id).
** $p < 0.01$, * $p < 0.05$, · $p < 0.1$

Table A.5: TWFE Estimates of PEPFAR Aid on Non-HIV Health Spending

	Gov Non-HIV Spending	**Private Non-HIV Spending**
PEPFAR Aid (Continuous)	0.5578* (0.2790)	1.190* (0.4865)
PEPFAR Aid (Binary)	0.1443* (0.0625)	0.2620* (0.1098)
Fixed Effects	Yes (Country, Year)	Yes (Country, Year)
Observations	2,928	2,928
R^2	0.9433	0.8830
Within R^2	0.0171	0.0592

Standard errors clustered at the country level.

$^*p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$