The Lifesaving Impact of Electronic Medical Records for HIV Patients

Laura Derksen, Anita M. McGahan and Leandro S. Pongeluppe*

June 19, 2025

Abstract

This paper shows that replacing paper-based records with electronic medical records (EMRs) improves HIV patient retention and prevents AIDS deaths in the low-income country of Malawi. An event study of 106 HIV clinics shows a 28 percent reduction in annual deaths five years after EMR implementation, with the greatest impact on children. Improvements in health outcomes appear due to efficiency gains, rather than to changes in the medical care provided at visits. These efficiency gains allow clinics to better manage patient data, trace lapsed patients and return them into care, and adapt to higher patient volumes over time.

JEL Classification: I18 I12 I15

*Derksen: Ragnar Frisch Centre for Economic Research and University of Toronto, laura.derksen@frisch.uio.no; McGahan: Rotman School of Management, University of Toronto, anita.mcgahan@rotman.utoronto.ca; Pongeluppe: Wharton School, University of Pennsylvania. pongelup@wharton.upenn.edu We are grateful for comments from Alessandro Acquisti, Marcella Alsan, Heski Bar-Isaac, April Franco, Rachel Glennerster, Avi Goldfarb, Mitchell Hoffman, Anne Karing, Jason Kerwin, Michael Kremer, Kory Kroft, Nicola Lacetera, Amalia Miller, Matt Mitchell, Heather Sarsons, and Olivier Sterck. We thank the editor, Tavneet Suri, and three referees for constructive comments. We thank Abdallah Chilungo, Joshua Berman, Jean Bourgeois, Sumeet Sodhi, Benjamin Kumwenda, and Baobab Health Trust for their collaboration, and Jessica Gallant, Francis Kamungu, Jonathan Kandodo, Macdonald Nazimera, and Dina O'Brien for research assistance. Funded by a Connaught Global Challenge Award and approved by IRBs at the Malawi College of Medicine (2572/2019) and University of Toronto (36913/2020). Errors and omissions are our own.

1 Introduction

Over 40 million people globally are living with Human Immunodeficiency Virus (HIV), with the majority residing in low-income countries. Untreated HIV leads to Acquired Immune Deficiency Syndrome (AIDS), opportunistic infection, and death. Ending the HIV/AIDS epidemic requires retaining patients consistently in care and on antiretroviral therapy (ART). Regular adherence to ART not only extends life by decades but also significantly reduces the risk of HIV transmission. However, many HIV patients lapse from care, leading to new infections and premature deaths. Efforts to trace patients are often hampered by the challenge of identifying those who have lapsed. In high-volume, under-staffed clinics, the ability to efficiently access and manage patient data is critical for ensuring effective treatment and follow-up.

In this paper, we study the rollout of an electronic medical record (EMR) system to 106 Malawian HIV treatment clinics between 2007 and 2019. Under the EMR system, paper records are comprehensively converted into electronic ones, and ART patient data is then updated and accessed electronically by clinic staff at the point of care. The system is designed to allow clinics to manage patient data more efficiently by simplifying data queries, facilitating continuity of care through the retrieval of patient records, and streamlining data entry tasks. In particular, the EMR system can quickly generate a list of HIV patients who have missed appointments, allowing clinic staff to trace them by phone or by visiting their homes. In non-EMR clinics, the staff must examine thousands of individual paper records to identify lapsed patients. The EMR system also supports adherence to healthcare protocols by providing guidance, prompts and alerts to clinical staff during patient visits.

We show that the introduction of the EMR system leads to a gradual reduction in patient deaths, with effects concentrated among younger patients. Our estimates are based on an event study specification with clinic and year fixed effects, as well as several robust alternatives. Five years after its introduction, the EMR system leads to an estimated 28 percent decrease in annual patient deaths. This effect is concentrated in patients under the age of 50, with large impacts for young children, who are particularly vulnerable to lapses in care. The effect is also larger in hospitals, where patients are younger and more likely to have advanced HIV.

The broad efficiency gains delivered by the EMR system enable clinics to manage higher patient volumes and retain patients more effectively without exceeding capacity limits. Among potential mechanisms, we find that the decrease in patient deaths can be primarily attributed to data-related efficiency gains, which enable clinics to better identify, trace and treat patients who have lapsed from care. In clinics that have adopted EMRs, patients are 5 percentage points less likely to be lapsed from care in a given year. This leads to a large and significant increase in patient retention and total patient volume at the clinic level, as well as improved patient health. We also find that EMRs support better continuity of care, as patients who return to an EMR clinic are more likely to attend later follow-up appointments. These improvements do not rely on an increase in clinic opening hours or by an increase in staffing levels.

We do not observe an increase in the frequency of scheduled visits, nor a change in the nature of medical care provided to patients at clinic visits. In fact, clinics appear to adapt to the influx of returning patients by reducing the frequency of scheduled visits. Despite prompts from the EMR system, we do not observe an increase in the likelihood that new patients are tested and treated for tuberculosis (TB, one of the leading causes of mortality among HIV patients), referred for other tests and services, or, in the case of underweight children, provided with effective nutrition support.

Aggregated over time, our findings imply that the adoption of EMRs has prevented 5,050 AIDS deaths. We estimate that the cost of EMR implementation is approximately USD $448 per life saved in the first five years. For a point of comparison, the most effective life-saving programs implemented by other charitable organizations have been estimated to cost USD $4,500 per life saved.[1] EMR implementation in Malawi is also much more cost-effective than implementing an EMR-system in the United States; Miller and Tucker (2011) find that implementing EMR in American hospitals costs USD $531,000 per baby's life saved.

Our findings carry implications for global health policy, research on the impact of EMR

[1] Source: GiveWell Available at: https://www.givewell.org/how-we-work/our-criteria/cost-effectiveness/cost-effectiveness-models. Accessed on April 30th, 2022.

and digitization, and demand for HIV care.

By many measures, the Malawian response to the AIDS epidemic has been a success, and has informed policy in Africa and beyond (Harries et al., 2016). This is remarkable in part because Malawi is a low-income country, with a GDP per capita of USD $645 (World Bank, 2022). Over the past decade, the rate of new infections has decreased by 72 percent and AIDS deaths have fallen by 68 percent.[2] As of 2021, 85 percent of Malawians living with HIV are on treatment and virally suppressed, compared to 73 percent in East and Southern Africa, and 75 percent across Western Europe and North America. Our findings indicate that the EMR system played a significant role in this success, and that other countries affected by the AIDS epidemic might expect large benefits from adopting EMRs and using them to engage in patient outreach.

This paper addresses a last-mile problem in global health: how do we ensure that lifesaving innovations reach those who need them the most? Medical innovations, such as vaccines and medicines, have led to dramatic health improvements worldwide, including in low-income countries. Yet, where clinics are under-resourced and understaffed, efficiency gains can allow clinics to reach many more patients and treat them more effectively. We contribute to a broader literature on healthcare policy that goes beyond the provision of medicines, by showing that managerial interventions such as EMR systems have important health impacts, and can be implemented at scale in low-resource settings. Such policies include incentives for immunization (Banerjee et al., 2010), the free distribution of health supplies to households (Dupas, 2009), and the use of scheduled healthcare appointments (Derksen et al., 2024). Solutions enabled by digital communication technology have an important role to play in improving the efficient delivery of healthcare, particularly in low-resource settings.

We find that in a country with limited healthcare resources and low staff-to-patient ratios, adopting an EMR system can lead to important improvements in patient outcomes. Prior studies on the impact of EMRs – and more broadly, information technology (IT) – have focused primarily on the United States (Miller and Tucker, 2009, 2011; Bhargava and Mishra, 2014; Dranove et al., 2014) and have generally reported only minor impacts on quality, staff productivity, and cost-effectiveness (Agha, 2014; Bhargava and Mishra, 2014; Dranove et al.,

[2] Source: UNAIDS (2022).

2014).[3] However, even in some highly resourced, highly organized contexts, EMR adoption does appear to improve quality of care, and can impact patient health including by reducing mortality (McCullough et al., 2010; Miller and Tucker, 2011; McCullough et al., 2016). Miller and Tucker (2011) show that EMRs can facilitate access to patient records, and thus reduce neonatal mortality rates. Consistent with this result, we find that the impact of tracing on HIV patient outcomes after EMR implementation in Malawi is strongest among children. In our setting, the mechanism is clear: EMRs allow healthcare staff to easily query a large patient database according to particular criteria (missed appointments). This mechanism is relevant in any context where follow-up is required for a subset of patients, for example, based on laboratory results or the date of a previous vaccine or procedure.[4] Low-income countries often face high health burdens with limited resources and severe staffing shortages. We find that in such a setting, the efficiency gains from EMR adoption can be particularly large.

This study identifies a managerial intervention, namely the implementation of an EMR system, that prevents AIDS deaths by returning lapsed patients into active care. Improving patient adherence to medical treatment is a multi-faceted challenge, with potential interventions that include education, behavioral tools, and incentives (Vermeire et al., 2001; Nattrass, 2019). Digital health interventions such as text message reminders have been found to improve adherence to ART treatment slightly (Pop-Eleches et al., 2011; Mbuagbaw et al., 2013; Wang et al., 2019). Related work has evaluated interventions that affect demand for HIV testing (Thornton, 2008; de Walque et al., 2015; Banerjee et al., 2019; Derksen et al., 2022; Macis et al., 2021; Derksen et al., 2024). While testing is an important first step towards treatment, in this paper we directly observe and report impacts on important outcomes such as retention in HIV treatment and averted patient deaths.

The remainder of the paper is organized as follows. Section 2 provides a description of the context. In Section 3, we discuss the data and the empirical approach. In Section 4, we

[3] Douglas et al. (2010) present qualitative evidence on EMR implementation challenges in developing countries, and Driessen et al. (2013) estimates the cost savings associated with Malawi's EMR system.

[4] See Shadmi et al. (2020) for an example from the Israeli healthcare system, where machine learning was used to identify patients at risk of readmission.

present our findings, and in Section 5, we conclude.

2 Context

2.1 HIV Care in Malawi

The global HIV/AIDS pandemic is ongoing, with approximately 40 million people currently infected (UNAIDS, 2022). Malawi is one of the most HIV-affected countries in the world. With 9.5 percent HIV of the population estimated to be HIV positive in 2019, the country ranks 9th worldwide in HIV prevalence. Approximately 13 thousand Malawians died of AIDS in 2019 out of a population of nearly 20 million. In 2004, at the peak of the epidemic, more than 70 thousand AIDS deaths were recorded.[5]

While the number of people living with HIV has increased over time, AIDS deaths have decreased dramatically due to the widespread availability of ART. ART is a combination of drugs, taken daily, that suppresses the virus over the long term. ART restores immune system function, prevents opportunistic infections, reverses the progression of AIDS, and prolongs life by decades (Teeraananchai et al., 2017). ART also prevents HIV transmission between sexual partners (Cohen et al., 2011) and from mother to child during pregnancy, childbirth and breastfeeding (Siegfried et al., 2011). At the community level, the availability of ART has improved mental health, and increased savings and human capital investment (Baranov et al., 2015; Baranov and Kohler, 2018). For these reasons, scaling up ART treatment has become central to public health policy.

ART is available in Malawi at no cost to patients. HIV clinics typically operate within larger clinics or hospitals. ART first became available in the early 2000's, with accessibility improving over time. Eligibility for treatment has expanded following changes in World Health Organization (WHO) guidelines, with implementation by the national Ministry of Health simultaneously across all clinics in Malawi (Harries et al., 2016). Initially, only those with advanced HIV (those with severe symptoms or a weak immune system as determined by a blood test) qualified for treatment. Since 2016, Malawi has adopted a universal test-and-

[5] Source: UNAIDS (2022).

treat policy: anyone diagnosed with HIV is encouraged to initiate treatment immediately. Patients must return to the clinic every few months for health assessments and to refill their prescriptions. The primary purpose of these return visits is to obtain ART medication. Some patients also benefit from changes in treatment regimen or treatment of comorbidities, but these impacts are likely to be marginal relative to the life-saving impact of the medication itself.

Approximately 79 percent of Malawians living with HIV are currently enrolled in ART programs, but retaining patients in care is a challenge. Patients and their families may avoid care for reasons that include (i) fear of stigmatization or relationship consequences, (ii) fear of healthcare or hospitalization, (iii) the need for frequent clinic visits, which impede the fulfillment of work and family responsibilities, (iv) concerns that children who are diagnosed as HIV-positive will die, be subjected to complex treatments, or be burdensome to the family, (v) aversion to side effects, and (vi) a lack of knowledge of the benefits of treatment (Genberg et al., 2015; Hoffman et al., 2017; Derksen et al., 2022; Gallant, 2023).

To prevent prolonged lapses in care, clinic staff formally engage in patient tracing: they track patients down by telephone or at their home address if they miss an appointment. Tracing patients who are lapsed from care is official health policy in Malawi (Ministry of Health, 2011). Staff in Malawian clinics are expected to review patient records regularly and to enter lapsed patient details into a paper-based Ministry of Health tracing register. A patient is formally classified as lapsed (or "defaulted") two months after the missed appointment date. Staff are instructed to attempt to contact patients any time from two weeks after a missed appointment, and counsel them to return to the clinic to continue treatment. The Ministry of Health guidelines acknowledge that tracing is very expensive and time-consuming, and HIV patients who have already initiated ART should be prioritized (as opposed to those ineligible for ART). This practice is intended to prevent community spread and save lives. In particular, proper adherence to ART is critical in preventing cancer and tuberculosis (TB), two leading causes of death for HIV patients (Williams and Dye, 2003; Yanik et al., 2013).

2.2 Electronic Medical Records

In 2005, the Ministry of Health began the process of migrating Malawian HIV clinics from paper-based HIV patient records to EMRs. Prior to that, all clinics relied on paper records to record HIV patient data. The standardized paper record included information on HIV testing, ART initiation, prescription refills, and death. The EMR system is also standardized across Malawian clinics, and is used almost exclusively for ART patients.[6] Patient records are entered and accessed on a touchscreen device at the point of care. The EMR rollout was managed by Baobab Health Trust, a Malawian NGO. The order of implementation of EMRs for HIV care across clinics in Malawi was determined primarily by the number of ART patients (Douglas et al., 2010). EMRs were first introduced in a large hospital in 2007, and were rolled out to 124 clinics by 2019 (out of more than 700 HIV clinics total).[7]

The EMR system is designed to increase the efficiency and accuracy of data management for clinic staff. The system enables staff to submit automated queries to the database. It also simplifies the retrieval of existing patient records during clinic visits and streamlines data entry.

In particular, the EMR system makes certain types of automated queries available to clinic staff. First, staff are able to view the list of scheduled patient refill visits on a given day. Staff can also generate a list of lapsed patients, with contact information, who are at least two months late for an ART appointment.[8] Finally, the system produces reports on aggregate patient numbers (including the number of patients lapsed and the number of deaths) for the Ministry of Health.

The system also ensures the continuity of patient care by linking each patient with the

[6] In some clinics, EMRs are created for HIV positive patients who have not yet initiated ART.

[7] For a more detailed description of the EMR system and rollout see Douglas et al. (2010), as well as data from qualitative interviews in Table A1. Table A1 contains extracts from 50 semi-structured interviews with clinic staff (nurses, physicians, data clerks, coordinators and tracing staff), Ministry of Health employees, NGO collaborators, and local community leaders, identified through a snowball procedure. A TED Talk video explaining the development of the Malawian EMR system is available at https://www.ted.com/talks/soyapi_mumba_medical_tech_designed_to_meet_africa_s_needs, last accessed on August 28th, 2024.

[8] The tracing list only includes HIV patients who had previously initiated ART.

patient's established record and medical history, even if the patient returns to care after a long absence or appears at a different clinic than the one at which the original diagnosis was made. A barcode on the patient's "health passport," a laminated card carried by the patient, allows staff to access a medical record instantly. Even if the patient does not present a health passport, the EMR system can quickly recover a record based on the patient's name, and allows for common alternate spellings.

The EMR system is also designed to improve adherence to protocolsduring a patient visit. The system alerts healthcare staff to medical issues, prompts the staff member to collect specific biometrics or conduct additional tests, and guides the dispensation of medication, ensuring adherence to up-to-date guidelines.[9] Importantly, the system prompts staff to screen patients for TB, one of the leading causes of sickness and death in those with HIV/AIDS (Williams and Dye, 2003). The system also prompts staff to order a CD4 count (a blood test that measures immune system health) and make referrals to other services. Finally, the system flags underweight patients, defined as having a body mass index (BMI) more than 2 standard deviations below the median (according to a distribution provided by the World Health Organization). These patients are supposed to receive extra nutritional support, with priority given to children and adolescents. The specific guidelines and procedures proposed by the EMR system prompts match those used in paper-based clinics. Staff are required to enter similar patient data regardless of whether they are using the EMR system or a paper form.

To save on costs, the Ministry of Health and Baobab Health Trust opted to implement an EMR system that would not require hiring additional staff (Douglas et al., 2010). A touchscreen workstation at the point of care was preferred because it was more user-friendly than alternatives. Regular clinic staff were trained to operate the system and record patients' information during a half-day-long orientation session (Driessen et al., 2013). Other than initial training and installation costs (detailed in Table A2), the adoption of the EMR system

[9] Using data from the US healthcare system, Abaluck et al. (2020) finds that improving adherence to guidelines can have a meaningful impact on health outcomes. A Youtube video illustrates the entire visit of an ART patient to a clinic in Malawi where the EMR system was implemented, available at https://www.youtube.com/watch?v=iaDGMufLfC0. Last accessed on August 28th, 2024.

did not increase labor costs in adopting clinics. In some clinics, the introduction of the EMR system coincided with the introduction of a power source, and with computer literacy training for staff. The EMR rollout was not combined with any separate policies to increase the number of clinical staff, tracing staff or other general resources (Douglas et al., 2010; Driessen et al., 2013).

It is important to note that the EMR rollout did not impact ART supply. There has never been a national stock-out of ART medication in Malawi, and stock-outs at individual clinics have been rare and short-lived (Harries et al., 2016). As described in Section 2.1, changes in ART eligibility were determined by WHO guidelines and implemented in all Malawian clinics simultaneously.

3 Data and Empirical Approach

3.1 Data Sources and Quality

Our analysis relies on a large proprietary dataset containing individual de-identified data on HIV patients obtaining ART from clinics that implemented the EMR system between 2007 and 2019 (see Figure A1 for a map of clinics).[10] We include clinics in the sample from one year after they begin offering ART. The panel is unbalanced, with clinics entering at various times between 2003 and 2012. The clinics maintain detailed administrative records for all ART patients, including on initiations, prescription refill visits, and deaths. The same basic patient data is collected at the point of care in clinics with and without EMRs. Upon EMR implementation, the history of a patient's paper record is captured electronically as existing data is digitized into the EMR system, including past initiations, deaths, and the most recent prescription refill visit. This pre-EMR data is required in order for clinics to perform basic functions, such as accessing patient records during a follow-up visit or generating a tracing list.

We are able to confirm that the administrative data is of high-quality and is comparable pre- and post-EMR. To use the EMR system, it was necessary for clinics to digitize their

[10] The data does not include information on other patients at the clinic, and does not contain complete or consistent information on HIV patient visits prior to ART initiation.

existing paper-based patient records. Indeed, our data includes detailed and complete information on all patient initiations, pre- and post-EMR, including for patients who never visit the clinic post-EMR. 52 percent of patient identifiers appear only in the pre-EMR period. We observe an initiation entry for every patient who subsequently visits the clinic. If a patient switches clinics, they will typically retain the same patient identifier. We observe switches for approximately 3 percent of patients in our sample. This is an underestimate of the true incidence of switching, however, as some patients may move to different clinics and adopt new patient identifiers. An HIV tracing study in Zambia found that 14 percent of patients flagged for tracing were in fact receiving care at a different clinic (Beres et al., 2021). While we cannot link these patients to previous identifiers, we can identify them as new or returning patients, based on whether they report having ever taken ART. In the analysis below, we show that the findings are robust to clinic switching (see Section refsec:resultslapses).

At the time of EMR adoption, clinics did not digitize the entire history of ART refill visits, but rather digitized information on the most recent pre-EMR visit for each patient. This history was included even if a patient never visited a clinic after EMR adoption. We are able to confirm that this digitization of pre-EMR visits took place, and note that this was necessary for the EMR system to function properly.[11] We observe at least one pre-EMR visit (or initiation) for every patient who obtains a prescription refill post-EMR. We also observe patients in every clinic who refill a prescription pre-EMR but never return post-EMR. There is no discontinuity in aggregate clinic visits in the quarters immediately preceding and immediately following EMR adoption (see Figure A3).

Patient deaths are also recorded consistently in the data, both pre- and post-EMR. We associate each patient death with the clinic at which the patient initiated care. A patient death can be recorded in one of two ways. First, a guardian, family member, clinic worker or community health worker may report the death, as well as the date of death, to the clinic. Otherwise, the date of death may be discovered when clinic staff attempt to trace the patient. Upon discovering that a patient has died, staff update the patient record with the date of death. This implies that some deaths that occurred pre-EMR were discovered post-EMR.

[11] In particular, it was not possible for the system to generate a list of lapsed patients without data on the most recent patient visit.

Regardless, these discovered deaths, including those which occurred pre-EMR, appear in our data. In fact, all of the clinics in our sample had adopted EMRs at the time of data collection, so the recording of retrospective death dates should have already taken place. The comparability of the recording of patient deaths before and after EMR implementation is further supported by the fact that we do not observe a spike in deaths after EMR adoption (see Section 4.1). Nonetheless, to further guard against possible measurement error, we exclude the most recent year of data from our main analysis to ensure that we are not missing deaths that are yet-to-be discovered. We also include a robustness test in which we drop late-adopting clinics (see Section 4.5). However, because patients are never recorded to have died without verification, deaths may be slightly under-counted overall. Reassuringly, the death rates that we observe for patients with advanced HIV are broadly consistent with the biomedical literature.[12]

3.2 Patient Lapses and Deaths

In our sample, we observe a total of 755,175 patient initiations and 55,181 patient deaths. As of 2018, a total of 358,843 patients were actively in care (i.e., visited a clinic at least once that year, see Figure 1). Because the typical prescription is for between one and six months, we consider a patient that has not visited the clinic for at least one year to be lapsed. Table A3 reports clinic-level averages including number of patients actively in care, new patients, returning patients, and patient deaths. The panel shows that, from 2008 to 2018, patient deaths fell steadily.

[Figure 1 about here.]

Patients frequently miss appointments for ART refills and lapse from care. Figure A2 reports missed appointments, lapse rates, and death within 2 years of initiation by gender and age.[13] 17 percent of patients lapse immediately after initiating ART. On average, fully

[12] In our data, patients with advanced HIV, as defined by the WHO, who lapse from care for at least 2 years face a 20 percent risk of death. This is consistent with estimates from Raffetti et al. (2016).

[13] Lapse rates may be measured with error as some patients might return to care under new identifiers or at different clinics (Beres et al., 2021) We discuss this further, and present

half of the patients within a clinic had lapsed from care at the time of EMR implementation. Younger adults (under age 50) are more likely to lapse, possibly due to HIV-related stigma and its consequences for dating and marriage (Derksen et al., 2022). At the time of EMR implementation, one-fifth of lapsed patients had died. Overall, 6 percent of patients die within two years of initiating ART, and 5 percent of patients die before age 30.

Children are particularly likely to lapse from HIV care and to die. 23 percent of children under age 10 lapse immediately after initiation, and 57 percent are lapsed at the time of EMR implementation. 7 percent of children under 10 die within 2 years of initiation. In Table A4, we use a multivariate regression to predict patient lapse pre-EMR by the number of months lapsed. For every duration of lapse, we find that children under the age of 10 are much more likely to lapse than patients in other age groups. Other individual-level predictors of patient lapse are gender (with male patients more likely to lapse), HIV stage at initiation, and years since initiation. Lapse from care is more common in hospitals, urban clinics, and small clinics.

3.3 Empirical Approach

To estimate the impact of EMR implementation on patient outcomes, we conduct an event study with year and clinic fixed effects (two-way fixed effects, or TWFE). We examine outcomes including patient deaths, lapses from care, the number of new and retained patients, health outcomes, and measures of medical care provided during visits. We investigate heterogeneity by gender, age, and clinic type. In practice, full EMR implementation takes several weeks. We define as our event the date from which all patient information was entered directly into the EMR system as a matter of routine clinical practice. In each year, we compare treated clinics (which had already implemented EMRs) to not-yet-treated clinics (which had not yet implemented EMRs). Not-yet-treated clinics serve as our control group (Abraham and Sun, 2020). Year fixed effects control for changes in ART eligibility criteria, which evolved at the national level, and clinic fixed effects control for level differences between treated and not-yet-treated clinics.

results that do not rely on patient identifiers and are robust to this type of error in Section 4.2.

For the TWFE estimator to be causal, three assumptions must be satisfied: parallel trends, no pre-treatment effect, and treatment-effect homogeneity by clinic (Abraham and Sun, 2020). The first assumption would fail if EMR adoption were endogenous to clinic performance. Policy documents as well as quantitative and qualitative evidence show that the criterion for EMR implementation was primarily the total number of patients enrolled at the clinic (Douglas et al., 2010), which is independent of the prevalence of lapsed patients and patient deaths.[14] Indeed, Table A3 shows a higher number of initiations among clinics that adopted earlier. We include a clinic fixed effect in our main analysis, which accounts for differences in clinic size, and confirm our results using various robustness tests. These include allowing for heterogeneous treatment effects by clinic size, and interacting time fixed effects with clinic size.[15] We also conduct an additional two-way fixed effects analysis in which clinics that adopt EMRs at the end of the study period (in 2019) serve as a pure control group.

Several recent papers have shown that, in staggered implementation designs, TWFE estimates, including estimates of pre-trends, may be biased due to heterogeneous treatment effects (Goodman-Bacon, 2018; Abraham and Sun, 2020; Callaway and Sant'Anna, 2020; de Chaisemartin and D'Haultfoeuille, 2020; Baker et al., 2021; Borusyak et al., 2021). To address this, we estimate time-varying treatment effects, which partially alleviates this concern (Goodman-Bacon, 2018; Abraham and Sun, 2020). We also perform additional analyses using four alternative estimators that are robust to heterogeneous effects across time and clinic (Abraham and Sun, 2020; Callaway and Sant'Anna, 2020; de Chaisemartin and D'Haultfœuille, 2020; Borusyak et al., 2021). Section 4.4 describes heterogeneity by clinic type, including by clinic size.

[14] For illustration, we provide two quotes from different informants explaining why particular clinics received the EMR system. *The cohort size of people who were on ART during that time was high. It was more than 2000. So the ministry produced a guideline that EMRs should be introduced to high volume sites and [Clinic Name] was one of them.* – Physician in Charge.

It was dependent upon cohort size the facility has. Like how many [ART] clients a facility has. Most facilities have a cohort size of more than 2000, so they received the EMRs. – District Technical Officer at Partner NGO.

[15] Including clinic-specific time trends leads to perfect multicollinearity, which means that we cannot include such trends as controls.

For outcomes measured at the clinic-year (or clinic-quarter) level, we estimate both average and dynamic treatment effects. Clinic-level outcomes include the number of patient deaths, the total number of patients active in care (i.e., visiting the clinic in a given year), patient retention (i.e., the number of patients who return to the clinic after having initiated in a previous year), and the number of new patients (i.e., those taking ART for the first time). To better understand the overall burden on the clinic, we also measure the number of patients tested and treated for TB, the number of patient referrals, and the number of underweight patients.[16]

We estimate average treatment effects using the following specification:

$$y_{ct} = \alpha_c + \lambda_t + POST_{ct} + \varepsilon_{ct} \tag{1}$$

Here, y_{ct} is the dependent variable for clinic c and year (or quarter) t. α_c is a clinic fixed effect, and λ_t is a year fixed effect. The dependent variable is the logarithm of the outcome. We take the logarithm in order to estimate elasticities, and to account for the fact that clinics vary in size and that outcomes are therefore skewed (see Figure A4). We add one to each variable to include clinics that are operating but have an outcome equal to zero.[17] In the appendix, we report results using the inverse hyperbolic sine transformation (Burbidge et al., 1988), and we discuss other ways of incorporating zero values in Section 4.5. We cluster standard errors at the clinic level (Cameron and Miller, 2015).

We also allow for time-varying treatment effects to capture both the immediate and the long-run effect of EMR implementation on outcomes. The dynamic estimation equation is as follows:

$$y_{ct} = \alpha_c + \lambda_t + \sum_{\substack{k=-6 \\ k\neq -1}}^{5} \tau_k \mathbf{1}\{t - E_c = k\} + \tau_6 \mathbf{1}\{t - E_c \geq 6\} + \varepsilon_{ct} \tag{2}$$

[16] We use the same definition of underweight as the EMR system: having a BMI more than 2 standard deviations below median, according to World Health Organization statistics (https://www.who.int/toolkits/child-growth-standards/standards/body-mass-index-for-age-bmi-for-age, https://www.who.int/tools/growth-reference-data-for-5to19-years/indicators/bmi-for-age).

[17] Zero deaths occur in fewer than one percent of clinic-year observations. For patients in care, the minimum observation is 160.

Here, k is the number of years since the event (EMR implementation), where the event occurs in clinic c at the beginning of year $t = E_c$ or $k = 0$. τ_k is the effect size k years after the event. Because all clinics adopt EMRs at some point, and because of imbalance in our panel, we exclude several time periods from the specification to avoid multicollinearity. In addition to the pre-period year $k = -1$, we elected to exclude all early pre-period years ($k < -6$) from the estimation, which allows data from these years to serve as a control group and increases statistical power. We also assume a constant long-run effect, and estimate a single coefficient, τ_6, for all periods greater than or equal to 6 years post-treatment. For outcomes related to clinic visits we are not able to estimate pre-trends because we only observe the patient's most recent visit prior to EMR (in addition to all initiation visits, and all visits post-EMR). For these outcomes, the year prior to EMR introduction serves as a control group.

Some outcomes are measured at the individual patient-year and patient-visit levels. At the patient-year level, we record whether a patient is lapsed (no visit for at least m months) as measured at the end of the year. At the patient-visit level, we record whether a patient misses their subsequent appointment by at least 60 days (the Ministry of Health's definition of a lapse from care). We also record patient health measures including viral load and an indicator representing whether the patient is underweight. Finally, we capture measures of the medical care provided during new patient visits at the patient-visit level by assessing whether (i) the patient is tested and treated for TB, (ii) the patient is referred for other services, (iii) a CD4 count is ordered, and (iv) an initiating child is effectively treated for malnutrition. For this last outcome, we restrict the sample to children who begin treatment while underweight and who visit the clinic for a follow-up visit within 3 to 6 months. Effective treatment for malnutrition is recorded if the child is no longer underweight upon the second visit.

For each of these patient-level outcomes, we estimate the following equation:

$$y_{ict} = \alpha_c + \lambda_t + POST_{ct} + \boldsymbol{\chi}'\mathbf{X_i} + \varepsilon_{ict} \tag{3}$$

Here, i represents an individual patient, and t represents the year (or, for some outcomes, the visit). We include a set of patient covariates $\mathbf{X_i}$ and cluster at the clinic level. To compare similar patients across EMR and non-EMR clinics, we control for gender, age category, years since initiation, an indicator for whether the patient was underweight at the time of initiation, and an indicator for whether the patient had advanced HIV at the time of initiation. In the appendix, we also report results without individual-level covariates.

4 Results

4.1 Patient Deaths

The introduction of EMRs causes a decrease in the number of patient deaths, particularly among children. In Panel A of Table 1, we report estimates of the impact of EMRs on the number of patient deaths using the TWFE specification at the clinic-year level (equation 1) as well as the alternative difference-in-difference estimator proposed by Borusyak et al. (2021). The TWFE estimates indicate an 11 percent decrease in patient deaths overall ($p < 0.1$), with similar results for male and female patients, and a 19 percent decrease in deaths among children under age 10 ($p < 0.01$).[18] The Borusyak et al. (2021) estimators indicate larger impacts: a 20 percent reduction in deaths overall ($p < 0.01$), and a 29 percent reduction in deaths among children under age 10 ($p < 0.01$). This discrepancy may be due to bias in the TWFE estimator, which could occur, for example, because of dynamic treatment effects or treatment effect heterogeneity.

[Table 1 about here.]

We do in fact find evidence of dynamic treatment effects; we observe a gradual decline in deaths over time among patients enrolled in EMR clinics (see Figure 2 and Table A5). Based on the TWFE estimates (2), we observe a 6 percent decrease in patient deaths in the year immediately subsequent to EMR implementation, and a significant 28 percent decrease after five years ($p < 0.01$). The effect is balanced across genders over the long run. We do not observe significant pre-trends for any gender or age category (see F-tests in Tables

[18] We report estimated effect sizes by translating log points into elasticities, following Kennedy et al. (1981).

A5 and A22). Because of potential bias in the dynamic TWFE estimator, we confirm these results using four robust estimators (Abraham and Sun, 2020; Callaway and Sant'Anna, 2020; de Chaisemartin and D'Haultfœuille, 2020; Borusyak et al., 2021), as well as in a specification where all late adopters (clinics who adopted EMRs in 2019) are classified as pure control clinics. The estimated effects are presented in Panel C of Figure 2. No pre-trends appear, and the results closely resemble our main findings.

[Figure 2 about here.]

The negative effect on deaths is concentrated among patients under the age of 50, and is particularly large for children under age 10 and for adults aged 18 to 49 (see Panel B of Figure 2). The implementation of EMRs decreases deaths among children under age 10 by approximately 14 percent in the first year ($p < 0.1$), and by approximately 44 percent after five years ($p < 0.01$). These results reflect the patterns we observe in Figure A2, which shows that younger patients are more likely to lapse from care, and are therefore more likely to benefit from improved tracing. The impact on deaths is more pronounced for younger children as opposed to adolescents. In fact, absent treatment, HIV acquired in infancy is particularly deadly (Newell et al., 2004). In our sample, children under the age of 10 (as compared to children aged 10 to 17) are 21 percent more likely to have advanced HIV at the time of ART initiation, and are 58 percent more likely to die.

4.2 Patient Lapses and Retention in Care

Patients are less likely to be lapsed from care following the implementation of EMR. Using equation 3, we estimate the impact of the EMR system on an indicator variable that represents whether an individual patient i has been lapsed for at least $m \in \{6, 12, ..., 36\}$ months, as measured at the end of year t. We find a significant ($p < 0.01$) reduction in lapse probability for every lapse duration m, with estimates ranging from -4 to -6 percentage points (Figure 3 and Tables A6, A7, and A8). The effect is similar for male and female patients (Panel A of Figure 3), and slightly larger for children and adolescents, with estimated impacts between -4 and -9 percentage points (Panel B of Figure 3). Indeed, children and adolescents are more likely to lapse from care and require tracing (see Figure A2 and Table A4).

[Figure 4 about here.]

Individual-level patient lapses may be subject to misclassification errors. In some cases, a patient might return to care after a long absence and, against policy, be assigned a new patient identifier. This is more common in non-EMR clinics, and when patients switch clinics, as it is more difficult for staff to match existing patients to their original identifiers. We therefore also examine the total number of patients in care at the clinic level, as well as the number of new and retained patients, to confirm that the introduction of the EMR system does in fact lead to an increase in the total number of patients active in care and receiving treatment.[19] The estimated impacts of EMRs on the total number of patients actively in care (new and returning) and on patient deaths are unlikely to be affected by classification errors because these outcomes include all patients who visit the clinic, without relying on specific patient identifiers. We are able to classify a patient as new or retained based on whether the patient reports having taken ART in the past. It is possible that we, nonetheless, overestimate the number of visits if a patient visits the clinic *and* reinitiates under a new identifier in the same year. This type of measurement error would cause us to underestimate the true impact of the EMR system on the number of patients in active care, because the EMR system allows clinic staff to easily link patients to their previous identifiers. In Section 4.5 we include a robustness test using quarterly data to alleviate the concern that this type of bias may be present.

EMR implementation has an immediate, persistent, and positive impact on patient retention at the clinic level, with a larger impact for children and adolescents. We identify a significant 17 percent increase in the total number of patients visiting a clinic in the year immediately following EMR implementation ($p < 0.01$, Panel A of Figure 4). This increase grows over time, and is driven entirely by patient retention (Panel C of Figure 4). We find no significant impact on the number of new patients enrolled at the clinic (Panel B of Figure 4 and Panel C of Table 1).[20]

[19] Here, we estimate equations 1 and 2 at the clinic-year level.
[20] Tables A9, A10 and A11 contain full dynamic TWFE estimates for these outcomes.

[Figure 4 about here.]

We also find that patients who visit a clinic after the introduction of EMRs are less likely to miss their next appointment, suggesting that the initial recovery of lapsed patients can have persistent effects on patient retention (Panel A of Figure 5). We estimate a variant of equation 3 at the patient-visit level, with separate indicators for new patient visits, the first refill visit post-EMR and later visits. We find that when a patient first visits a clinic for an ART refill post-EMR adoption (as opposed to pre-EMR), they are 7 percentage points less likely to miss their *next* appointment ($p < 0.01$). This effect is similar across demographics, and grows over time after the initial post-EMR visit ($p < 0.01$, Table A12, Panel A).

[Figure 5 about here.]

This improvement in patient retention appears to improve patient health, as observed at clinic visits. We estimate equation 3 at the patient-visit level, with health outcomes including an underweight indicator (based on the patient's height, weight and age) as well as viral suppression.[21] We do find an improvement in the first measure after the introduction of EMR (Panel B of Figure 5 and Table A12), and a slight improvement in viral suppression among adolescents in the long run (Panel C). These estimates, however, must be interpreted with caution, as the patients who visit a clinic form a selected sample. For example, healthier patients might choose to avoid or delay treatment. We can to some extent control for selection by comparing the initial post-EMR visit to subsequent visits, and the impact does appear larger in the longer run, particularly for young patients. Viral suppression estimates (Panel C) are imprecise, as viral load measures are ordered infrequently (at one percent of visits), and are likely offered to a selected sample; clinics often measure viral load when they suspect poor treatment adherence. While these patterns support the interpretation that retention in care leads to an improvement in patient health outcomes, we are not able to estimate the full magnitude of the effect, as we do not observe the health status of patients while they are lapsed from care.

[21] Viral suppression is defined as having less than 200 copies of HIV per milliliter of blood. We cannot include symptoms or CD4 count, as this data is only collected at the time of ART initiation.

4.3 Mechanisms

The Malawian EMR system for HIV patients might impact patient outcomes, and in particular patient deaths, through several mechanisms. By improving data management efficiency, the system provides clinic staff with the ability to quickly identify and trace lapsed patients, as well as to retrieve accurate patient records and provide tailored counselling and services during visits. Data-specific efficiency gains may translate to broader clinic-level efficiency gains, allowing clinics to see more patients, to see patients more frequently, or to provide higher quality medical care during patient visits. Finally, the system may affect the provision of medical care directly, by prompting clinic staff to conduct diagnostic tests or offer additional services.

The reduction in deaths we observe appears to be driven by an increase in the number of lapsed patients who return to care. Indeed, ART provision is by far the most important benefit of a clinic visit, because without treatment, AIDS is fatal. We find a reduction in patient deaths accompanied by a significant decrease in patient lapse rates and an increase in the number of patients in active care (see Figures 2, 3 and 4). These effects are strongest for children under the age of 10, who are at the highest risk of lapse (see Figure 3 and Table A4). This is consistent with the fact that the system makes it feasible for staff to accurately and automatically identify individual lapsed patients. The system also allows staff to produce accurate reports showing the total volume of lapsed patients for submission to the Ministry of Health. Clinic staff may therefore be particularly motivated to trace patients back into care. They may also have more time to do so, as they spend less time searching through patient records.

Without an EMR system, it is difficult for clinic staff to identify lapsed patients (see qualitative data in Table A1) despite the general practice of maintaining accurate and organized records. At the time of EMR implementation, the average clinic had over 3,000 ART patients enrolled. It is challenging to organize records in a way that both facilitates quick access for returning patients, and enables health workers to quickly identify lapsed patients. Furthermore, HIV patients typically do not adhere to a precise appointment date, and simply visit the clinic in the few weeks before or after their medication runs out. So, organizing records by appointment date, as opposed to name, would make day-to-day clinic operations difficult. Even paper-based clinics are expected to keep a record of lapsed patients in tracing registers. But because the tracing list is not accurate unless updated continuously, and because catching a missed appointment requires searching through thousands of paper-based patient records, tracing may not occur systematically under the paper-based system. Celhay et al. (2019) find that health workers may not adopt effective outreach strategies if the cost of effort is too high. Qualitative interviews with staff at various clinics and levels describe significant improvements in effective patient tracing and retention after the adoption of the EMR system (Table A1).

Patients who return to care under the EMR system are more likely to attend their next scheduled appointment (Figure 5), suggesting improved patient counselling as a potential second mechanism. The EMR system improves continuity of care by allowing staff to easily link a returning patient to their existing record even after a long absence. By accessing the existing patient record, clinic staff can offer informed counselling based on the patient's history. In particular, if clinic staff are aware of past lapses from care, they may effectively encourage the patient to remain active in care going forward. The fact that the EMR system supports staff to return lapsed patients *sooner*, and ensure consistent care thereafter, may be key to explaining the reduction in patient deaths. In a study of HIV tracing efforts in Zambia, Beres et al. (2021) find that while nearly half of traced patients return to care within one month of tracing, many non-traced patients do later return to care, and the impact of tracing, while positive, is not significant two years after lapse. The fact that patient records are linked electronically may also allow clinic staff to offer tailored diagnostics and services, though we do not find evidence in support of this mechanism (see below for an analysis of the medical care provided at visits).

A third potential mechanism relies on broader clinic-level efficiency gains, as clinical staff are able to devote more time to patient care, both by attending to more patients and by increasing the duration, frequency or quality of clinic visits. However, these gains may be offset by the fact that EMR clinics contend with a larger number of patients actively in care, as previously lapsed patients were traced, returned to care, and retained in greater numbers.

We find that broader efficiency gains did arise as clinics saw an increased number of patients after EMR implementation with the same personnel and opening hours as before EMR implementation. Shortly after EMR adoption, the number of ART patient visits per quarter rose sharply despite no change in the number of staff or in the number of days per month that the HIV clinic was open (Panel A of Figure 6). Here, we regress outcomes, measured at the clinic-quarter level, on the number of quarters post-EMR adoption. We do not include any pre-EMR periods, as we cannot observe the total number of visits or staffing levels pre-EMR.[22] We therefore include only clinic fixed effects, and observe changes in outcomes over the short-term after the first post-EMR period. We find that the total number of patient visits per quarter rises steadily, with no significant change to the number of staff nor the number of operating days. Over time, clinics do slightly increase the number of days on which they schedule refill visits. Indeed, some clinics set aside certain days of the week to schedule ART refill visits. An increase in this number of "refill days" may reflect a need to adapt to a higher number of returning patients, or an increase in capacity driven by efficiency gains at the clinic.

The frequency of scheduled visits did not increase with EMR implementation. In fact, EMR clinics appear to *reduce* the frequency of patient visits and offset the increase in the number of patients in care by increasing the dispensing interval for the average patient. In Column 1 of Table 2, we estimate equation 3 at the patient-visit level, where the dependent variable is days of medication provided.[23] The introduction of EMR leads to a 4.6-day increase in the dispensing interval. At the clinic level, the dispensing interval rises slightly in the year leading up to full EMR implementation, rises further in the following year, and remains higher five years later (Panel B of Figure 6).[24]

[22] Recall that only the last pre-EMR refill visit is recorded in the data. In EMR clinics, staff scan cards to confirm their identity during patient visits. We do not have accurate data on staff for pre-EMR visits.

[23] The sample is restricted to include only ART initiation visits, as the patients who visit the clinic for prescription refills post-EMR form a selected sample. For example, clinics typically schedule more frequent visits for patients who previously lapsed from care. We estimate this specification without covariates in Table A13.

[24] We do not interpret the significant effect in event-year $k = -1$ as a pre-trend but rather as an anticipation effect; clinic staff scheduled appointments slightly further apart while undergoing EMR system training and while preparing for implementation. This preparation period does not appear to have affected new patient initiations nor death rates (see pre-trend estimates in Tables A5 and A10).

[Table 2 about here.]

[Figure 6 about here.]

This suggests that despite the efficiency gains brought about by the EMR system, clinics adjust their operations to address the increase in patient retention. The change in clinical practice to extend the time between visits may be due, in part, to the confidence that lapsed patients can be traced effectively. Clinics typically operate on the assumption that patients are more likely to be retained in care if visits are scheduled frequently. Yet, Hoffman et al. (2021) have shown that an adjustment to lengthen the dispensing interval does not necessarily affect patients adversely. In our data, there is substantial variation in average dispensing intervals across EMR clinics (Figure A10). However, clinics with longer intervals between patient visits do not administer fewer lab tests or referrals per-patient, and in fact have lower rates of patient lapse (Table A23). Clinics appear to take individual patient characteristics account when scheduling appointments, including gender, age, health and history (Table A24). The first appointment after initiation is typically scheduled with a shorter interval than later appointments, with 85 percent of new patients given an appointment within one month. Controlling for patient characteristics, new patients who are given a very short interval (two weeks) are slightly less likely to lapse from care, but one month is not better than two or more (Panel A of Figure A11). Once a patient is established in care, longer spacing intervals of three or four months appear to be optimal (Panel B of Figure A11). This implies that longer dispensing intervals may in fact increase adherence for some patients.

While the efficiency gains associated with the EMR system facilitated clinical care for a larger volume of patients, we should not interpret our results as *driven by* an increase in overall clinic capacity. It is highly unlikely that established patients, for whom adherence is a clinical priority, would be turned away from the clinic even prior to EMR implementation. Clinic capacity can typically be managed by extending the time between visits rather than turning away patients altogether. Indeed, Column 1 of Table 2 shows that clinics do manage capacity in this way after the introduction of EMRs. Moreover, among clinic-years in our sample, the median time between initiation and the first scheduled visit is never more than one month, suggesting that clinics do manage to treat new patients in a timely manner.[25]

A fourth mechanism that may explain outcomes involves the medical care provided to patients during visits. Indeed, the EMR system is designed to support clinic staff in adhering to medical protocols. The system flags underweight children who require nutritional support. It also prompts staff to test for TB, a common HIV comorbidity, to refer patients for additional lab tests and health services, to measure CD4 counts, and to follow ART initiation guidelines. These prompts, in addition to broad efficiency gains, may lead to differnt levels of care provision.

We are able to test whether the EMR system prompts lead to a changein the medical care offered at each patient visit. We again estimate equation 3 at the patient-visit level, with dependent variables including whether an initially underweight child shows improvement at a follow-up visit, and whether staff test and treat an initiating patient for TB, offer a referral, or order a CD4 count.

First, we restrict the sample to children who initiate before age 18, are underweight at the time of initiation, and return for a refill three to six months later.[26] We compare children who initiate before versus after the EMR system is introduced. The outcome is an indicator for whether the child remains underweight at the follow-up visit.[27] This allows us to estimate the impact of the EMR system on a child's weight *conditional* on the child's retention in care within six months months. We find no significant impact on the child's likelihood of being underweight at the follow-up visit, and the point estimate is in fact positive (Column 3 of Table 2). This suggests that while the system's ability to flag underweight children may lead to action on the part of the clinic, it does not lead to meaningful weight gain for the child before the next clinic visit. This also suggests that the impact on child weight we documented in Panel B of Figure 5 is driven by increased compliance with a schedule of clinic visits and adherence to ART, as opposed to other medical care received *conditional* on attending a visit.

Second, we show that staff in EMR and non-EMR clinics are equally likely to offer

[25] In 99 percent of clinic-years, the median time between visits is one month or less.

[26] We must also exclude patients who initiate more than six months pre-EMR as we may not observe the follow-up visit.

[27] We use the last visit within a six-month period post-initiation.

additional health services at a given patient initiation visit. We focus on three outcomes: (i) the diagnosis and treatment of TB, (ii) all referrals, including lab tests and referrals to other departments (including for TB), and (iii) whether the clinic orders a CD4 count. For all three outcomes, we find that the EMR system has a small and statistically insignificant impact, with point estimates often negative (Columns 5, 7, and 9 of Table 2).

Finally, the EMR system did not lead to significantly higher patient enrollment. During patient registration, EMR prompts are designed to ensure that patients initiate ART if and only if they qualify based on the most recent guidelines (based on their HIV stage or CD4 count). While the number of new patients rises slightly after EMR introduction, this increase is not statistically significant (see Panel B of Figure 4).

In sum, we do not observe an effective increase in the number of additional services delivered at each patient visit. The relative ineffectiveness of EMR prompts may be due to the fact that, in our context, clinics see a much higher number of patients post-EMR. If clinics were instead able to leverage EMR efficiency gains to devote more time to each patient visit, they may be better able to respond to EMR prompts, and thus offer additional medical care.

4.4 Heterogeneity by Clinic Type

We next explore heterogeneity by clinic type at the clinic level. We estimate variants of equations 1 and 2 in which post-EMR indicators are interacted with clinic type. We compare hospitals to regular clinics, urban clinics to rural clinics, and large clinics to small clinics.[28]

The effect of the EMR system on patient deaths is similar across clinic types, with one exception: the reduction in deaths is more pronounced in hospitals (Column 1 of Table 3). In hospitals, this reduction becomes statistically significant beginning two years after the introduction of EMRs (Panel A of Figure A9). This heterogeneity may be explained by the fact that hospitals are more likely to treat young children as well as patients who have advanced HIV (Panel B of Figure A9), and that hospital patients are more likely to lapse

[28] Large clinics are defined as clinics with an above-median number of initiating patients as measured in 2013. We do not estimate pre-trends for this analysis to improve statistical power.

from care (Table A4). An alternative explanation is that hospitals, unlike regular clinics, are able to offer additional medical care post-EMR, for example by referring patients to other hospital departments as required. However, our findings do not support this explanation. Hospitals are not more likely than regular clinics to (i) improve outcomes for underweight children, (ii) test and treat TB, (iii) refer patients for additional services, or (iv) order CD4 counts (Table 2). Hospitals may be better equipped to deal with the influx of returning patients post-EMR. Hospitals space patient visits by an additional 3 days (rather than 6) after the introduction of EMR, but this discrepancy between hospitals and regular clinics is not statistically significant (Column 2 of Table 2).

[Table 3 about here.]

The increase in patient retention we observe is also similar across different types of clinics (Column 4 of Table 3). Indeed, the patient tracing protocol is standardized across Malawian clinics regardless of clinic type (Ministry of Health, 2011). It is perhaps surprising that effects are similar for urban and rural clinics, though slightly larger for urban clinics. There is one important difference between rural and urban clinics: urban patients are 6 percentage points more likely than rural patients to be lapsed for 12 months or more and thus to require tracing (Column 2 of Table A4). One potential explanation for this is that urban clinics afford anonymity to patients travelling from rural communities, who incur travel costs and therefore lapse from care. Patients visiting urban clinics for privacy reasons may also lapse due to fear of stigma. To locate lapsed patients, rural clinics typically rely on strong local networks, patient guardians and community health workers to locate lapsed patients. Urban clinics are more likely to rely on mobile phones and easy local transportation. In both cases, the EMR system allows clinics to *identify* lapsed patients more efficiently, before taking steps to locate the patients and return them to care. It is less surprising that tracing efficiency improves in both large and small clinics, as the number of dedicated tracing staff is scaled according to patient volume. While patients lapse more often in small clinics, the magnitude of the difference is small (1 to 2 percentage points, Table A4).

4.5 Robustness

Our first robustness test involves testing for pre-trends on the number patient deaths and the number of new patients for each gender and age group.[29] We test for pre-trends using standard F-tests in the TWFE specification (Tables A5 and A10) as well as the robust F-tests recommended by Borusyak et al. (2021) (Table A22). The p-values are insignificant, with one exception. We find a significant pre-trend on the number of new patients aged 10 to 17 in the standard TWFE specification ($p = 0.03$), which appears to be driven by a slightly higher number of new adolescent patients in event-year $k = -5$ (Table A10). This may have occurred by chance; the p-value is not significant in the robust specification (Table A22).

As discussed in Section 4.2, our estimates are not explained by patients switching clinics. By aggregating at the clinic-quarter level, we also show that our measure of the number of patients in care is not greatly affected by patient double-counting. It is unlikely that a patient would visit a clinic under two different identifiers within the same few months. The estimated impact of the intervention on the total number of patients in care, whether measured at the annual or quarterly level, is similar (see Figure 4 and Figure A3).

This robustness test, along with our examination of pre-trends, also allows us to examine the possibility of measurement error in digitized administrative data. It was necessary for clinics to digitize existing patient records before they could make use of the new EMR system. Yet, we might be concerned about misplaced records, or about selective digitization of recent data. However, we observe no significant discontinuity in patient visits immediately before and after EMR adoption, nor in patient initiations or deaths in the four quarters preceding and following EMR adoption (Figure A3). In the six years leading up to EMR adoption, we find no evidence of pre-trends in terms of new patient initiations or deaths.

We perform several additional robustness tests. As described above, we estimate the impact on patient deaths using alternate non-TWFE estimators, and allowing for heterogeneous effects by clinic type.[30] We also estimate a specification with time fixed effects

[29] We can test for pre-trends on outcomes for which we have complete pre-EMR data. We cannot test for pre-trends on patient lapses, the total number of patients in care, or patient retention, because we do not have data on all pre-EMR ART refill visits.

[30] In addition to patient deaths, we can use the alternate estimators to estimate the impact on the number of new patients, as we have complete pre-EMR data for this outcomes (see

interacted with an indicator of clinic size (i.e., having an above-median number of initiating patients as of 2013), and find similar impacts (Table A17). In Table A18, we report results excluding clinics that adopted EMRs after 2017. While this affects precision, the results are consistent with those in our main specification. Finally, we explore three alternative ways of incorporating zero values for clinic-level outcomes in our analysis: using inverse hyperbolic sine (Table A19, as in Burbidge et al. 1988), adding 0.01 before taking the logarithm (Table A20), and scaling outcomes by clinic size (Table A21). The findings are similar to our primary analysis.[31]

4.6 Impact and Cost-Effectiveness

The findings presented in Section 4.1 allow us to estimate the total number of deaths averted due to the introduction of EMRs, as well as the cost-effectiveness of the EMR system. For detailed calculations, see Appendix Section A.1.

We estimate that, as of 2019, the EMR system prevented 5,050 patient deaths in total. After five years, the average EMR clinic averted 76 deaths, implying a short-run cost of approximately USD $448 per life saved, or USD $185 per disability-adjusted life year. This does not accurately reflect the long run impact of EMRs. Most of the deaths averted in the short-run were among children and young adults who are likely to live long past the initial five year period. To approximate the longer-run impact of EMRs, we calibrate and simulate a Cox Proportional Hazard model using moments from our data. We find that over the longer term, EMRs extend patient lives by approximately 1.3 DALYs. This implies a cost-per-DALY of USD $7.37.[32] These figures do not include the cost of providing ART, which is approximately USD $260 per year (Opuni et al., 2023); yet, this cost is likely offset

Figure A5).

[31] If we do not use a log-transformation, we must scale outcomes to account for the variation in clinic size (see Figure A7). If we use the number of patients actively in care as an indicator for clinic size. To avoid endogeneity concerns, we use a fixed denominator, the number of active patients as of 2017, scaled by the relevant subgroup size, and we include only clinics that adopted EMRs before 2017. The broad patterns are consistent with our primary analysis, though estimates are imprecise. The magnitudes of the coefficients are smaller as they represent percentage point changes.

[32] See Table A26 for sensitivity analysis.

by the savings arising from the reduction in new infections due to ART adherence. Indeed, an approximation based on evidence provided in the medical literature (Tanser et al., 2013) suggests a per-clinic reduction of approximately 30 new infections per year due to EMRs, which more than offsets the increased cost of providing medication to longer-lived patients.

Although we document a large relative decrease in deaths among young children, the EMR system prevented the greatest number of deaths among young adults (aged 18 to 49). This is because many more HIV patients are young adults than children (see Figure 1). As of 2019, the adoption of EMRs had prevented approximately 457 deaths among young children, 80 deaths among adolescents, and 3,812 deaths among young adult patients.

5 Conclusion

This paper analyzes the impact of an EMR system in Malawi, a low-income country that ranks 9th in the world in HIV prevalence. The EMR system improved access to patient records and facilitated patient tracing by clinic staff. It led to an immediate and persistent improvement in the recovery of lapsed patients into care. After five years of implementation, the EMR system led to a 28 percent annual reduction in deaths. The effects are more pronounced for patients under the age of 50, and are notably large for young children.

Among potential mechanisms, the primary explanation for the reduction in patient deaths is the ability of EMR-equipped clinics to manage and query patient data efficiently. The EMR system enables clinics to quickly identify and trace patients who had lapsed from care. We observe a significant decrease in patient lapse rates after the adoption of EMR, accompanied by an overall increase in patient retention and total volume of patients in care, as well as improvements in patient health measures. In interviews, clinic staff explain that the EMR system made patient tracing routinely feasible by facilitating access to patient records. Rather than sifting through thousands of paper records, clinic staff could use a system-generated list to identify patients in need of follow-up.

One policy implication of this work is that the adoption of an EMR system can deliver outsized, life-saving benefits in developing countries, where staff are overburdened and under-resourced. This highlights the potential impact of health information technology more generally – from analytics to artificial intelligence and automation – in developing countries.

Efficient patient data management, and in particular, the ability to identify and trace lapsed patients benefits society by preventing the spread of HIV and by preventing AIDS deaths. It likely also benefits individual patients directly since it lowers the risk of premature death. The tracing of children with the intention of recovering them into HIV care is especially beneficial. In 2021, approximately 85,000 children under the age of 10 died of AIDS worldwide. In our data, young children represent only 2 percent of patients in care, but 6 percent of deaths. Young children are particularly vulnerable to lapses in care. They therefore benefit disproportionately from an improved patient tracing process.

Many questions for further research arise from this analysis, including the need to develop a clearer understanding of the barriers to care for HIV patients. This will enable the design and delivery of interventions that include not only the scheduled administration of medication, but also the social contact and support that keep patients engaged in care and avert premature death. Patients might be unaware that delaying HIV treatment significantly increases the risk of mortality (Insight Start Study Group, 2015). The effectiveness of tracing implies that social pressure, or conversely, social support, is often sufficient to offset a patient's concerns about returning to care. Indeed, social dynamics can influence healthcare decisions in contexts ranging from HIV testing to childhood immunization (Godlonton and Thornton, 2012; Karing, 2018). Social pressure may also help patients overcome problems of limited self-control; procrastination has been identified as a significant barrier to HIV testing (Derksen et al., 2024; Macis et al., 2021). Patients may in fact view tracing as a form of social support: a sign that the clinic cares about them and their wellbeing. Some of our qualitative data is consistent with this view.

Several factors likely contribute to the alarmingly high lapse rate among young children specifically. Parents might find it difficult to accept a child's diagnosis, and having an HIV-infected child may carry additional stigma for the child and the entire family. Moreover, caring for an HIV-positive child is a great burden for families. Many HIV-positive children have lost one or both parents to HIV. Children cannot typically visit a clinic on their own, adhere to medication or keep track of appointments, and therefore rely on caregivers who are often busy with other work and family responsibilities.

Digital data management solutions to global health challenges may hold promise well beyond the use of EMRs in HIV care. The efficiency gains available through the implementation of an EMR system are likely to be relevant for other low-income settings. Nevertheless, the EMR system evaluated in this study was designed to address the particular challenges brought about by the HIV epidemic. HIV treatment clinics serve chronically ill patients in high volumes, with frequently scheduled follow-up visits. Lapses from care are common and often fatal. The specific magnitude of the impact of an EMR system likely depends on the particular health context. Nonetheless, the dramatic improvement in outcomes observed in the Malawian context suggests that, for other resource-constrained settings with serious health burdens, EMR implementation may have a large impact on patient health. Broadening our understanding of the benefits of digital solutions for health data management across low-income country contexts remains an important direction for future research.

References

Abaluck, J., Agha, L., Chan, D., Singer, D., and Zhu, D. (2020). Fixing Misallocation with Guidelines: Awareness vs. Adherence. *National Bureau of Economic Research Working Paper Series*, No. 27467.

Abraham, S. and Sun, L. (2020). Estimating Dynamic Treatment Effects in Event Studies With Heterogeneous Treatment Effects. *Journal of Econometrics*.

Agha, L. (2014). The effects of health information technology on the costs and quality of medical care. *Journal of Health Economics*, 34(1):19–30.

Baker, A. C., Larcker, D. F., and Wang, C. C. Y. (2021). How Much Should We Trust Staggered Difference-In-Differences Estimates? *SSRN Electronic Journal*, pages 1–65.

Banerjee, A., La Ferrara, E., and Orozco-Olvera, V. H. (2019). The entertaining way to behavioral change: Fighting HIV with MTV.

Banerjee, A. V., Duflo, E., Glennerster, R., and Kothari, D. (2010). Improving immunisation coverage in rural India: clustered randomised controlled evaluation of immunisation campaigns with and without incentives. *BMJ*, 340.

Baranov, V., Bennett, D., and Kohler, H. P. (2015). The Indirect Impact of Antiretroviral

Therapy: Mortality Risk, Mental health, and HIV-negative Labor Supply. *Journal of Health Economics*, 44:195–211.

Baranov, V. and Kohler, H.-P. (2018). The impact of AIDS treatment on savings and human capital investment in Malawi. *American Economic Journal: Applied Economics*, 10(1):266–306.

Beres, L. K., Mody, A., Sikombe, K., Nicholas, L. H., Schwartz, S., Eshun-Wilson, I., Somwe, P., Simbeza, S., Pry, J. M., Kaumba, P., McGready, J., Holmes, C. B., Bolton-Moore, C., Sikazwe, I., Denison, J. A., and Geng, E. H. (2021). The effect of tracer contact on return to care among adult, "lost to follow-up" patients living with HIV in Zambia: an instrumental variable analysis. *Journal of the International AIDS Society*, 24(12).

Bhargava, H. K. and Mishra, A. N. (2014). Electronic medical records and physician productivity: Evidence from panel data analysis. *Management Science*, 60(10):2543–2562.

Borusyak, K., Jaravel, X., and Spiess, J. (2021). Revisiting Event Study Designs: Robust and Efficient Estimation. *Work in Progress*, pages 1–48.

Burbidge, J. B., Magee, L., and Robb, A. L. (1988). Alternative transformations to handle extreme values of the dependent variable. *Journal of the American Statistical Association*, 83(401):123–127.

Callaway, B. and Sant'Anna, P. H. (2020). Difference-in-Differences with Multiple Time Periods. *Journal of Econometrics*.

Cameron, A. C. and Miller, D. L. (2015). A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resouces*, 50(2):317–372.

Celhay, P. A., Gertler, P. J., Giovagnoli, P., and Vermeersch, C. (2019). Long-run effects of temporary incentives on medical care productivity. *American Economic Journal: Applied Economics*, 11(3):92–127.

Cohen, M. S., Chen, Y. Q., McCauley, M., Gamble, T., Hosseinipour, M. C., Kumarasamy, N., Hakim, J. G., Kumwenda, J., Grinsztejn, B., Pilotto, J. H., and Godbole, S. V. (2011). Prevention of HIV-1 infection with early antiretroviral therapy. *The New England Journal of Medicine*, 365(6):493–505.

de Chaisemartin, C. and D'Haultfœuille, X. (2020). Difference-in-differences estimators of intertemporal treatment effects. *Available at SSRN 3731856*.

de Chaisemartin, C. and D'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–2996.

de Walque, D., Gertler, P. J., Bautista-Arredondo, S., Kwan, A., Vermeersch, C., de Dieu Bizimana, J., Binagwaho, A., and Condo, J. (2015). Using provider performance incentives to increase HIV testing and counseling services in Rwanda. *Journal of Health Economics*, 40:1–9.

Derksen, L., Kerwin, J. T., Reynoso, N. O., and Sterck, O. (2024). Healthcare appointments as commitment devices. *The Economic Journal, forthcoming*.

Derksen, L., Muula, A., and van Oosterhout, J. (2022). Love in the time of HIV: How beliefs about externalities impact health behavior. *Journal of Development Economics*, 159:102993.

Douglas, G. P., Gadabu, O. J., Joukes, S., Mumba, S., McKay, M. V., Ben-Smith, A., Jahn, A., Schouten, E. J., Lewis, Z. L., van Oosterhout, J. J., Allain, T. J., Zachariah, R., Berger, S. D., Harries, A. D., and Chimbwandira, F. (2010). Using Touchscreen electronic medical record systems to support and monitor national scale-up of antiretroviral therapy in Malawi. *PLoS Medicine*, 7(8).

Dranove, D., Forman, C., Goldfarb, A., and Greenstein, S. (2014). The trillion dollar conundrum: Complementarities and health information technology. *American Economic Journal: Economic Policy*, 6(4):239–270.

Driessen, J., Cioffi, M., Alide, N., Landis-Lewis, Z., Gamadzi, G., Gadabu, O. J., and Douglas, G. (2013). Modeling return on investment for an electronic medical record system in Lilongwe, Malawi. *Journal of the American Medical Informatics Association*, 20(4):743–748.

Dupas, P. (2009). What matters (and what does not) in households' decision to invest in malaria prevention? *American Economic Review*, 99(2):224–230.

Gallant, J. (2023). Time Costs and Take-up of Free HIV Care . *Working Paper*.

Genberg, B. L., Lee, Y., Rogers, W. H., and Wilson, I. B. (2015). Four Types of Barriers to Adherence of Antiretroviral Therapy are Associated with Decreased Adherence Over Time. *AIDS and Behavior*, 19(1):85–92.

Godlonton, S. and Thornton, R. (2012). Peer effects in learning HIV results. *Journal of development economics*, 97(1):118–129.

Goodman-Bacon, A. (2018). Difference-in-differences with variation in treatment timing. Technical report, National Bureau of Economic Research.

Harries, A. D., Ford, N., Jahn, A., Schouten, E. J., Libamba, E., Chimbwandira, F., and Maher, D. (2016). Act local, think global: how the Malawi experience of scaling up antiretroviral treatment has informed global policy. *BMC public health*, 16(1):1–9.

Hoffman, R. M., Moyo, C., Balakasi, K. T., Siwale, Z., Hubbard, J., Bardon, A., Fox, M. P., Kakwesa, G., Kalua, T., Nyasa-Haambokoma, M., et al. (2021). Multimonth dispensing of up to 6 months of antiretroviral therapy in Malawi and Zambia (INTERVAL): a cluster-randomised, non-blinded, non-inferiority trial. *The Lancet Global Health*, 9(5):e628–e638.

Hoffman, R. M., Phiri, K., Parent, J., Grotts, J., Elashoff, D., Kawale, P., Yeatman, S., Currier, J. S., and Schooley, A. (2017). Factors associated with retention in Option B+ in Malawi: a case control study. *Journal of the International AIDS Society*, 20(1):21464.

Insight Start Study Group (2015). Initiation of Antiretroviral Therapy in Early Asymptomatic HIV Infection. *New England Journal of Medicine*, 373(9):795–807.

Karing, A. (2018). Social Signaling and Childhood Immunization: A Field Experiment in Sierra Leone. *Working Paper, UC Berkeley.*

Kennedy, P. E. et al. (1981). Estimation with correctly interpreted dummy variables in semilogarithmic equations [the interpretation of dummy variables in semilogarithmic equations]. *American Economic Review*, 71(4):801–801.

Macis, M., Grunauer, M., Gutierrez, E., Izurieta, R., Phan, P., Reina Ortiz, M., Rosas, C., and Teran, E. (2021). Using Incentives and Nudging to Improve Non-Targeted HIV Testing in Ecuador: A Randomized Trial. *AIDS and Behavior*, 25(8):2542–2550.

Mbuagbaw, L., Van Der Kop, M. L., Lester, R. T., Thirumurthy, H., Pop-Eleches, C., Ye, C., Smieja, M., Dolovich, L., Mills, E. J., and Thabane, L. (2013). Mobile phone text messages for improving adherence to antiretroviral therapy (ART): An individual patient data meta-analysis of randomised trials. *BMJ Open*, 3(12):1–5.

McCullough, J. S., Casey, M., Moscovice, I., and Prasad, S. (2010). The effect of health information technology on quality in U.S. hospitals. *Health Affairs*, 29(4):647–654.

McCullough, J. S., Parente, S. T., and Town, R. (2016). Health information technology and patient outcomes: The role of information and labor coordination. *RAND Journal of Economics*, 47(1):207–236.

Miller, A. R. and Tucker, C. E. (2009). Privacy protection and technology diffusion: The case of electronic medical records. *Management Science*, 55(7):1077–1093.

Miller, A. R. and Tucker, C. E. (2011). Can health care information technology save babies? *Journal of Political Economy*, 119(2):289–324.

Ministry of Health, M. (2011). Malawi integrated clinical HIV guidelines 1st edition.

Nattrass, N. (2019). South african aids activism, art and academia: A memoir from the 2000s. *Journal of Southern African Studies*, 45(1):215–227.

Newell, M.-L., Coovadia, H., Cortina-Borja, M., Rollins, N., Gaillard, P., Dabis, F., et al. (2004). Mortality of infected and uninfected infants born to HIV-infected mothers in Africa: a pooled analysis. *The Lancet*, 364(9441):1236–1243.

Opuni, M., Sanchez-Morales, J. E., Figueroa, J. L., Salas-Ortiz, A., Banda, L. M., Olawo, A., Munthali, S., Korir, J., DiCarlo, M., and Bautista-Arredondo, S. (2023). Estimating the cost of hiv services for key populations provided by the LINKAGES program in Kenya and Malawi. *BMC Health Services Research*, 23(1):337.

Pop-Eleches, C., Thirumurthy, H., Habyarimana, J. P., Zivin, J. G., Goldstein, M. P., de Walque, D., MacKeen, L., Haberer, J., Kimaiyo, S., Sidle, J., Ngare, D., and Bangsberg, D. R. (2011). Mobile phone technologies improve adherence to antiretroviral treatment in a resource-limited setting: a randomized controlled trial of text message reminders. *AIDS*, 27; 25(6):825–834.

Raffetti, E., Postorino, M. C., Castelli, F., Casari, S., Castelnuovo, F., Maggiolo, F., Di Filippo, E., D'Avino, A., Gori, A., Ladisa, N., et al. (2016). The risk of late or advanced presentation of HIV infected patients is still high, associated factors evolve but impact on overall mortality is vanishing over calendar years: results from the Italian MASTER Cohort. *BMC Public Health*, 16(1):1–10.

Shadmi, E., Zeltzer, D., Shir, T., Flaks-Manov, N., Einav, L., and Balicer, R. D. (2020). Can targeting high-risk patients reduce readmission rates? Evidence from Israel. *Journal of Applied Economics*, 23(1):729–745.

Siegfried, N., van der Merwe, L., Brocklehurst, P., and Sint, T. T. (2011). Antiretrovirals for reducing the risk of mother-to-child transmission of HIV infection. *Cochrane database of systematic reviews*, (7).

Tanser, F., Bärnighausen, T., Grapsa, E., Zaidi, J., and Newell, M.-L. (2013). High coverage of ART associated with decline in risk of HIV acquisition in rural KwaZulu-Natal, South Africa. *Science*, 339(6122):966–971.

Teeraananchai, S., Kerr, S., Amin, J., Ruxrungtham, K., and Law, M. (2017). Life expectancy of HIV-positive people after starting combination antiretroviral therapy: a meta-analysis. *HIV medicine*, 18(4):256–266.

Thornton, R. L. (2008). The demand for, and impact of, learning HIV status. *American Economic Review*, 98(5):1829–1863.

UNAIDS (2022). UNAIDS data 2022. *Programme on HIV/AIDS*.

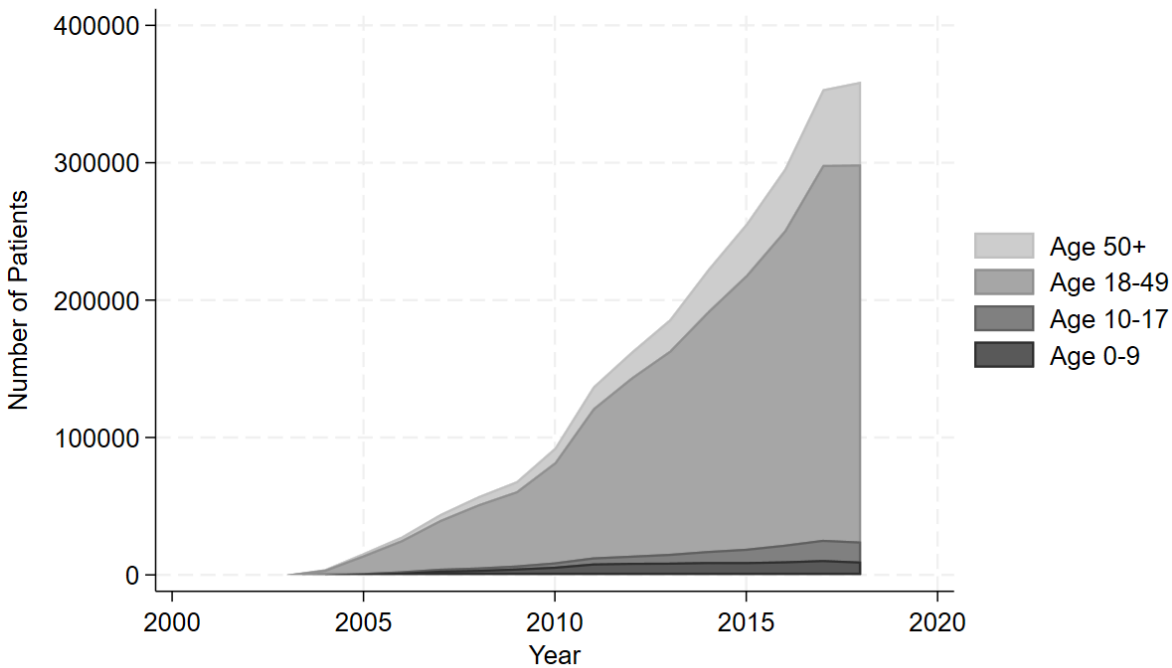
Vermeire, E., Hearnshaw, H., Van Royen, P., and Denekens, J. (2001). Patient adherence to treatment: Three decades of research. A comprehensive review. *Journal of Clinical Pharmacy and Therapeutics*, 26(5):331–342.

Wang, Z., Zhu, Y., Cui, L., and Qu, B. (2019). Electronic health interventions to improve adherence to antiretroviral therapy in people living with HIV: Systematic review and meta-analysis. *JMIR mHealth and uHealth*, 7(10).

Williams, B. G. and Dye, C. (2003). Antiretroviral drugs for tuberculosis control in the era of HIV/AIDS. *Science*, 301(5639):1535–1537.

Yanik, E. L., Napravnik, S., Cole, S. R., Achenbach, C. J., Gopal, S., Olshan, A., Dittmer, D. P., Kitahata, M. M., Mugavero, M. J., Saag, M., et al. (2013). Incidence and timing of cancer in HIV-infected individuals following initiation of combination antiretroviral therapy. *Clinical infectious diseases*, 57(5):756–764.

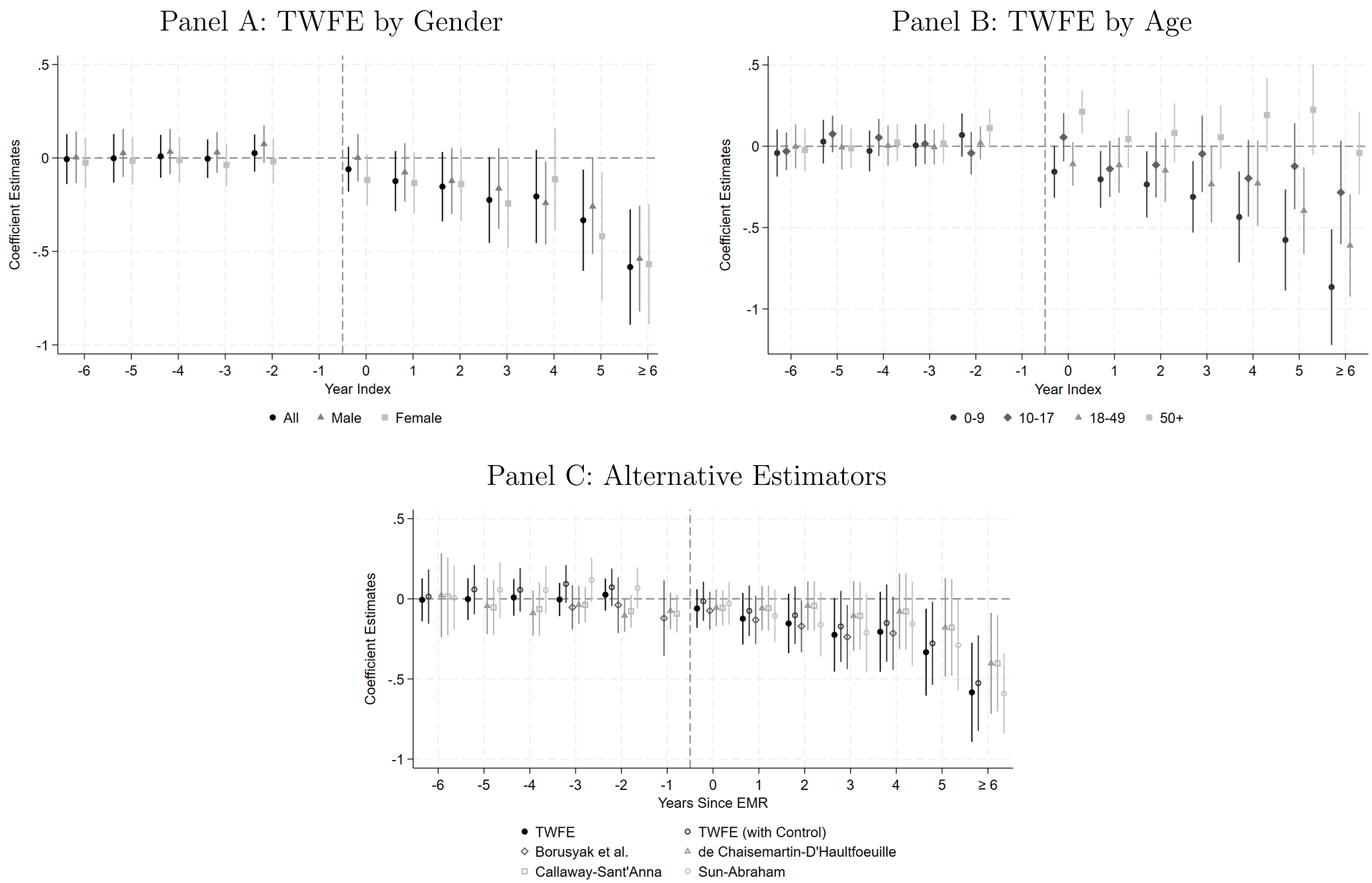
Figure 1
Number of Patients In Care by Age



Notes: Total number of patients in active care (at least one ART initiation or refill visit) in all sample clinics by year and age group.

Figure 2

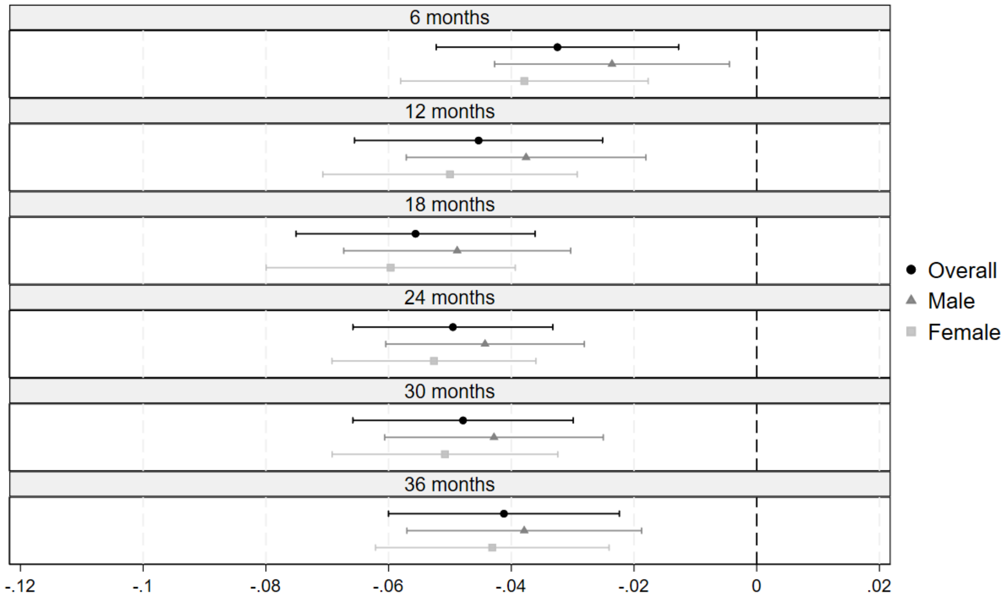
Dynamic Effect of EMR Adoption on log(Patient Deaths)



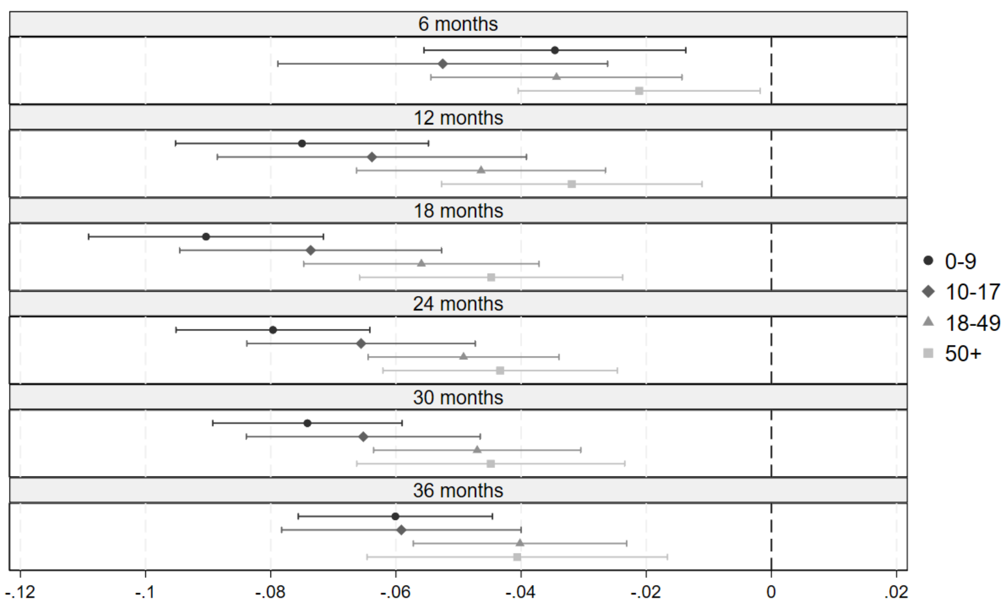
Notes: *Patient Deaths* is the number of ART patients that died, measured at the clinic-year level. We add 1 to the outcome measure. $k = 0$ is the year immediately following EMR implementation. Each regression includes year and clinic fixed effects following (2). 95% C.I. based on heteroskedasticity-robust standard errors clustered at the clinic level.

Figure 3
Effect of EMR System on Patient Lapse Probability

Panel A: Patient Lapse by Gender



Panel B: Patient Lapse by Age

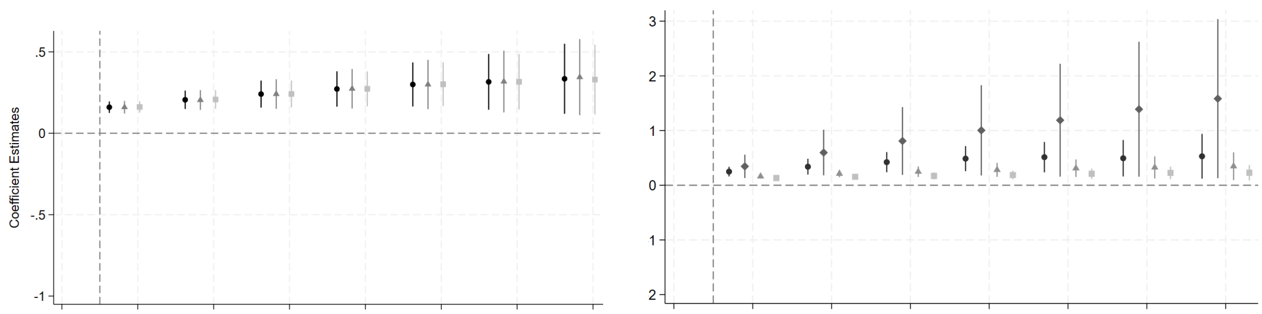


Notes: Coefficient estimate (Post-EMR) on x-axis. *Patient Lapse* is measured at the individual patient-year (t) level, on the sample of patients who have initiated care by year t. It is an indicator for no clinic visit in the m months prior to the end of year t. Regressions include year and clinic fixed effects as well as controls, as in (3). Controls include gender, age category, years since initiation, an indicator for whether the patient was underweight at the time of initiation, and an indicator for whether the patient had advanced HIV at the time of initiation. 95% C.I. based on heteroskedasticity-robust standard errors clustered at the clinic level.

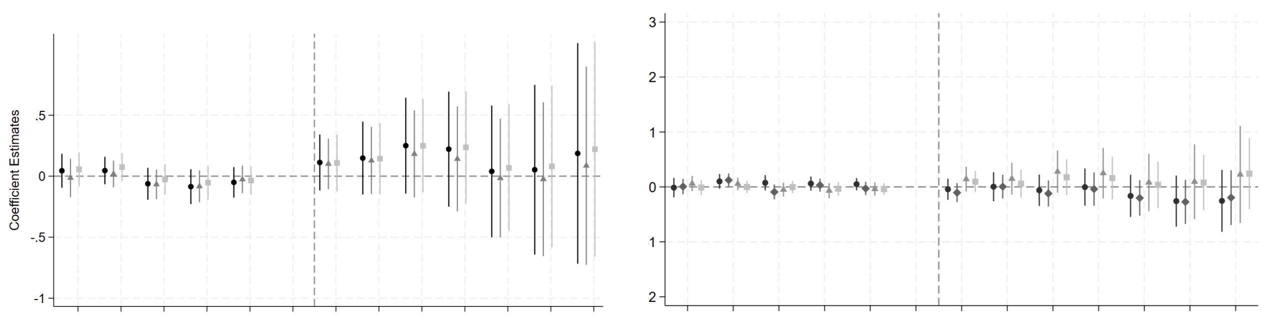
Figure 4

Dynamic Effect of EMR Adoption on log(Number of Patients)

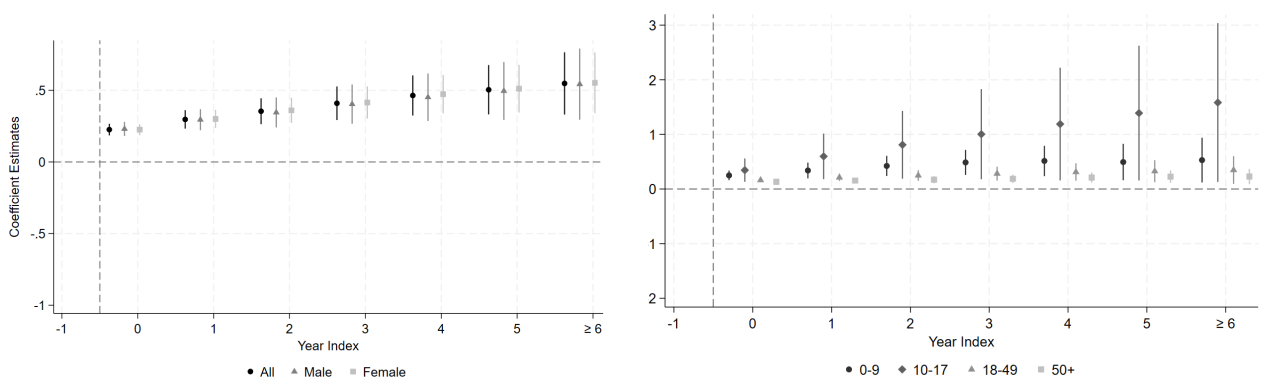
Panel A: log(Patients In Care)



Panel B: log(New Patients)



Panel C: log(Retention)

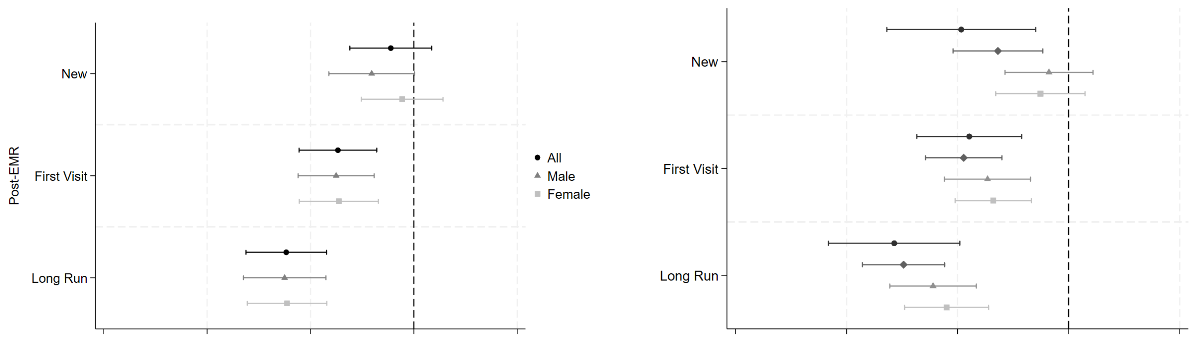


Notes: Patients In Care is the number of ART patients that visited the clinic at least once, measured at the clinic-year level. *New Patients* is the number of new ART patients that initiated ART for the first time. *Patient Retention* is the number of returning (did not initiate for the first time that year) ART patients that visited the clinic. We add 1 to the outcome measure. $k = 0$ is the year immediately following EMR implementation. Each regression includes year and clinic fixed effects following (2). 95% C.I. based on heteroskedasticity-robust standard errors clustered at the clinic level.

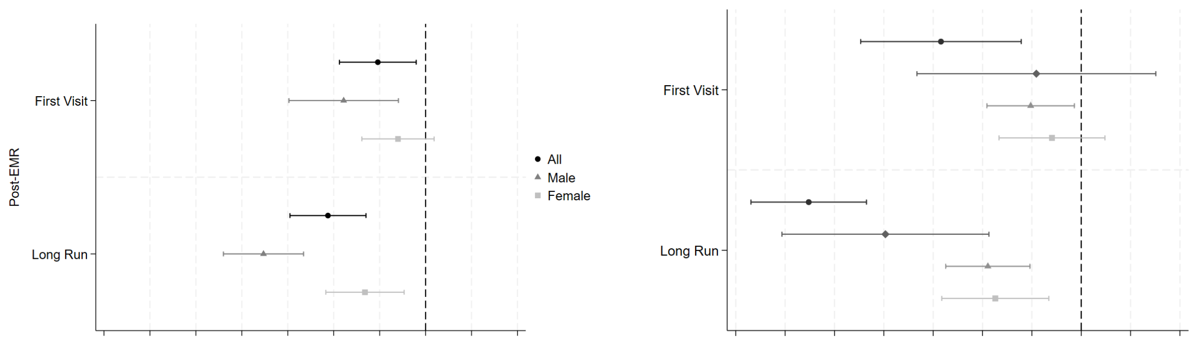
Figure 5

Effect of EMR Adoption on Patient Adherence and Health

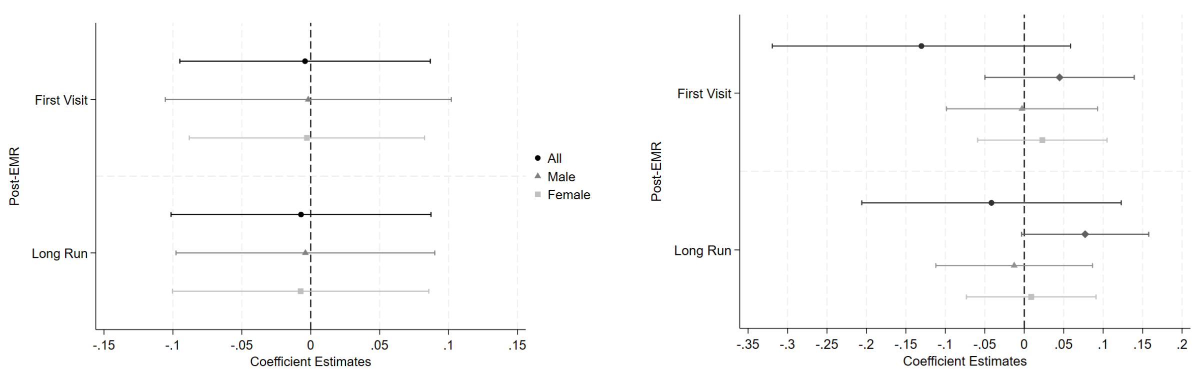
Panel A: Missed Next Appointment



Panel B: Underweight



Panel C: Viral Suppression

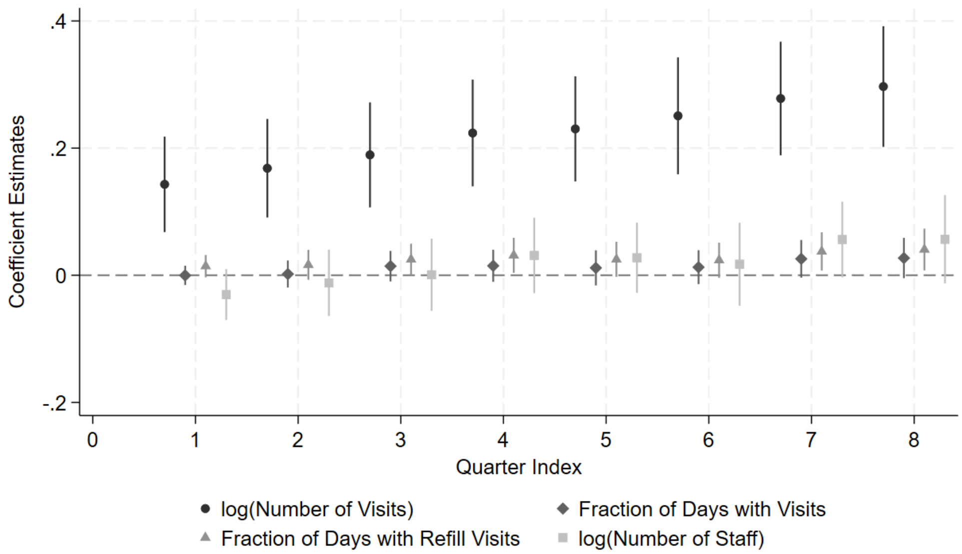


Notes: Coefficient estimate (Post-EMR) on x-axis. *Missed next appointment*: 2 months late for next refill. *Underweight*: BMI 2 SD below median. *Viral Suppression*: < 200 copies of HIV per ml of blood. Outcomes at patient-visit level. Regressions include year and clinic fixed effects and controls, as in (3). Controls include gender, age category, years since initiation, an indicator for whether the patient was underweight at the time of initiation, and an indicator for whether the patient had advanced HIV at the time of initiation. Results reported separately for first refill visit post-EMR, later visits. 95% C.I. based on heteroskedasticity-robust standard errors clustered at clinic level.

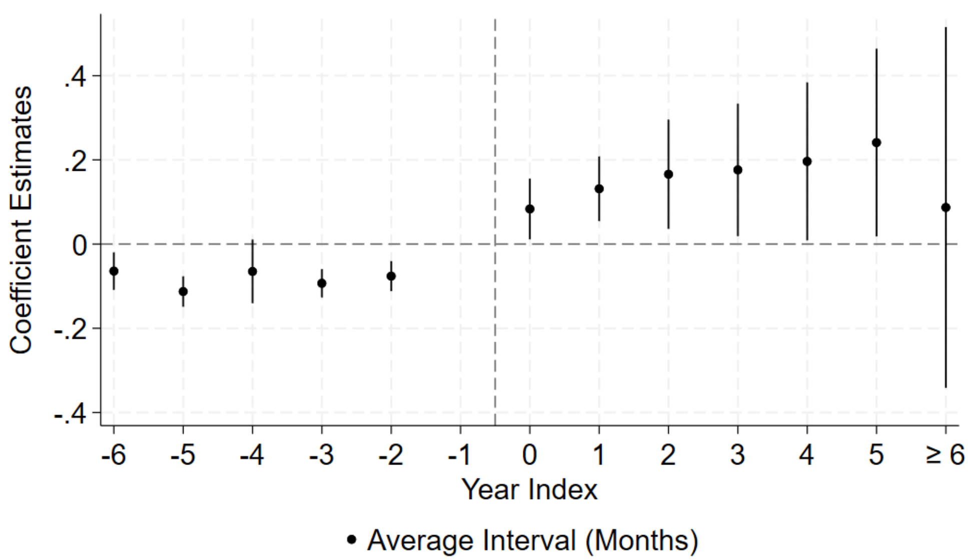
Figure 6

Clinic Efficiency Measures

Panel A: Short-Term Efficiency Gains



Panel B: Effect of EMRs on Visit Spacing



Notes: A: clinic-quarter level number of patient visits (+1), fraction of days with ART visits, and number of staff (+1). Variant of (2) with only clinic fixed effect. $k = 0$ is quarter immediately after EMR implementation and reference period. Sample: $0 \leq k \leq 8$. B: mean months between initiation and first scheduled visit, (2) at clinic-year level. $k = -1$ is reference period. All: 95% C.I. based on heteroskedasticity-robust standard errors clustered at the clinic level.

Table 1
Effect of EMR Adoption on Clinic-Level Patient Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	All	Gender		Age			
		Male	Female	≤ 9	10 to 17	18 to 49	≥ 50
Panel A: log(Patient Deaths)							
Post-EMR	-0.111*	-0.100	-0.101	-0.214***	-0.0428	-0.126*	0.0800
	(0.0655)	(0.0616)	(0.0685)	(0.0699)	(0.0532)	(0.0689)	(0.0587)
BJS							
Post-EMR	-0.218***	-0.210***	-0.197**	-0.345***	-0.105**	-0.239***	0.0456
	(0.0803)	(0.0740)	(0.0809)	(0.0800)	(0.0532)	(0.0818)	(0.0706)
Observations	1,166	1,166	1,166	1,166	1,166	1,166	1,166
R-squared	0.790	0.779	0.766	0.555	0.530	0.779	0.719
Panel B: log(Patients In Care)							
Post-EMR	0.151***	0.148***	0.154***	0.252***	0.226***	0.153***	0.124***
	(0.0168)	(0.0180)	(0.0168)	(0.0432)	(0.0464)	(0.0183)	(0.0151)
Observations	411	411	411	411	411	411	411
R-squared	0.985	0.984	0.985	0.949	0.960	0.983	0.988
Panel C: log(New Patients)							
Post-EMR	0.205	0.167	0.178	-0.0899	-0.0614	0.213	0.124
	(0.169)	(0.151)	(0.164)	(0.130)	(0.106)	(0.163)	(0.130)
BJS							
Post-EMR	0.278*	0.209	0.258*	-0.115	-0.0614	0.292**	0.202*
	(0.152)	(0.136)	(0.151)	(0.134)	(0.103)	(0.148)	(0.121)
Observations	1,166	1,166	1,166	1,166	1,166	1,166	1,166
R-squared	0.785	0.781	0.785	0.735	0.719	0.789	0.740
Panel D: log(Retention)							
Post-EMR	0.199***	0.201***	0.200***	0.340***	0.294***	0.200***	0.161***
	(0.0216)	(0.0237)	(0.0215)	(0.0467)	(0.0543)	(0.0227)	(0.0215)
Observations	411	411	411	411	411	411	411
R-squared	0.982	0.981	0.982	0.944	0.954	0.981	0.988

Notes: Patient Deaths is the number of ART patients that died, measured at the clinic-year level. *Patients In Care* is the number of ART patients that visited the clinic at least once, measured at the clinic-year level. *New Patients* is the number of new ART patients that initiated ART for the first time. *Patient Retention* is the number of returning (did not initiate for the first time that year) ART patients that visited the clinic, measured at the clinic-year level. We add 1 to each outcome measure. Each regression includes year and clinic fixed effects following (1). ***BJS*** indicates alternative DiD estimation based on Borusyak et al. (2021). Heteroskedasticity-robust standard errors clustered at the clinic level, in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 2

Medical Care Provided at Visit

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Days of Medication Provided at Initiation		Child Underweight at 3-6 Months		TB Treatment at Initiation		All Referrals at Initiation		Ordered CD4 Count at Initiation	
Post-EMR	4.574***	5.568***	0.0200	0.0430	-0.00283	-0.00970	-0.0476	-0.0254	0.00219	0.00172
	(0.773)	(1.529)	(0.0321)	(0.0385)	(0.00691)	(0.00909)	(0.0322)	(0.0339)	(0.0244)	(0.0234)
Hospital Post-EMR	-2.265		-0.0533		0.0149		-0.0572		0.00127	
	(3.742)		(0.0569)		(0.00930)		(0.0416)		(0.0315)	
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Observations	733,891	719,780	3,267	3,256	755,066	740,917	755,066	740,917	755,066	740,917
R-squared	0.124	0.125	0.098	0.098	0.025	0.026	0.160	0.161	0.129	0.128

Notes: (1-2): the number of days' worth of medication provided at initiation. (3-4): indicator for an underweight child being underweight at 3-6 month follow-up visit. (5-6): indicator for TB treatment at initiation. (7-8): number of referrals to other departments at initiation. (9-10): indicator for CD4 count order at initiation. Outcomes measured at individual-visit level. Regressions include year and clinic fixed effects as well as controls, as in (3). Controls include gender, age category, years since initiation, an indicator for whether the patient was underweight at the time of initiation, and an indicator for whether the patient had advanced HIV at the time of initiation. Results without controls in Appendix Table A13. Heteroskedasticity-robust standard errors clustered at the clinic level, in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 3
EMR Heterogeneous Treatment Effect by Clinic Type

	(1) log(Patient Deaths)	(2) log(Patients In Care)	(3) log(New Patients)	(4) log(Retention)
Panel A: Hospital or Regular Clinic				
Post Hospital	-0.220**	0.0810***	0.421	0.125***
	(0.0873)	(0.0291)	(0.326)	(0.0354)
Post Regular	0.00523	0.207***	0.0916	0.250***
	(0.0732)	(0.0243)	(0.186)	(0.0277)
BJS				
Post Hospital	-0.265***	N/A	0.231	N/A
	(0.0980)		(0.197)	
Post Regular	-0.142	N/A	0.233*	N/A
	(0.0924)		(0.123)	
Observations	1,147	407	1,147	407
R-squared	0.792	0.986	0.787	0.983
Panel B: Clinic Size				
Post Large	-0.108	0.182***	0.167	0.247***
	(0.0853)	(0.0196)	(0.237)	(0.0284)
Post Small	-0.113	0.128***	0.244	0.163***
	(0.0737)	(0.0258)	(0.268)	(0.0301)
BJS				
Post Large	-0.165	N/A	-0.0171	N/A
	(0.182)		(0.200)	
Post Small	-0.244***	N/A	0.270	N/A
	(0.0857)		(0.233)	
Observations	1,166	411	1,166	411
R-squared	0.790	0.985	0.786	0.982
Panel C: Clinic Location				
Post Urban	-0.110	0.203***	0.436	0.250***
	(0.0875)	(0.0217)	(0.280)	(0.0313)
Post Rural	-0.101	0.110***	-0.0653	0.151***
	(0.0738)	(0.0234)	(0.226)	(0.0285)
BJS				
Post Urban	-0.216**	N/A	0.0627	N/A
	(0.103)		(0.244)	
Post Rural	-0.200**	N/A	0.108	N/A
	(0.0903)		(0.152)	
Observations	1,147	407	1,147	407
R-squared	0.791	0.985	0.788	0.982

Notes: Outcomes measured at the clinic-year level (+1). Each regression includes year and clinic fixed effects following (1). *BJS* indicates alternative DiD estimation based on Borusyak et al. (2021). Large clinics have above-median number of initiating patients as of 2013. Heteroskedasticity-robust standard errors clustered at the clinic level, in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

The Lifesaving Impact of Electronic Medical Records for HIV Patients

Laura Derksen, Anita M. McGahan and Leandro S. Pongeluppe

A Appendix – For Online Publication

A.1 Cost-Effectiveness Calculations

The estimates of number of deaths averted presented in Section 4.6 are based on the dynamic treatment effect estimates (Table A5), which we use to calculate the counterfactual number of deaths that would have occurred in each clinic and year in the absence of the EMR system. We then calculate the approximate number of deaths averted in each clinic and year after EMR implementation, and aggregate over the clinic-years in our sample. The estimate of total deaths averted is likely an underestimate, as an additional 18 clinics, from which we did not obtain complete data, had adopted EMRs as of 2019. We calculate the number of disability-adjusted life years (DALYs) averted in the first five years using the same approach, and aggregating additional years of life across the deaths averted in each year. We scale this gain in life years by 0.053, the disability weight associated with treated HIV (Salomon, 2013). This produces an estimate of 184 DALYs averted, per clinic, in the first five years after EMR adoption.

The total cost of EMR implementation for an average clinic with three workstations is USD \$34,050.[33] Cost-per-life-saved (USD \$448) is based on the clinic-level cost of EMR (USD \$34,050), and the clinic-level average number of patient deaths prevented in the first five years (76). The calculation relies on data from clinics that have had EMRs for at least five years, while the total number of deaths averted (5,050) includes all EMR clinics.

We can similarly construct counterfactuals for the number of patients active in care, and approximate the impact on new HIV infections. After five years, EMR clinics provided care to 1,841 more patients on average than clinics without EMR. This large increase in the number of patients on ART inevitably led to a reduction in new HIV infections, which are not captured in the cost-effectiveness calculation above. An approximation based on (Tanser

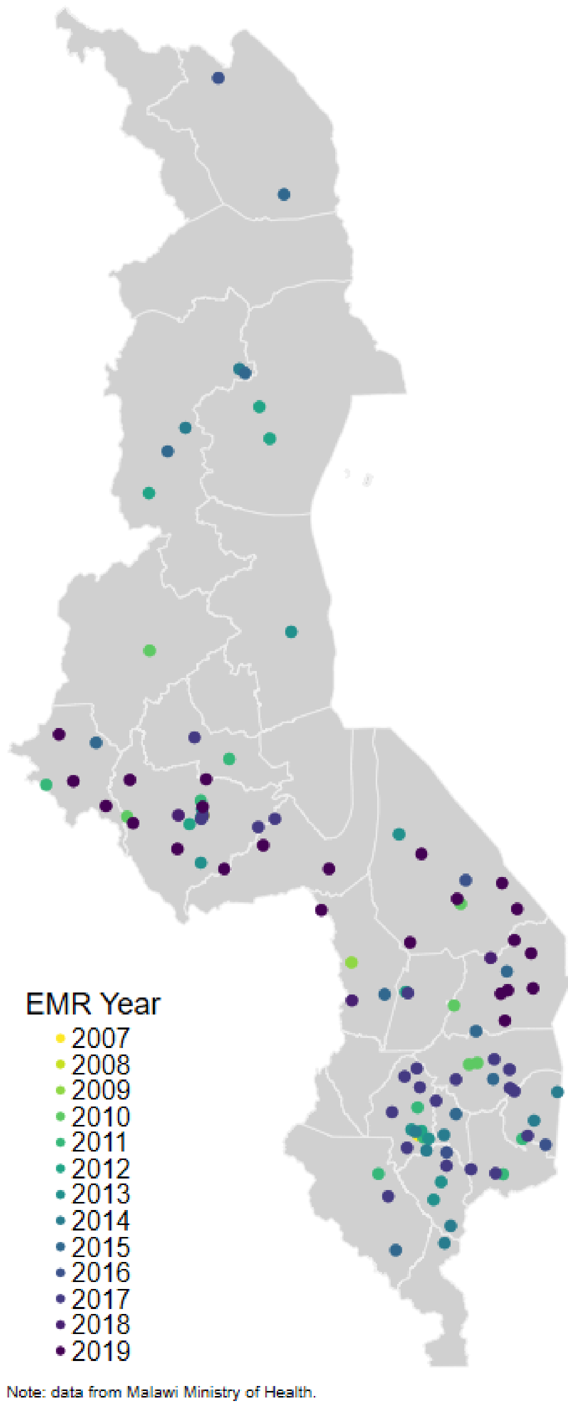
[33] Table A2 breaks down costs of EMR implementation. These values were provided by Baobab Health Trust and from Driessen et al. (2013).

et al., 2013) suggests a per-clinic reduction of approximately 30 new infections per year due to EMR implementation. Five years after EMR implementation, the fraction of HIV patients in active care increased by 34 percent. In 2019, approximately 69 percent of Malawians living with HIV were in active treatment (UNAIDS, 2022). We assume that half of patients had access to an EMR clinic at that point (124 out of 297 clinics), so treatment rates were 79 percent for patients with access to EMRs and 59 percent elsewhere. This suggests a 20 percentage point increase in treatment rates in EMR areas. Tanser et al. (2013) find that a 1 percent increase in ART coverage is associated with a 1.4 percent reduction in new HIV transmissions. Scaled by 20 percentage points, EMR implementation leads to an approximate 28 percent reduction in incidence. In 2019, incidence was 1.33 per 1000 Malawian residents (UNAIDS, 2022). This suggests incidence in non-EMR areas incidence would be 1.55 per 1000, versus 1.11 in EMR areas. The population of Malawi was 18.63 million in 2019, and with HIV prevalence at 10.6 (Malawi DHS 2015-16), this implies an uninfected population of 16.66 million. We therefore calculate approximately 3,665 annual new infections averted, or 30 per EMR clinic.

To approximate the longer run impact of EMRs, we calibrate and simulate a Cox Proportional Hazard model. To estimate the control group hazard rate, we use average death rates by age group for non-EMR clinics in years 2016 and later. For the treatment group hazard rate, we scale these death rates by the long-run elasticities-by-age-group associated with our dynamic treatment effect estimates (Table A5). We then simulate 1,000,000 patient lifespans, with initial age distribution representative of the average clinic in 2018. These model simulation rely on some assumptions, most important of which is the assumption that the impact on the death rate is constant in the long run. We find that over the longer term, EMRs extend patient lives by 1.4 years, or 1.3 DALYs on average. Scaling by the average clinic population in 2018 (3,455), this implies a cost per DALY of USD $7.37. This estimate is an underestimate as it does not take into account the fact that the clinic population may grow over time as a result of new HIV infections.

Appendix Figure A1

EMR Implementation by Year in Malawi

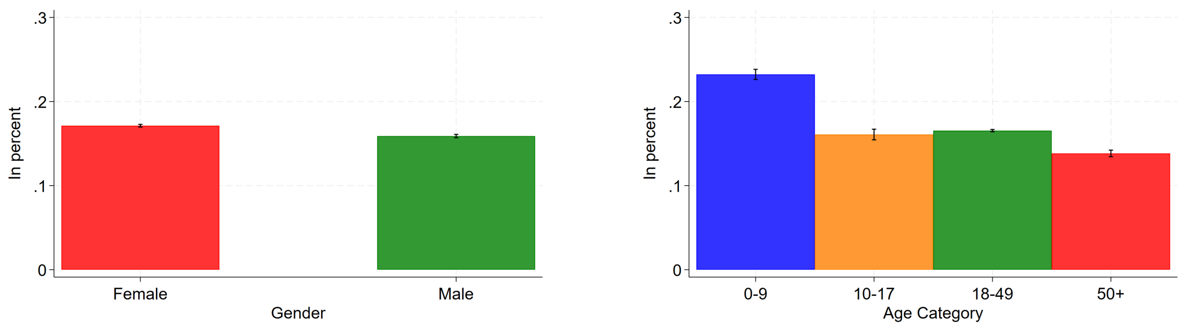


Notes: Geospatial data from 103 out of 106 clinics. The latitude and longitude data of clinics was hand-collected using the name of clinics in google maps. For three clinics in the sample, we could not precisely identify their geospatial location.

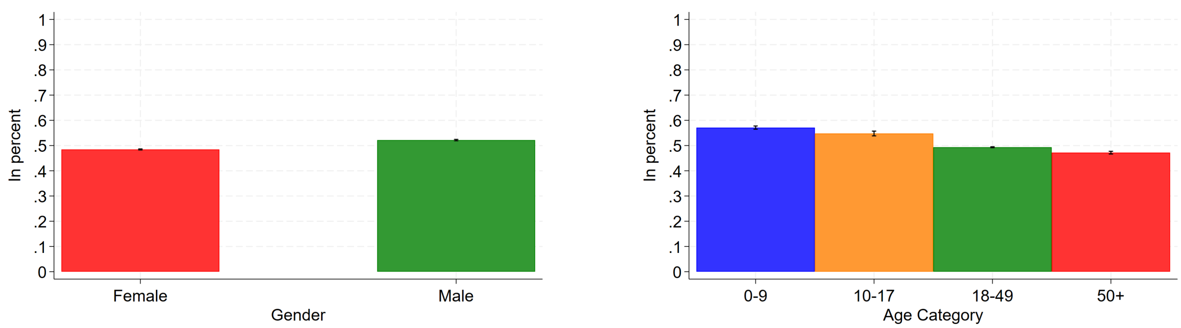
Appendix Figure A2

Descriptive Statistics by Gender and Age

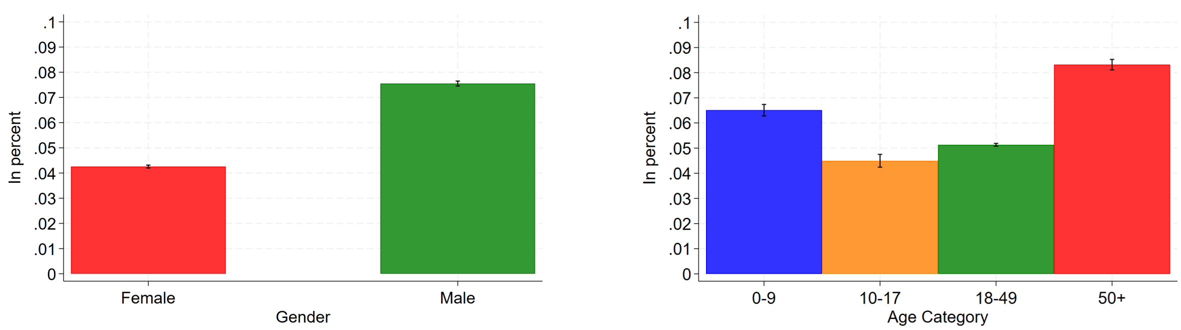
Panel A: Missed First Appointment



Panel B: Lapsed at EMR Adoption



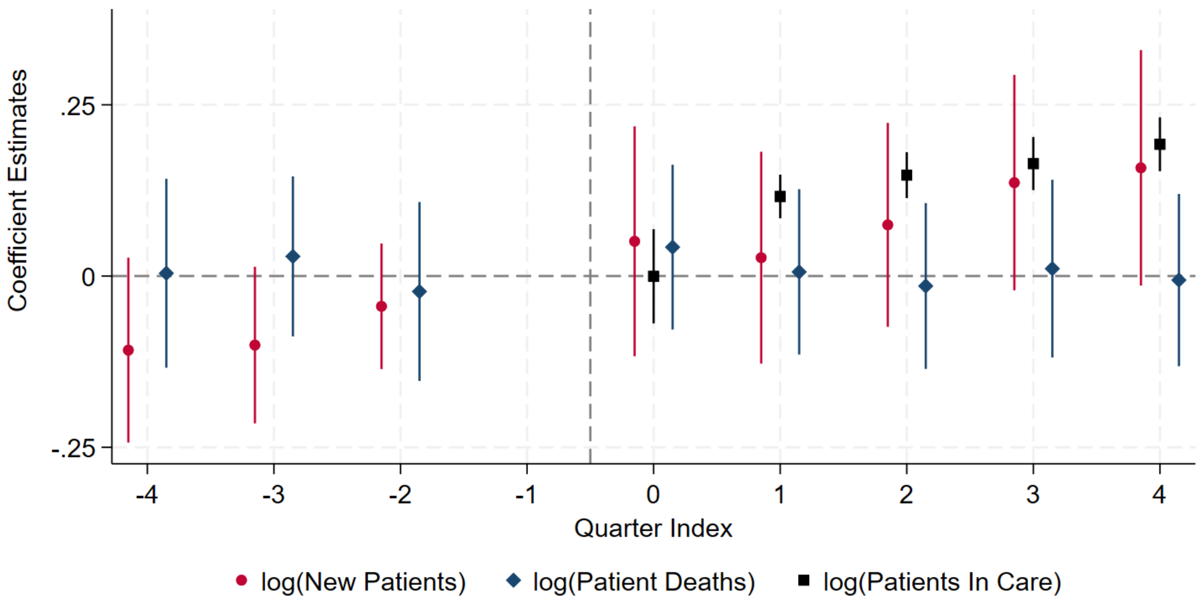
Panel C: Death Within 2 Years



Notes: Panel A: an indicator for patients who missed their first appointment after ART initiation (at least 2 months late). Panel B: an indicator for no clinic visit in the year immediately preceding EMR implementation. Includes only patients who initiate at least 1 year prior to EMR implementation. Panel C: an indicator for death within 2 years of ART initiation.

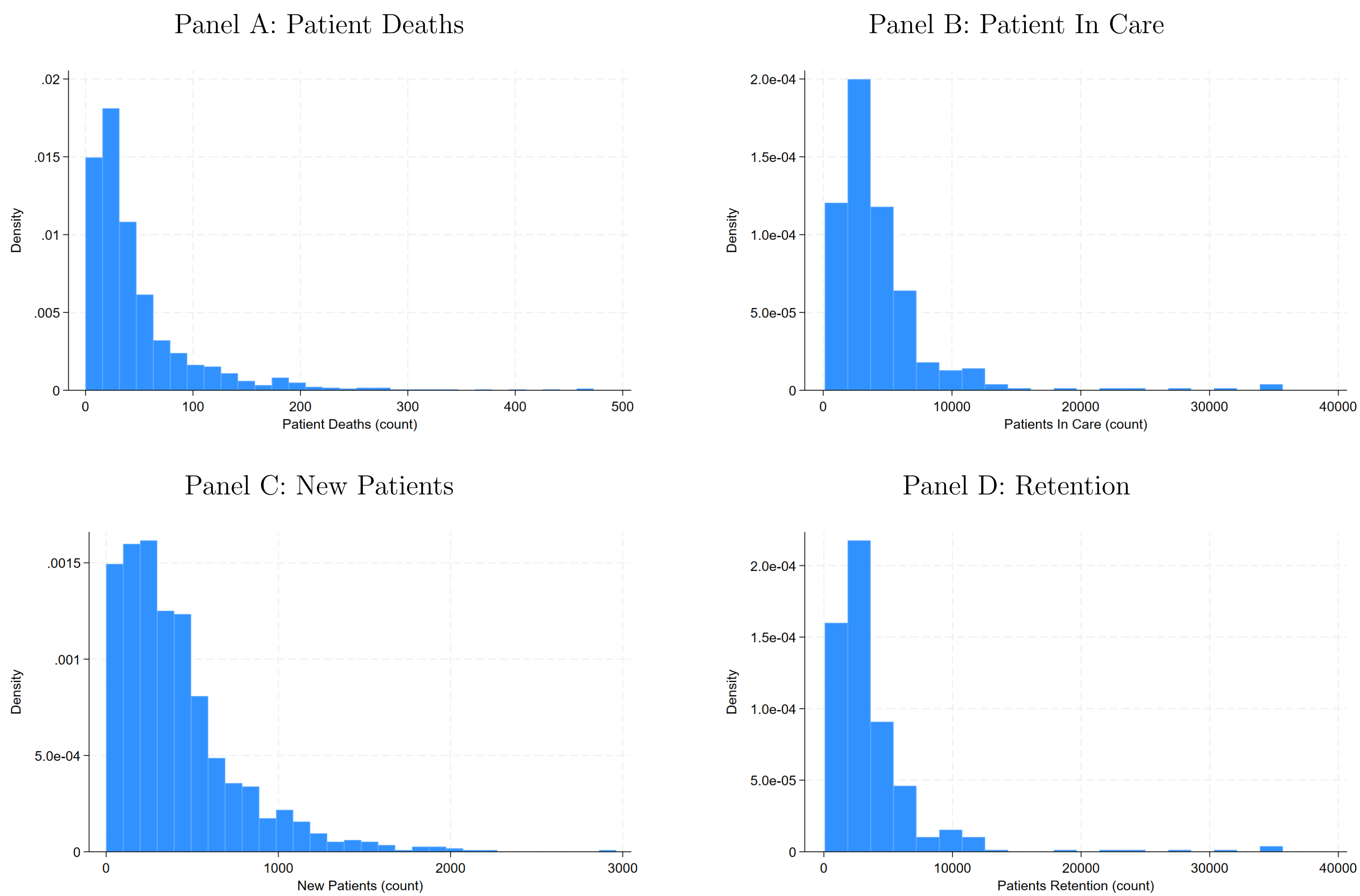
Appendix Figure A3

Clinic-Level Outcomes by Quarter



Notes: Outcomes measured at clinic-quarter level. We add 1 to all outcome measures. $k = 0$ is the quarter immediately following EMR implementation, $k = -1$ serves as the reference period. Regressions include only clinic fixed effects. 95% C.I. based on heteroskedasticity-robust standard errors clustered at the clinic level.

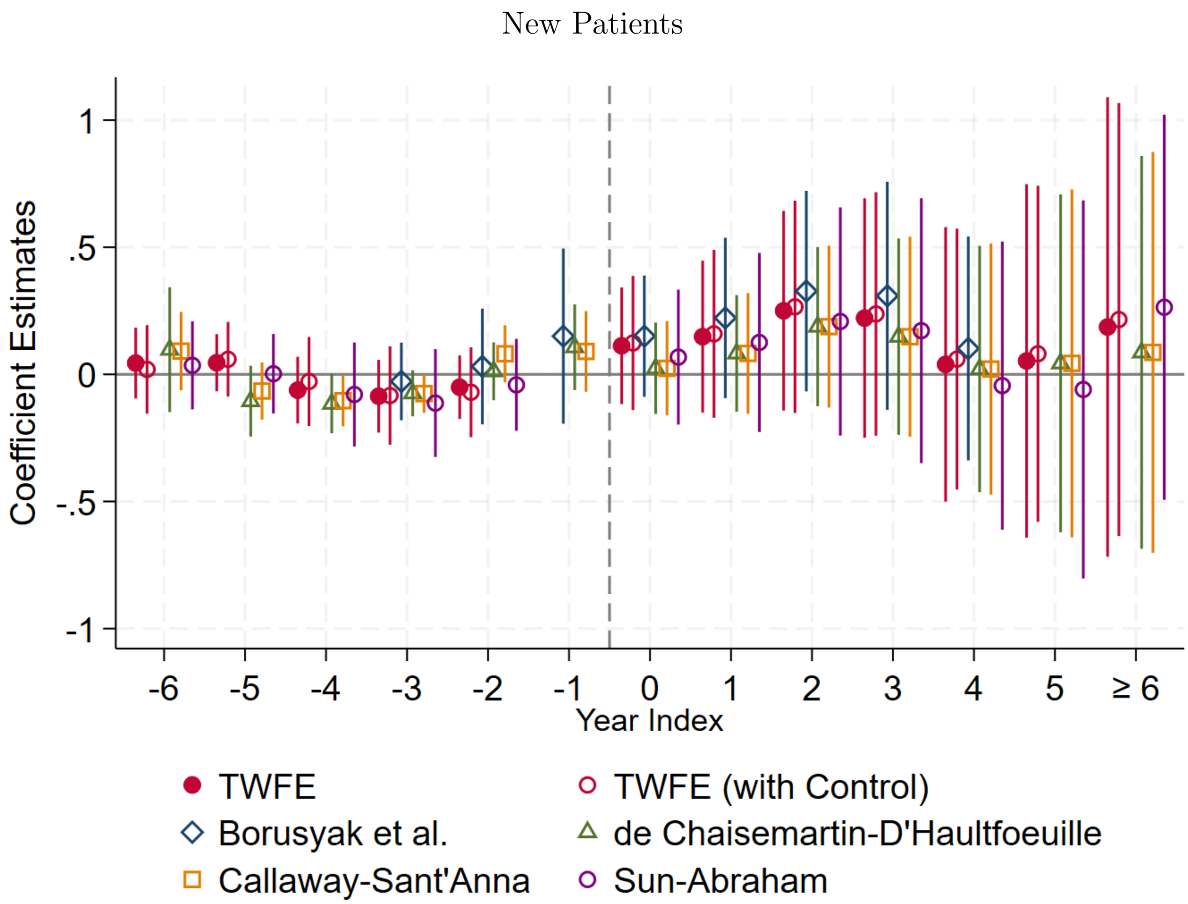
Appendix Figure A4
Distribution of Patient Outcomes



Notes: Histograms illustrating skew in outcomes due to variations in clinic size.

Appendix Figure A5

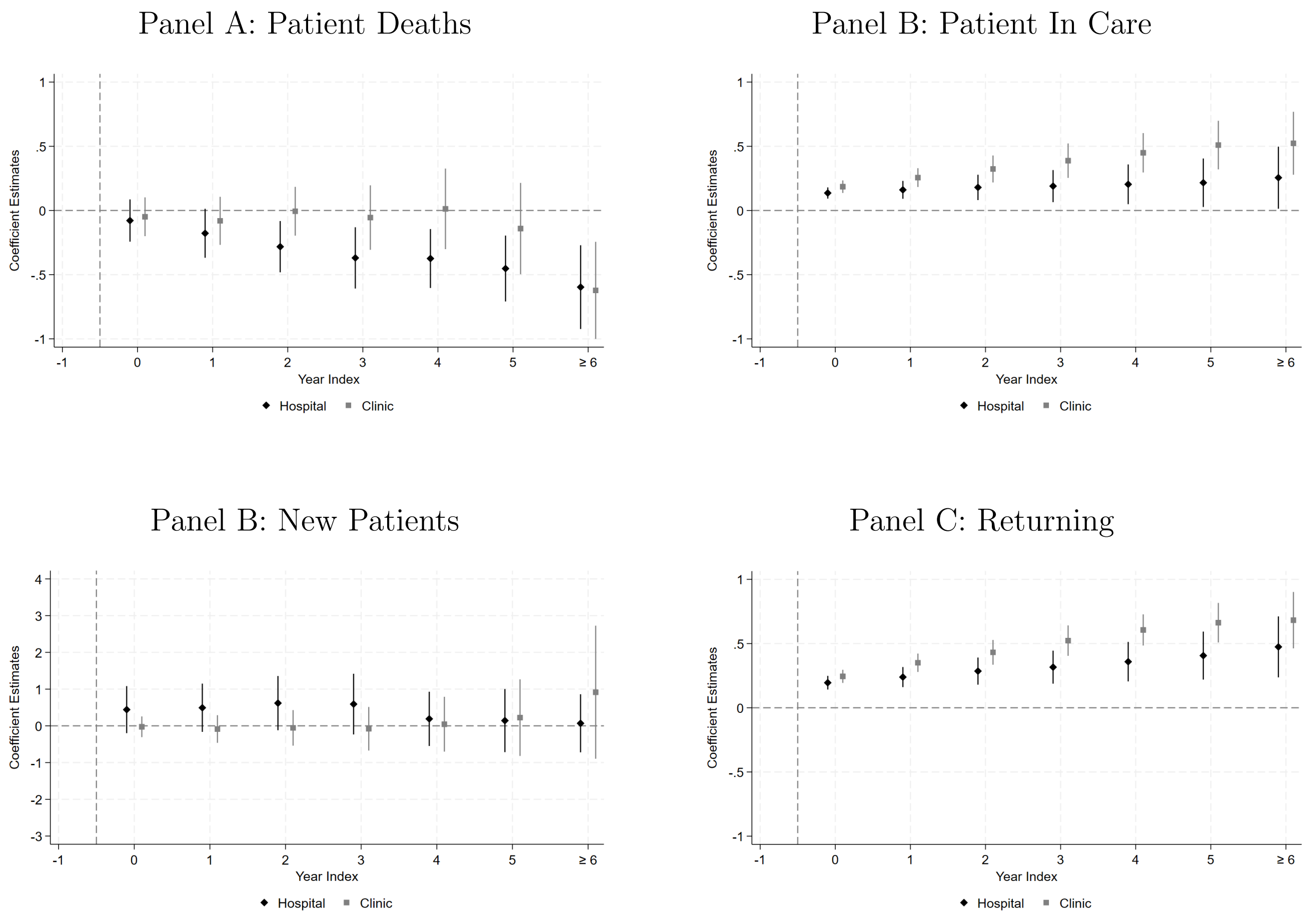
EMR Treatment Effect on New Patients, Alternative Estimators



Notes: Outcomes measured at clinic-year level. We add 1 to the outcome measure. Alternative estimators following Abraham and Sun (2020); Callaway and Sant'Anna (2020); de Chaisemartin and D'Haultfœuille (2020); Borusyak et al. (2021). Standard errors were estimated using 1,000 bootstrap replications in de Chaisemartin and D'Haultfoeuille (2020). $k = 0$ is the year immediately following EMR implementation. 95% C.I. based on heteroskedasticity-robust standard errors clustered at the clinic level.

Appendix Figure A6

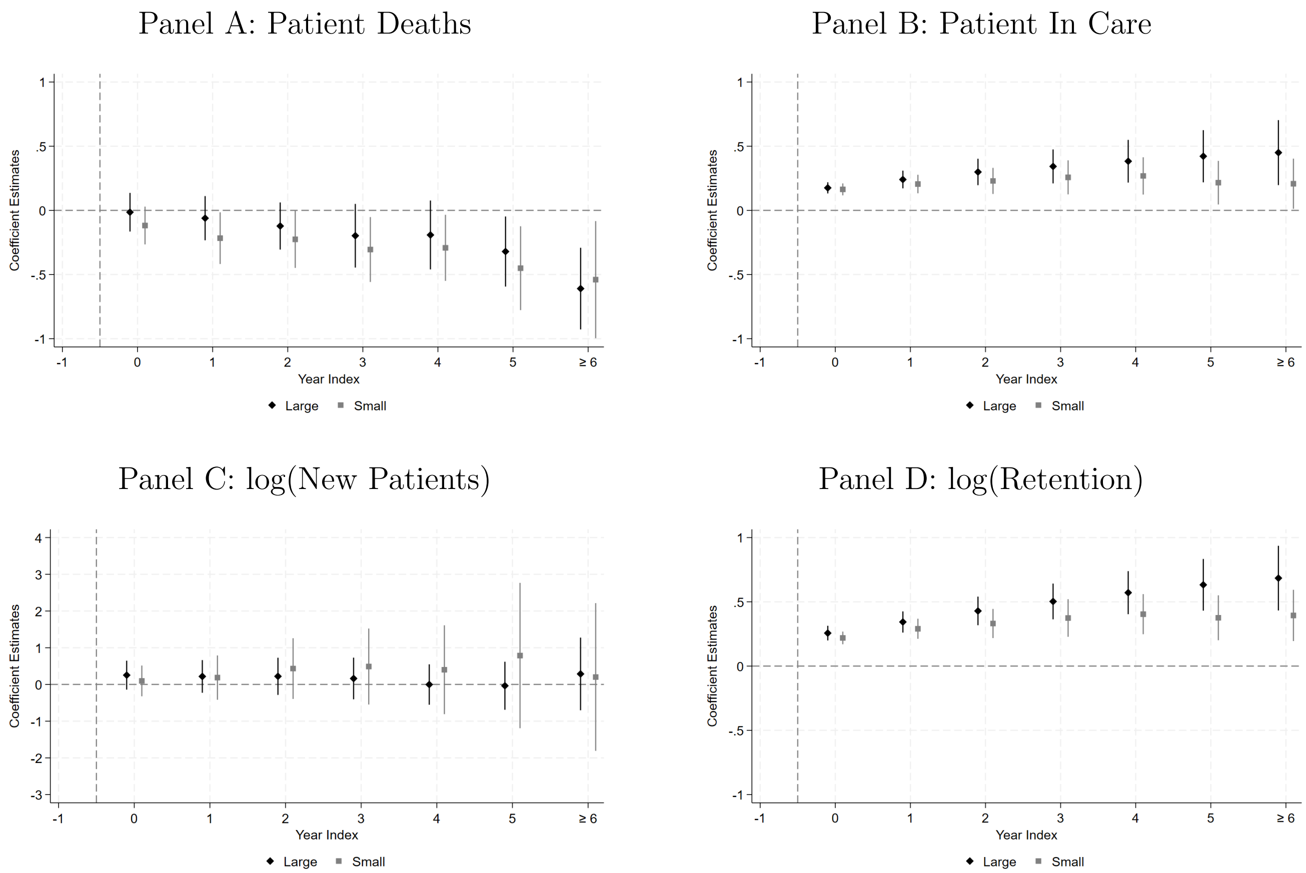
Effect of EMR on Patient Outcomes by Clinic Type (Hospital vs. Clinic)



Notes: Outcomes measured at clinic-year level. Panel A: number of ART patient deaths. Panel B: number of ART patients visiting the clinic at least once for ART initiation and/or prescription refill. Panel C: the number of patients who visit the clinic for ART initiation. Panel D: number of patients who visit the clinic only for prescription refill. Clinic type defined as hospitals or regular clinics. We add 1 to all outcome measures. $k = 0$ is the year immediately following EMR implementation, $k \leq -1$ serve as the reference periods. Regressions include year and clinic fixed effects. 95% C.I. based on heteroskedasticity-robust standard errors clustered at the clinic level.

Appendix Figure A7

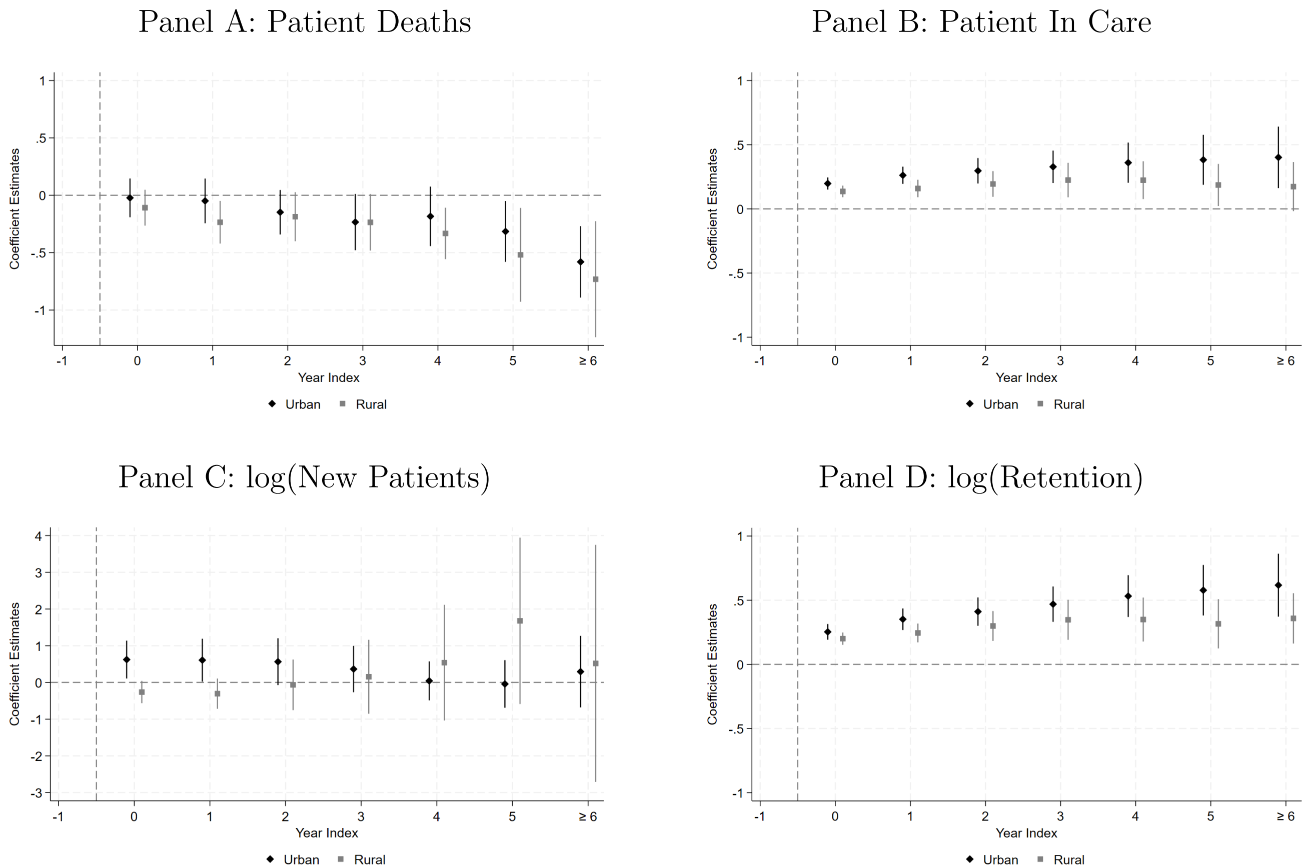
Effect of EMR on Patient Outcomes by Clinic Size (Large vs. Small)



Notes: Outcomes measured at clinic-year level. Panel A: number of ART patient deaths. Panel B: number of ART patients visiting the clinic at least once for ART initiation and/or prescription refill. Panel C: the number of patients who visit the clinic for ART initiation. Panel D: number of patients who visit the clinic only for prescription refill. Clinic size defined as above or below mean number of new patients in 2013. We add 1 to all outcome measures. $k = 0$ is the year immediately following EMR implementation, $k \leq -1$ serve as the reference periods. Regressions include year and clinic fixed effects. Large clinics have above-median number of initiating patients as of 2013. 95% C.I. based on heteroskedasticity-robust standard errors clustered at the clinic level.

Appendix Figure A8

Effect of EMR on Patient Outcomes by Clinic Location (Urban vs. Rural)

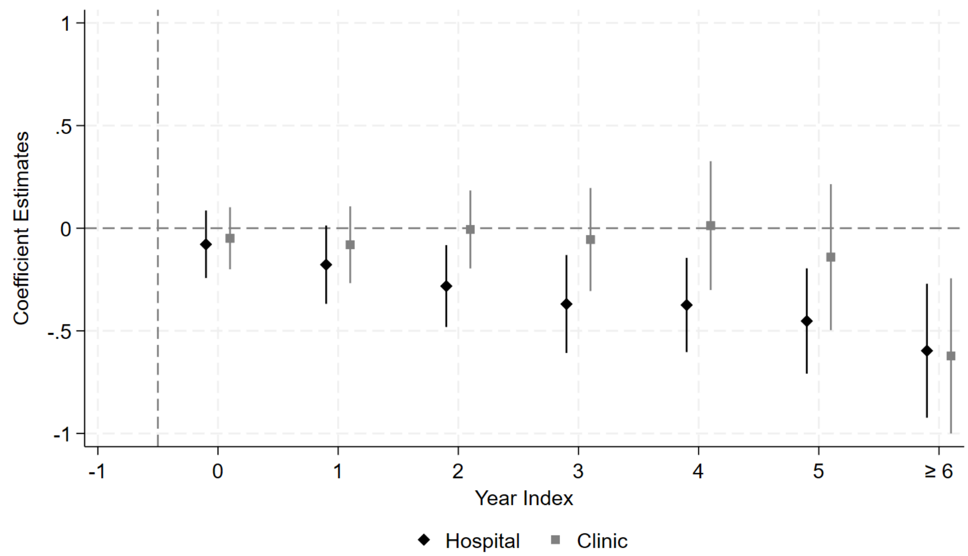


Notes: Outcomes measured at clinic-year level. Panel A: number of ART patients visiting the clinic at least once for ART initiation and/or prescription refill. Panel B: number of ART patient deaths. Panel C: the number of patients who visit the clinic for ART initiation. Panel D: number of patients who visit the clinic only for prescription refill. Clinic location defined as in urban or rural areas. We add 1 to all outcome measures. $k = 0$ is the year immediately following EMR implementation, $k \leq -1$ serve as the reference periods. Regressions include year and clinic fixed effects. 95% C.I. based on heteroskedasticity-robust standard errors clustered at the clinic level.

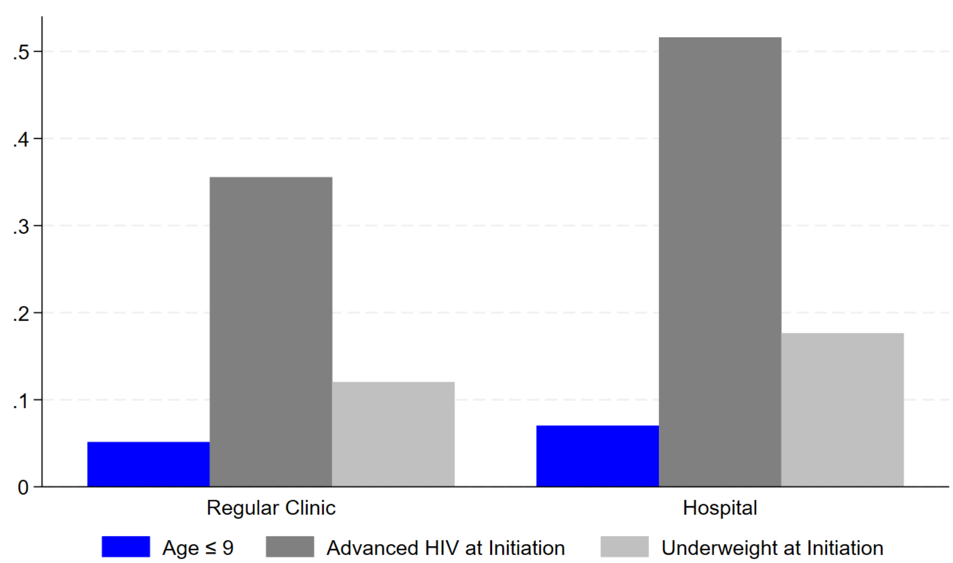
Appendix Figure A9

Dynamic Effect of EMR Adoption in Hospitals vs. Regular Clinics

Panel A: log(Patient Deaths)

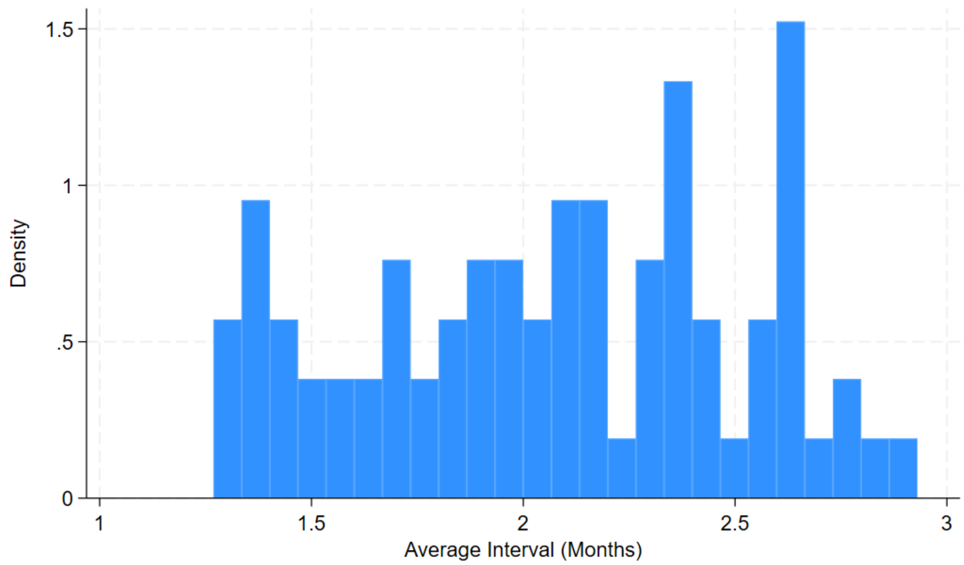


Panel B: Patient Selection



Notes: Patient Deaths is the number of ART patients that died, measured at the clinic-year level. We add 1 to the outcome measure. $k = 0$ is the year immediately following EMR implementation. $k = -1$ serves as the reference period. Regressions include year and clinic fixed effects. 95% C.I. based on heteroskedasticity-robust standard errors clustered at the clinic level.

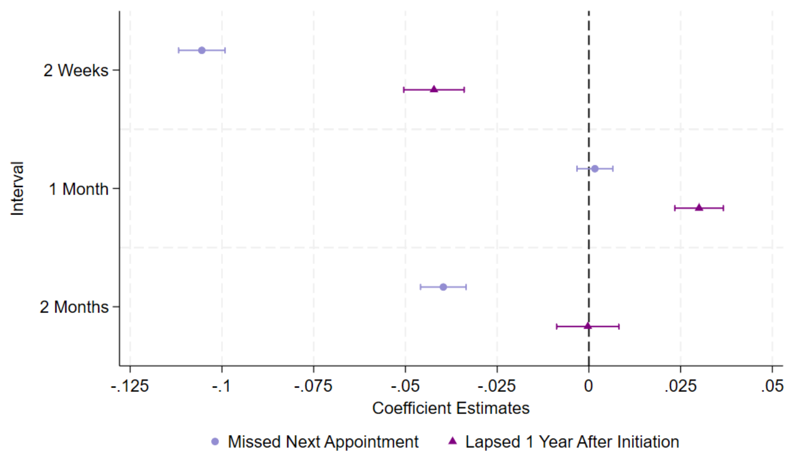
Appendix Figure A10
Clinic-Level Average Dispensing Interval



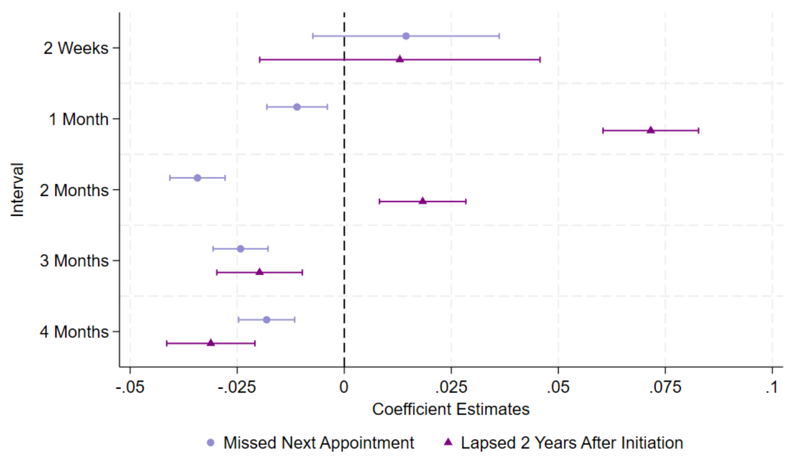
Notes: Clinic-level average interval between scheduled appointments (2018).

Appendix Figure A11
Appointments Intervals (in Months) and Patient Lapse

Panel A: At Initiation



Panel B: One Year After Initiation



Notes: Missing Next Appointment is an indicator for missed next next scheduled appointment. *Lapse* is an indicator for no clinic visit in the six months before the end of the year. Panel B: measures captured at the first visit in the six months before the end of the year. Coefficient estimates from a patient-visit-level regression including year fixed effects and controls for age category, gender, advanced HIV at initiation, and underweight at initiation, and, in Panel B, whether the patient missed their last scheduled appointment.

Appendix Table A1

Qualitative Data from Interviews

Theme	Qualitative Evidence
EMR Features, Linking Data to Patient IDs, Back-Entry	"Client's health passport is where we paste the client's visits, all the vitals, and also next appointment so that the client should know when she/he supposed to go to the facility to get drugs. The information we paste in the book is the same information we paste on the master card. [When a patient loses her/his health passport] It's very easy for us. We just search hers/his name in the system, and the system brings out the information of the client." **-- ART Clerk**
	"Sometimes the master cards can be missing and makes it difficult to access certain information and when a client has come to the clinic with no information and it becomes difficult to start searching the master card with no enough information. These clients come to the clinic and they will only tell you their name but when you asks them the number which is in the master card they will say that they have forgotten and you begin searching by looking individual cards which is very difficult. Sometimes they leave the health passport home, while on the EMR you can just punch a button and it brings all the information." **-- Nurse #1**
	"Currently, the system is localized. The fact that I am on ART it's only Naisi system that knows that. It has not yet migrated, but plans are there that if a client upon scanning of his barcode whether he was getting drugs from Lilongwe and now he is in Zomba when you scan the barcode, it will be able to show you where he is getting drugs from and all that." **-- Clinician #2**
	"EMR reminds you things that you are supposed to do to a client. Even a patient who has a high viral load, once you enter the Respondent ID, it indicates to you that this person has a high viral load. It blinks because of the counseling sessions, and it assists on how to assist every patient coming to the clinic." **-- Nurse #1**
	"(…) the coming of EMR reduced time taken to treat a patient. There is some information on the vitals from the patients which the health provider forgets to collect from a patient, and the EMR reminds the provider to collect such information (…)" **-- EMR Data Statistician**
	"[When asked which data has been back entered in EMR] We were entering all the data the facility had. For instance, if the facility had a cohort of more than 2000, we would enter data from patient one up to patient 2000. We were doing back data entry. (…) we were entering data in chronological order. (…) You were guided by the card to enter everything that was on the card until everything has been entered." **-- NGO Strategic Information and Evaluation Officer**
	[When asked which data was back-entered into EMR] "I can say it was all the data; let us say; if the clinic started its operations in 2008; that means we will register all the patients from our patient number one. So we need to enter all the information of the patients and their outcomes whether they died or they transferred" **-- Clinician #4**

Notes: Qualitative data from 50 semi-structured interviews with clinic staff (nurses, physicians, data clerks, coordinators and tracing staff), MoH employees, NGO collaborators, and local community leaders. Interviews conducted over the phone December 2020 to March 2021 due to COVID-19 restrictions. Interviews recorded and transcribed verbatim. Topics included EMRs and clinic management, patient tracing, clinic quality, number of initiations, lapse from care, and deaths. The table includes an informative selection of quotes from a total of 380 pages transcribed, with information triangulated across respondents and validated by policy documents.

Theme	Qualitative Evidence
Improvements in Efficiency, Quality of Care and Outcomes due to EMR	"[EMR was introduced] mainly it was to ease the burden of data management especially in the terms of reporting but later on it was embedded with decision making abilities to help the clinicians with their work; for example, if they want to decide which regimen they should give the clients and also it should remind them when to take the viral load and other things. So the main reasons were one; the report writing and decision making ability. (…) right now the goals [of EMR] have been met and the cohort has grown in number of patients alive on ART. It is a very huge number and cannot be managed on paper " **– Government Monitoring and Evaluation Officer**
	"Yeah, in a way, because in EMR, there are things such as signs and symptoms of other diseases are opportunistic infections, or to say infections that get advantage because of low immunity, such as Tuberculosis. Because when clients come, the system reminds you to ask if the client is coughing or if the client has a fever. So in such a way, the client is detected early, also starts treatment early, and the outcomes come out good that we prevent deaths. So the system has really helped. Even things like weight, the system already calculates that this one is going into severe malnutrition, and it alerts you to do something. So adult malnutrition is dangerous and can cause death. So the system helps us to detect malnutrition in people." **-- Physician and Hospital Manager**
	"EMR helps in doing your work very fast, and it is consistent; instead of searching through the master cards and the files, you just use the EMR, and it brings whatever you want. Sometimes the master cards can be missing and make it difficult to access certain information, and when a client has come to the clinic with no information, and it becomes difficult to start searching the master card with no enough information. These clients come to the clinic, and they will only tell you their name, but when you ask them the number which is in the master card, they will say that they have forgotten, and you begin searching by looking at individual cards, which is very difficult. Sometimes they leave the health passport home, while on the EMR you can just punch a button, and it brings all the information" **-- Nurse #1**
	"The system is able to flag out the people that need to check viral load. (…) when you are entering the data into the EMR, the system is able to remind you to get the viral load from the client. Another way is that the system is able to show you the type of drug the client is supposed to get according to the weight of the client. You don't have to go to the pharmacy to ask because the system is able to show you which is helping patients in care." **-- NGO Strategic Information and Evaluation Officer**
	"Indirectly [EMR contributed to reducing deaths] in the sense that first, as I said that, it created a systematic review of the client. So since we are following the client systematically, the issues such as malnutrition can be elicited fast, issues of TB can be diagnosed fast, issues of high viral load can be elicited fast, issues of drug combinations that you cannot mix such kind of drugs the EMR is able to let you know. So in that direction, I would say that it has helped to improve the quality of life for clients. Secondly, by giving us the number of defaulters, it helps us to trace those people, and in the long-run, we put those people back in care. (…) So it did not help prevent deaths directly, but indirectly it has improved the quality of life for these clients and thereby prevents other deaths." **-- Clinic Coordinator**
	"Another advantage [of EMR] is that data is safe because, with the hard copies, you could find that pages have been torn and some data has been lost, and it was very difficult for us to trace when the client started drugs. (…) when we are at the consultation for us to give the client drugs, the time we are taking is very little. Because the system already calculates. The drugs, time, next appointment, everything is already calculated in the system. So a patient does not take long to get drugs at a facility. (…) in the past, we were having trouble with calculating. In those days, we lacked calendars, so we had to get calculators and do mathematics until we help the client. But now it's straight forward. When we are at the consultation, we are having a wonderful time with the clients." **-- ART Clerk**
	"Now, when you want to generate a report, you don't bother to get the master cards and the register; you just go into the system and click buttons, and you download a report. So within 30 minutes, you are able to finish the report, unlike in the past with the paper-based system, it could take three days, including nights, to write a report. Similarly, when you are attending a client, you spend less time with him. You don't bother to write anything because things are already in the system. So clients are being helped faster." **-- NGO District Technical Officer**
	"More clients are coming because as of now our cohort reached 5000 from 4000. (...) in general even the clients themselves are happy with the system because what they cite include; the speed of service delivery, and secondly their heath passports are preserved. This is so because the print outs that we paste in the passport saves space than the manual way of writing using a pen. But so far what I can say is that most people are coming to ART" **-- Nurse #2**
	"I would say more clients came because we started doing things a little faster than the way we used to because writing in the cards was time-consuming. So since some clients would be in a hurry to work, to school, they were opting not to come because they would feel that they will be late. That's just an opinion. But now, with the EMR, we are a bit faster, and I would say more clients have come. Secondly, with the EMR, it's easy to elicit those clients who have defaulted treatment, so it's easy to trace them. So basically, I would say more clients have come because of the EMR." **-- Clinic Coordinator**

Notes: Qualitative data from 50 semi-structured interviews with clinic staff (nurses, physicians, data clerks, coordinators and tracing staff), MoH employees, NGO collaborators, and local community leaders. Interviews conducted over the phone December 2020 to March 2021 due to COVID-19 restrictions. Interviews recorded and transcribed verbatim. Topics included EMRs and clinic management, patient tracing, clinic quality, number of initiations, lapse from care, and deaths. The table includes an informative selection of quotes from a total of 380 pages transcribed, with information triangulated across respondents and validated by policy documents.

Appendix Table A2
EMR Implementation Detailed Cost Structure

EMR Cost Structure		
Workstations		Unit Cost
Touchscreen Computer	$	570.00
Printer	$	905.00
Scanner	$	150.00
Integrated Computer	$	550.00
Dashboard	$	320.00
2D Barcode Scanner	$	50.00
Cables and Aditional Equipment	$	1,007.00
Total Workstation	*$*	*3,552.00*
Server Cabinet		*Unit Cost*
Server	$	8,100.00
Cisco Switch	$	300.00
Server Cabinet	$	725.00
Cables and Aditional Equipment	$	237.00
Total Server Cabinet	*$*	*9,362.00*
Power Cabinet		*Unit Cost*
Batteries (200Ah and 96Ah) and Battery Charger	$	710.00
Solar Panels (300 Watts) and Solar Panel Charger Controller (MPPT)	$	744.00
Charger/Inverter	$	1,965.00
Circuit breakers, Switch, Voltage Monitor, and Additional Equipment	$	1,026.00
Cisco SPA122 ATA with Router for VOIP Phone	$	100.00
Total Power Cabinet	*$*	*4,545.00*
Power and Network Cabling		*Unit Cost*
Remote Site Monitoring: (SPM-200 Site Power Monitor)	$	900.00
CAT-6 Ethernet Cable (Indoor and Outdoor)	$	547.00
Installation Equipments	$	1,244.00
Additional Cables, Materials, and Equipments	$	412.00
Total Power and Network Cabling	*$*	*3,103.00*
Connectivity Equipment		*Unit Cost*
Tower	$	4,000.00
Sector Antenna (2.4Ghz and 5Ghz)	$	449.00
Cisco Mikrotic Router	$	200.00
Rocket5ac PtP Radio	$	219.00
Additional Equipment	$	940.00
Total Connectivity Equipment	*$*	*5,808.00*
Other Supplies		*Unit Cost*
Miscleneous hardware (frames etc)	$	100.00
Mastercards, Ribbons, Labels, and Other Materials	$	476.00
Total Other Supplies	*$*	*576.00*
Total Cost for an Average Clinic (3 Workstations)	**$**	**34,050.00**
Total Number of Clinics in the Program		106
Total EMR Program Cost (Estimated)	**$**	**3,609,300.00**

Notes: Data provided by Baobab Health Trust and Driessen et al. (2013).

Appendix Table A3

Descriptive Statistics

Year	Number of Clinics in the Database	Proportion of Clinics that Adopted EMR	Proportion of Hospitals that Adopted EMR	Proportion of Urban Clinics that Adopted EMR	Average Number of Deaths (all clinics)	Average Number of New Patients (all clinics)	Average Number of Deaths among EMR Clinics	Average Number of Patients In Care among EMR Clinics	Average Number of New Patients among EMR Clinics	Average Number of Retained Patients among EMR Clinics
2003	4	0.0%	0.0%	0.0%	4.0	13.3	N/A	N/A	N/A	N/A
2004	12	0.0%	0.0%	0.0%	29.9	75.3	N/A	N/A	N/A	N/A
2005	33	0.0%	0.0%	0.0%	49.8	167.4	N/A	N/A	N/A	N/A
2006	51	0.0%	0.0%	0.0%	66.6	281.8	N/A	N/A	N/A	N/A
2007	57	1.8%	3.3%	3.0%	79.4	382.9	146.0	4854.0	586.0	4810.0
2008	67	3.0%	3.1%	5.7%	68.4	405.9	141.0	5926.5	574.0	5855.0
2009	75	4.0%	6.3%	8.3%	60.8	431.0	112.0	5640.0	872.3	5448.0
2010	83	10.8%	21.9%	21.6%	57.3	407.4	123.9	5164.1	1031.0	5040.7
2011	97	17.5%	36.4%	40.0%	48.1	411.7	103.7	4779.1	844.6	4609.2
2012	106	22.6%	48.5%	52.5%	42.5	409.4	92.5	5147.4	675.9	5035.3
2013	106	30.2%	57.6%	65.0%	37.2	353.2	71.2	4977.0	594.5	4846.6
2014	106	38.7%	69.7%	70.0%	34.3	369.6	56.8	4788.4	542.3	4681.0
2015	106	48.1%	84.8%	77.5%	36.5	360.3	55.2	4647.3	503.1	4539.5
2016	106	51.9%	90.9%	77.5%	35.5	443.8	50.7	4885.8	580.3	4766.2
2017	106	74.5%	97.0%	92.5%	34.8	487.9	41.4	4358.6	566.9	4221.8
2018	106	77.4%	97.0%	92.5%	27.8	373.1	32.3	4237.9	415.5	4101.4
2019	106	100.0%	100.0%	100.0%	N/A	N/A	N/A	N/A	N/A	N/A

Notes: This table describes the dataset used for analysis. Averages are measured at the clinic level.

Appendix Table A4

Predictors of Patient Lapse Pre-EMR

	(1)	(2)	(3)	(4)	(5)	(6)
	Lapse (in Months)					
	6 months	12 months	18 months	24 months	30 months	36 months
Years Since Initiation	0.0583***	0.0692***	0.0780***	0.0838***	0.0880***	0.0906***
	(0.000307)	(0.000297)	(0.000286)	(0.000273)	(0.000259)	(0.000244)
Male	0.0339***	0.0295***	0.0271***	0.0233***	0.0204***	0.0189***
	(0.00156)	(0.00151)	(0.00146)	(0.00139)	(0.00132)	(0.00124)
Age 10-17	-0.0803***	-0.0793***	-0.0698***	-0.0603***	-0.0494***	-0.0399***
	(0.00522)	(0.00505)	(0.00487)	(0.00464)	(0.00441)	(0.00415)
Age18 to 45	-0.0888***	-0.0788***	-0.0684***	-0.0575***	-0.0486***	-0.0401***
	(0.00357)	(0.00346)	(0.00333)	(0.00318)	(0.00302)	(0.00284)
Age ≥ 50	-0.148***	-0.136***	-0.124***	-0.109***	-0.0969***	-0.0842***
	(0.00399)	(0.00386)	(0.00372)	(0.00354)	(0.00337)	(0.00317)
Underweight at Initiation	-0.0729***	-0.0687***	-0.0625***	-0.0626***	-0.0611***	-0.0580***
	(0.00208)	(0.00202)	(0.00194)	(0.00185)	(0.00176)	(0.00166)
Advanced HIV at Initiation	0.0294***	0.0319***	0.0305***	0.0399***	0.0406***	0.0368***
	(0.00158)	(0.00153)	(0.00147)	(0.00140)	(0.00133)	(0.00126)
Hospital	0.0571***	0.0494***	0.0427***	0.0310***	0.0256***	0.0195***
	(0.00159)	(0.00154)	(0.00149)	(0.00142)	(0.00135)	(0.00127)
Large Clinic	-0.0183***	-0.0164***	-0.0144***	-0.0198***	-0.0215***	-0.0228***
	(0.00179)	(0.00173)	(0.00167)	(0.00159)	(0.00151)	(0.00142)
Urban Clinic	0.0723***	0.0622***	0.0537***	0.0418***	0.0330***	0.0269***
	(0.00178)	(0.00173)	(0.00166)	(0.00158)	(0.00150)	(0.00142)
Observations	408,624	408,624	408,624	408,624	408,624	408,624
R-squared	0.101	0.139	0.178	0.218	0.255	0.289

Notes: We regress an indicator for patient lapse, measured at the individual-year level, on patient and clinic-level covariates. Sample restricted to the pre-EMR period. No fixed effects. Standard errors in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A5

Dynamic Effect of EMR Adoption on Number of Patient Deaths

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	log(Patient Deaths)						
		Gender		Age			
Event Year	All	Male	Female	≤ 9	10 to 17	18 to 49	≥ 50
k = -6	-0.00627	0.00295	-0.0253	-0.0416	-0.0310	-0.00155	-0.0235
	(0.0681)	(0.0704)	(0.0683)	(0.0741)	(0.0590)	(0.0690)	(0.0669)
k = -5	-0.00178	0.0266	-0.0150	0.0282	0.0753	-0.00621	-0.0121
	(0.0664)	(0.0658)	(0.0657)	(0.0685)	(0.0574)	(0.0697)	(0.0612)
k = -4	0.00869	0.0335	-0.0106	-0.0288	0.0537	0.00383	0.0224
	(0.0585)	(0.0625)	(0.0623)	(0.0638)	(0.0582)	(0.0633)	(0.0579)
k = -3	-0.00405	0.0301	-0.0376	0.00519	0.0142	-0.00515	0.0172
	(0.0526)	(0.0556)	(0.0581)	(0.0654)	(0.0626)	(0.0551)	(0.0628)
k = -2	0.0262	0.0743	-0.0181	0.0688	-0.0422	0.0202	0.111*
	(0.0508)	(0.0511)	(0.0611)	(0.0675)	(0.0666)	(0.0513)	(0.0600)
k = -1	0	0	0	0	0	0	0
k = 0	-0.0602	0.000870	-0.118*	-0.156*	0.0553	-0.109	0.212***
	(0.0613)	(0.0650)	(0.0699)	(0.0823)	(0.0754)	(0.0681)	(0.0686)
k = 1	-0.124	-0.0765	-0.134	-0.203**	-0.140	-0.116	0.0451
	(0.0819)	(0.0798)	(0.0832)	(0.0890)	(0.0878)	(0.0869)	(0.0921)
k = 2	-0.153	-0.123	-0.140	-0.235**	-0.115	-0.149	0.0815
	(0.0946)	(0.0898)	(0.102)	(0.104)	(0.102)	(0.0982)	(0.0928)
k = 3	-0.225*	-0.163	-0.243**	-0.311***	-0.0468	-0.233*	0.0556
	(0.117)	(0.110)	(0.122)	(0.112)	(0.120)	(0.120)	(0.1000)
k = 4	-0.206	-0.241**	-0.114	-0.435***	-0.197	-0.228*	0.192*
	(0.127)	(0.113)	(0.140)	(0.142)	(0.120)	(0.134)	(0.115)
k = 5	-0.333**	-0.259**	-0.419**	-0.576***	-0.123	-0.397***	0.223
	(0.138)	(0.130)	(0.175)	(0.159)	(0.135)	(0.136)	(0.141)
k ≥ 6	-0.583***	-0.539***	-0.568***	-0.866***	-0.284*	-0.610***	-0.0417
	(0.157)	(0.144)	(0.163)	(0.182)	(0.163)	(0.160)	(0.129)
F-test on pre-trends	0.109	0.461	0.132	0.474	1.178	0.0624	0.838
p-value of F-test	0.990	0.805	0.985	0.795	0.325	0.997	0.526
Observations	1,166	1,166	1,166	1,166	1,166	1,166	1,166
R-squared	0.795	0.785	0.771	0.572	0.538	0.784	0.723

Notes: Patient Deaths is the number of ART patients that died, measured at the clinic-year level. We add 1 to the outcome measure. $k = 0$ is the year immediately following EMR implementation. Each regression includes year and clinic fixed effects following (2). Heteroskedasticity-robust standard errors clustered at the clinic level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A6

Effect of EMR System on Patient Lapse Probability, With and Without Controls

	(1)	(2)	(3)	(4)	(5)	(6)
	Lapse (in Months)					
Panel (A)	Without Controls					
	6 months	12 months	18 months	24 months	30 months	36 months
Post-EMR	-0.0369***	-0.0503***	-0.0610***	-0.0551***	-0.0535***	-0.0469***
	(0.00892)	(0.00900)	(0.00926)	(0.00845)	(0.00983)	(0.0104)
Controls	N	N	N	N	N	N
Observations	3,475,826	3,475,826	3,475,826	3,475,826	3,475,826	3,475,826
R-squared	0.025	0.030	0.034	0.038	0.043	0.048
Panel (B)	With Controls					
	6 months	12 months	18 months	24 months	30 months	36 months
Post-EMR	-0.0325***	-0.0453***	-0.0556***	-0.0495***	-0.0479***	-0.0412***
	(0.00996)	(0.0102)	(0.00982)	(0.00821)	(0.00905)	(0.00949)
Controls	Y	Y	Y	Y	Y	Y
Observations	3,475,609	3,475,609	3,475,609	3,475,609	3,475,609	3,475,609
R-squared	0.117	0.156	0.191	0.221	0.248	0.271

Notes: *Patient Lapse* is measured at the individual patient-year (t) level, on the sample of patients who have initiated care by year t. It is an indicator for no clinic visit in the m months prior to the end of year t. Regressions include year and clinic fixed effects as well as controls (in Panel B), as in (3). Controls include gender, age category, years since initiation, an indicator for whether the patient was underweight at initiation, and an indicator for whether the patient had advanced HIV at initiation. Heteroskedasticity-robust standard errors clustered at the clinic level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A7

Effect of EMR System on Patient Lapse Probability by Gender

	(1)	(2)	(3)	(4)	(5)	(6)
	Lapse (in Months)					
Panel (A)	Male					
	6 months	12 months	18 months	24 months	30 months	36 months
Post-EMR	-0.0236**	-0.0376***	-0.0488***	-0.0443***	-0.0428***	-0.0379***
	(0.00965)	(0.00985)	(0.00933)	(0.00816)	(0.00898)	(0.00965)
Controls	Y	Y	Y	Y	Y	Y
Observations	1,307,118	1,307,118	1,307,118	1,307,118	1,307,118	1,307,118
R-squared	0.126	0.168	0.204	0.236	0.263	0.286
Panel (B)	Female					
	6 months	12 months	18 months	24 months	30 months	36 months
Post-EMR	-0.0379***	-0.0500***	-0.0597***	-0.0526***	-0.0508***	-0.0431***
	(0.0102)	(0.0105)	(0.0102)	(0.00838)	(0.00928)	(0.00960)
Controls	Y	Y	Y	Y	Y	Y
Observations	2,168,491	2,168,491	2,168,491	2,168,491	2,168,491	2,168,491
R-squared	0.112	0.149	0.183	0.213	0.238	0.261

Notes: *Patient Lapse* is measured at the individual patient-year (t) level, on the sample of patients who have initiated care by year t. It is an indicator for no clinic visit in the m months prior to the end of year t. Regressions include year and clinic fixed effects as well as controls, as in (3). Controls include gender, age category, years since initiation, an indicator for whether the patient was underweight at initiation, and an indicator for whether the patient had advanced HIV at initiation. Heteroskedasticity-robust standard errors clustered at the clinic level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A8

Effect of EMR System on Patient Lapse Probability by Age

	(1)	(2)	(3)	(4)	(5)	(6)
	Lapse (in Months)					
Panel (A)	Age ≤ 9					
	6 months	12 months	18 months	24 months	30 months	36 months
Post-EMR	-0.0346***	-0.0750***	-0.0903***	-0.0796***	-0.0741***	-0.0601***
	(0.0105)	(0.0102)	(0.00946)	(0.00781)	(0.00763)	(0.00782)
Controls	Y	Y	Y	Y	Y	Y
Observations	128,777	128,777	128,777	128,777	128,777	128,777
R-squared	0.152	0.209	0.255	0.295	0.324	0.344
Panel (B)	Age 10 to 17					
	6 months	12 months	18 months	24 months	30 months	36 months
Post-EMR	-0.0525***	-0.0638***	-0.0736***	-0.0656***	-0.0652***	-0.0591***
	(0.0133)	(0.0125)	(0.0106)	(0.00920)	(0.00942)	(0.00965)
Controls	Y	Y	Y	Y	Y	Y
Observations	140,466	140,466	140,466	140,466	140,466	140,466
R-squared	0.150	0.191	0.223	0.250	0.271	0.290
Panel (C)	Age 18 to 49					
	6 months	12 months	18 months	24 months	30 months	36 months
Post-EMR	-0.0343***	-0.0464***	-0.0559***	-0.0492***	-0.0470***	-0.0402***
	(0.0101)	(0.0100)	(0.00948)	(0.00769)	(0.00835)	(0.00860)
Controls	Y	Y	Y	Y	Y	Y
Observations	2,627,400	2,627,400	2,627,400	2,627,400	2,627,400	2,627,400
R-squared	0.118	0.159	0.196	0.229	0.256	0.280
Panel (D)	Age ≥ 50					
	6 months	12 months	18 months	24 months	30 months	36 months
Post-EMR	-0.0211**	-0.0319***	-0.0448***	-0.0433***	-0.0448***	-0.0406***
	(0.00975)	(0.0105)	(0.0106)	(0.00944)	(0.0108)	(0.0121)
Controls	Y	Y	Y	Y	Y	Y
Observations	578,966	578,966	578,966	578,966	578,966	578,966
R-squared	0.114	0.137	0.158	0.178	0.197	0.215

Notes: *Patient Lapse* is measured at the patient-year level, on sample of patients who initiated care by year t. It is an indicator for no clinic visit in the m months prior to the end of year t. Regressions include year and clinic fixed effects as well as controls, as in (3). Controls include gender, age category, years since initiation, an indicator for whether the patient was underweight at initiation, and an indicator for whether the patient had advanced HIV at initiation. Heteroskedasticity-robust standard errors clustered at the clinic level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A9

Dynamic Effect of EMR Adoption on Number of Patients In Care

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	log(Patients In Care)						
	All	Gender		Age			
Event Year		Male	Female	≤ 9	10 to 17	18 to 49	≥ 50
k = -1	0	0	0	0	0	0	0
k = 0	0.160***	0.160***	0.162***	0.249***	0.345***	0.163***	0.132***
	(0.0179)	(0.0197)	(0.0179)	(0.0452)	(0.109)	(0.0207)	(0.0141)
k = 1	0.206***	0.204***	0.208***	0.339***	0.598***	0.212***	0.155***
	(0.0287)	(0.0309)	(0.0293)	(0.0737)	(0.212)	(0.0331)	(0.0236)
k = 2	0.241***	0.241***	0.241***	0.422***	0.809**	0.248***	0.171***
	(0.0423)	(0.0463)	(0.0423)	(0.0941)	(0.316)	(0.0490)	(0.0330)
k = 3	0.272***	0.273***	0.273***	0.487***	1.004**	0.281***	0.187***
	(0.0554)	(0.0618)	(0.0548)	(0.117)	(0.421)	(0.0649)	(0.0398)
k = 4	0.300***	0.300***	0.301***	0.514***	1.189**	0.311***	0.210***
	(0.0690)	(0.0772)	(0.0683)	(0.142)	(0.527)	(0.0817)	(0.0479)
k = 5	0.316***	0.318***	0.316***	0.494***	1.390**	0.326***	0.226***
	(0.0874)	(0.0966)	(0.0868)	(0.170)	(0.630)	(0.103)	(0.0588)
k ≥ 6	0.335***	0.345***	0.330***	0.530**	1.583**	0.348***	0.229***
	(0.110)	(0.119)	(0.109)	(0.209)	(0.742)	(0.130)	(0.0717)
F-test on pre-trends	N/A	N/A	N/A	N/A	N/A	N/A	N/A
p-value of F-test	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Observations	411	411	411	411	411	411	411
R-squared	0.985	0.984	0.985	0.951	0.968	0.984	0.988

Notes: *Patients In Care* is the number of ART patients that visited the clinic at least once, measured at the clinic-year level. We add 1 to the outcome measure. $k = 0$ is the year immediately following EMR implementation. Each regression includes year and clinic fixed effects following (2). * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A10

Dynamic Effect of EMR Adoption on Number of New Patients

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	log(New Patients)						
	All	Gender		Age			
Event Year		Male	Female	≤ 9	10 to 17	18 to 49	≥ 50
k = -6	0.0444	-0.0146	0.0560	-0.0123	0.00677	0.0618	-0.0117
	(0.0711)	(0.0803)	(0.0714)	(0.0900)	(0.0718)	(0.0709)	(0.0707)
k = -5	0.0458	0.0180	0.0746	0.100	0.124**	0.0554	-0.000635
	(0.0574)	(0.0556)	(0.0583)	(0.0678)	(0.0616)	(0.0574)	(0.0564)
k = -4	-0.0622	-0.0666	-0.0258	0.0758	-0.0920	-0.0505	-0.00410
	(0.0668)	(0.0617)	(0.0638)	(0.0710)	(0.0688)	(0.0655)	(0.0565)
k = -3	-0.0859	-0.0831	-0.0534	0.0610	0.0320	-0.0696	-0.0336
	(0.0731)	(0.0665)	(0.0707)	(0.0656)	(0.0607)	(0.0708)	(0.0637)
k = -2	-0.0506	-0.0262	-0.0354	0.0502	-0.0290	-0.0382	-0.0390
	(0.0637)	(0.0575)	(0.0596)	(0.0567)	(0.0617)	(0.0619)	(0.0546)
k = -1	0	0	0	0	0	0	0
k = 0	0.112	0.101	0.108	-0.0429	-0.105	0.141	0.0960
	(0.117)	(0.106)	(0.118)	(0.0978)	(0.0880)	(0.115)	(0.0998)
k = 1	0.148	0.128	0.143	0.00414	0.00516	0.149	0.0593
	(0.152)	(0.140)	(0.149)	(0.135)	(0.111)	(0.147)	(0.130)
k = 2	0.250	0.182	0.249	-0.0612	-0.120	0.278	0.175
	(0.200)	(0.182)	(0.195)	(0.146)	(0.121)	(0.195)	(0.167)
k = 3	0.221	0.142	0.236	-0.00353	-0.0391	0.249	0.161
	(0.240)	(0.219)	(0.235)	(0.172)	(0.155)	(0.235)	(0.198)
k = 4	0.0393	-0.0156	0.0689	-0.164	-0.201	0.0816	0.0377
	(0.275)	(0.249)	(0.266)	(0.196)	(0.165)	(0.266)	(0.216)
k = 5	0.0528	-0.0245	0.0807	-0.258	-0.273	0.0943	0.0789
	(0.355)	(0.321)	(0.339)	(0.237)	(0.205)	(0.347)	(0.258)
k ≥ 6	0.186	0.0859	0.220	-0.254	-0.195	0.226	0.242
	(0.461)	(0.414)	(0.449)	(0.287)	(0.255)	(0.451)	(0.331)
F-test on pre-trends	1.143	1.134	1.093	0.878	2.665	1.190	0.198
p-value of F-test	0.343	0.347	0.369	0.499	0.0261	0.319	0.963
Observations	1,166	1,166	1,166	1,166	1,166	1,166	1,166
R-squared	0.786	0.781	0.786	0.737	0.722	0.790	0.741

Notes: *New Patients* is the number of new ART patients that initiated ART for the first time, measured at the clinic-year level. We add 1 to the outcome measure. Each regression includes year and clinic fixed effects following (2). Heteroskedasticity-robust standard errors clustered at the clinic level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A11

Dynamic Effect of EMR Adoption on Patient Retention

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	log(Retention)						
	All	Gender		Age			
Event Year		Male	Female	≤ 9	10 to 17	18 to 49	≥ 50
k = -1	0	0	0	0	0	0	0
k = 0	0.226***	0.231***	0.226***	0.351***	0.434***	0.229***	0.185***
	(0.0207)	(0.0250)	(0.0197)	(0.0429)	(0.122)	(0.0226)	(0.0204)
k = 1	0.297***	0.295***	0.300***	0.473***	0.701***	0.305***	0.237***
	(0.0328)	(0.0374)	(0.0323)	(0.0735)	(0.236)	(0.0360)	(0.0318)
k = 2	0.354***	0.346***	0.360***	0.591***	0.957***	0.364***	0.273***
	(0.0462)	(0.0534)	(0.0447)	(0.0942)	(0.350)	(0.0515)	(0.0417)
k = 3	0.410***	0.404***	0.415***	0.652***	1.159**	0.424***	0.314***
	(0.0598)	(0.0700)	(0.0571)	(0.116)	(0.467)	(0.0673)	(0.0510)
k = 4	0.464***	0.452***	0.473***	0.716***	1.376**	0.483***	0.360***
	(0.0715)	(0.0846)	(0.0682)	(0.133)	(0.582)	(0.0820)	(0.0592)
k = 5	0.504***	0.496***	0.512***	0.728***	1.613**	0.523***	0.399***
	(0.0879)	(0.103)	(0.0849)	(0.163)	(0.694)	(0.101)	(0.0719)
k ≥ 6	0.548***	0.543***	0.553***	0.782***	1.817**	0.572***	0.429***
	(0.111)	(0.127)	(0.108)	(0.204)	(0.821)	(0.127)	(0.0894)
F-test on pre-trends	N/A	N/A	N/A	N/A	N/A	N/A	N/A
p-value of F-test	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Observations	411	411	411	411	411	411	411
R-squared	0.983	0.982	0.982	0.946	0.962	0.981	0.983

Notes: *Patient Retention* is the number of returning (did not initiate for the first time that year) ART patients that visited the clinic, measured at the clinic-year level. We add 1 to the outcome measure. $k = 0$ is the year immediately following EMR implementation. Each regression includes year and clinic fixed effects following (2). Heteroskedasticity-robust standard errors clustered at the clinic level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A12

Effect of EMR System on Patient-Level Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Missed Next Appointment						
	All	Gender		Age			
		Male	Female	≤ 9	10 to 17	18 to 49	≥ 50
Post-EMR	-0.0712***	-0.0724***	-0.0706***	-0.0834***	-0.0916***	-0.0707***	-0.0667***
	(0.0180)	(0.0177)	(0.0183)	(0.0223)	(0.0167)	(0.0185)	(0.0170)
Long Run Post-EMR	-0.0511***	-0.0511***	-0.0511***	-0.0700***	-0.0553***	-0.0502***	-0.0427***
	(0.00434)	(0.00400)	(0.00478)	(0.0116)	(0.00533)	(0.00466)	(0.00554)
Controls	Y	Y	Y	Y	Y	Y	Y
Observations	7,422,750	2,659,951	4,762,799	267,219	335,752	5,712,703	1,107,074
R-squared	0.031	0.031	0.031	0.055	0.028	0.027	0.025

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Underweight						
	All	Gender		Age			
		Male	Female	≤ 9	10 to 17	18 to 49	≥ 50
Post-EMR	-0.00288	-0.00181	-0.00375	-0.0193*	0.00311	-0.00471	0.0130**
	(0.00382)	(0.00427)	(0.00400)	(0.0103)	(0.0107)	(0.00423)	(0.00506)
Long Run Post-EMR	-0.0146***	-0.0252***	-0.00855***	-0.0332***	-0.0356***	-0.0117***	-0.0199***
	(0.00277)	(0.00353)	(0.00261)	(0.00584)	(0.00508)	(0.00301)	(0.00318)
Controls	Y	Y	Y	Y	Y	Y	Y
Observations	7,882,940	2,828,415	5,054,525	299,020	352,169	6,069,572	1,162,179
R-squared	0.207	0.221	0.200	0.132	0.096	0.214	0.244

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Viral Load						
	All	Gender		Age			
		Male	Female	≤ 9	10 to 17	18 to 49	≥ 50
Post-EMR	-7.941*	-2.569	-11.56*	-13.09	-7.557	-12.03**	3.515
	(4.601)	(7.186)	(5.914)	(13.55)	(5.968)	(5.607)	(7.506)
Long Run Post-EMR	-0.925	1.076	-0.930	5.141	-2.924	-1.373	2.254
	(1.520)	(3.827)	(1.154)	(5.855)	(5.299)	(1.879)	(4.447)
Controls	Y	Y	Y	Y	Y	Y	Y
Observations	82,338	27,705	54,627	2,084	4,036	60,918	15,290
R-squared	0.003	0.007	0.004	0.032	0.018	0.003	0.013

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Viral Suppression						
	All	Gender		Age			
		Male	Female	≤ 9	10 to 17	18 to 49	≥ 50
Post-EMR	-0.00546	0.000775	-0.00609	-0.120	0.0448	-0.00683	0.0315
	(0.0456)	(0.0491)	(0.0445)	(0.0883)	(0.0467)	(0.0486)	(0.0389)
Long Run Post-EMR	-0.00244	-0.00308	-0.00331	0.0855**	0.0325	-0.00841	-0.0175
	(0.0157)	(0.0224)	(0.0168)	(0.0418)	(0.0308)	(0.0157)	(0.0223)
Controls	Y	Y	Y	Y	Y	Y	Y
Observations	82,338	27,705	54,627	2,084	4,036	60,918	15,290
R-squared	0.500	0.461	0.522	0.211	0.255	0.516	0.524

Notes: *Missed next appointment* is at least 2 months late. *Underweight* is an indicator for having a BMI more than 2 SD below median, according to WHO. *Viral Load* is measured in 1000 particles per ml of blood. *Viral Suppression*: < 200 copies of HIV per ml of blood. Outcome is at patient-visit level. Regressions include year and clinic fixed effects, controls, as in (3). Controls include gender, age category, years since initiation, an indicator for whether the patient was underweight at initiation, and an indicator for whether the patient had advanced HIV at initiation. Long-Run Post-EMR is an indicator for visits after the first visit Post-EMR. Heteroskedasticity-robust standard errors clustered at the clinic level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A13

Medical Care Provided at Visit, Without Controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	Days of Medication Provided at Initiation		Child Underweight at 3-6 Months		TB treatment at Initiation		All Referrals at Initiaiton		Ordered CD4 Count at Initiation		Underweight at Initiation		Advanced HIV at Initiation	
Post-EMR	4.436***	5.090***	0.0271	0.0455	-0.00193	-0.00740	-0.0459	-0.0212	0.00211	-0.000265	0.0131	0.0160*	0.0338	0.0923
	(0.863)	(1.550)	(0.0337)	(0.0403)	(0.00644)	(0.00807)	(0.0317)	(0.0347)	(0.0245)	(0.0243)	(0.00943)	(0.00961)	(0.0372)	(0.0725)
Hospital Post-EMR		-1.525		-0.0447		0.0120		-0.0627		0.00570		-0.00669		-0.126
		(3.727)		(0.0587)		(0.00880)		(0.0465)		(0.0316)		(0.0149)		(0.103)
Controls	N	N	N	N	N	N	N	N	N	N	N	N	N	N
Observations	733,999	719,884	3,267	3,256	755,175	741,022	755,175	741,022	755,175	741,022	755,175	741,022	755,175	741,022
R-squared	0.113	0.114	0.062	0.061	0.016	0.017	0.158	0.159	0.121	0.120	0.023	0.023	0.234	0.235

Notes: (1-2): the number of days' worth of medication provided at initiation. (3-4): indicator for an underweight child being underweight at 3-6 month follow-up visit. (5-6): indicator for TB treatment at initiation. (7-8): number of referrals at initiation. (9-10): indicator for CD4 count order at initiation. Outcomes measured at individual-visit level. Regressions include year and clinic fixed effects as in (3). Heteroskedasticity-robust standard errors clustered at the clinic level, in parentheses: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A14

Robustness: Heterogeneity by Clinic Type (Hospital vs. Clinic)

	(1)	(2)	(3)	(4)
Event Year	log(Patient Deaths)	log(Patients In Care)	log(New Patients)	log(Retention)
k = -1 × Hospital	0	0	0	0
k = 0 × Hospital	-0.0784	0.136***	0.441	0.195***
	(0.0839)	(0.0222)	(0.327)	(0.0271)
k = 1 × Hospital	-0.178*	0.161***	0.492	0.239***
	(0.0973)	(0.0354)	(0.335)	(0.0399)
k = 2 × Hospital	-0.282***	0.180***	0.618	0.286***
	(0.102)	(0.0504)	(0.377)	(0.0539)
k = 3 × Hospital	-0.369***	0.190***	0.592	0.316***
	(0.122)	(0.0638)	(0.421)	(0.0656)
k = 4 × Hospital	-0.374***	0.204**	0.190	0.359***
	(0.117)	(0.0789)	(0.377)	(0.0784)
k = 5 × Hospital	-0.452***	0.216**	0.142	0.406***
	(0.131)	(0.0963)	(0.438)	(0.0953)
k ≥ 6 × Hospital	-0.597***	0.255**	0.0694	0.474***
	(0.166)	(0.123)	(0.403)	(0.121)
k = -1 × Hospital	0	0	0	0
k = 0 × Regular Clinic	-0.0491	0.185***	-0.0260	0.245***
	(0.0771)	(0.0249)	(0.144)	(0.0260)
k = 1 × Regular Clinic	-0.0806	0.256***	-0.0863	0.351***
	(0.0954)	(0.0373)	(0.192)	(0.0365)
k = 2 × Regular Clinic	-0.00601	0.323***	-0.0555	0.432***
	(0.0969)	(0.0535)	(0.247)	(0.0492)
k = 3 × Regular Clinic	-0.0552	0.388***	-0.0795	0.523***
	(0.128)	(0.0684)	(0.303)	(0.0604)
k = 4 × Regular Clinic	0.0125	0.449***	0.0441	0.606***
	(0.160)	(0.0782)	(0.381)	(0.0619)
k = 5 × Regular Clinic	-0.141	0.509***	0.223	0.662***
	(0.181)	(0.0965)	(0.531)	(0.0785)
k ≥ 6 × Regular Clinic	-0.622***	0.523***	0.915	0.682***
	(0.193)	(0.125)	(0.924)	(0.112)
Observations	1,147	407	1,147	407
R-squared	0.798	0.988	0.792	0.985

Notes: Patient Deaths is the number of ART patients that died, measured at the clinic-year level. *Patients In Care* is the number of ART patients that visited the clinic at least once, measured at the clinic-year level. *New Patients* is the number of new ART patients that initiated ART for the first time. *Patient Retention* is the number of returning (did not initiate for the first time that year) ART patients that visited the clinic, measured at the clinic-year level. We add 1 to the outcome measure. $k = 0$ is the year immediately following EMR implementation. Each regression includes year and clinic fixed effects following (2). Heteroskedasticity-robust standard errors clustered at the clinic level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A15

Robustness: Heterogeneity by Clinic Size (Large vs. Small)

Event Year	(1) log(Patient Deaths)	(2) log(Patients In Care)	(3) log(New Patients)	(4) log(Retention)
k = -1 × Large Clinic	0	0	0	0
k = 0 × Large Clinic	-0.0143	0.176***	0.253	0.257***
	(0.0767)	(0.0221)	(0.201)	(0.0290)
k = 1 × Large Clinic	-0.0609	0.241***	0.218	0.343***
	(0.0879)	(0.0352)	(0.227)	(0.0420)
k = 2 × Large Clinic	-0.122	0.299***	0.221	0.429***
	(0.0937)	(0.0525)	(0.258)	(0.0568)
k = 3 × Large Clinic	-0.198	0.343***	0.162	0.503***
	(0.127)	(0.0674)	(0.290)	(0.0709)
k = 4 × Large Clinic	-0.192	0.383***	-0.00268	0.571***
	(0.137)	(0.0848)	(0.281)	(0.0854)
k = 5 × Large Clinic	-0.321**	0.422***	-0.0360	0.632***
	(0.139)	(0.104)	(0.334)	(0.103)
k ≥ 6 × Large Clinic	-0.609***	0.450***	0.286	0.684***
	(0.162)	(0.129)	(0.505)	(0.128)
k = -1 × Small Clinic	0	0	0	0
k = 0 × Small Clinic	-0.118	0.164***	0.0955	0.219***
	(0.0751)	(0.0237)	(0.214)	(0.0251)
k = 1 × Small Clinic	-0.216**	0.205***	0.186	0.290***
	(0.103)	(0.0368)	(0.308)	(0.0398)
k = 2 × Small Clinic	-0.226**	0.229***	0.432	0.331***
	(0.114)	(0.0518)	(0.422)	(0.0578)
k = 3 × Small Clinic	-0.305**	0.257***	0.488	0.374***
	(0.129)	(0.0676)	(0.528)	(0.0746)
k = 4 × Small Clinic	-0.292**	0.269***	0.400	0.404***
	(0.131)	(0.0741)	(0.617)	(0.0796)
k = 5 × Small Clinic	-0.450***	0.216**	0.786	0.376***
	(0.167)	(0.0866)	(1.010)	(0.0891)
k ≥ 6 × Small Clinic	-0.540**	0.208**	0.203	0.394***
	(0.233)	(0.0996)	(1.026)	(0.102)
Observations	1,166	411	1,166	411
R-squared	0.795	0.986	0.788	0.984

Notes: Patient Deaths is the number of ART patients that died, measured at the clinic-year level. *Patients In Care* is the number of ART patients that visited the clinic at least once, measured at the clinic-year level. *New Patients* is the number of new ART patients that initiated ART for the first time. *Patient Retention* is the number of returning (did not initiate for the first time that year) ART patients that visited the clinic, measured at the clinic-year level. We add 1 to the outcome measure. $k = 0$ is the year immediately following EMR implementation. Each regression includes year and clinic fixed effects following (2). Large clinics have above-median number of initiating patients as of 2013. Heteroskedasticity-robust standard errors clustered at the clinic level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A16

Robustness: Heterogeneity by Clinic Location (Urban vs. Rural)

Event Year	(1) log(Patient Deaths)	(2) log(Patients In Care)	(3) log(New Patients)	(4) log(Retention)
k = -1 × Urban Clinic	0	0	0	0
k = 0 × Urban Clinic	-0.0224	0.198***	0.623**	0.253***
	(0.0862)	(0.0241)	(0.263)	(0.0312)
k = 1 × Urban Clinic	-0.0490	0.262***	0.607**	0.352***
	(0.0994)	(0.0342)	(0.298)	(0.0428)
k = 2 × Urban Clinic	-0.148	0.297***	0.565*	0.411***
	(0.0989)	(0.0505)	(0.326)	(0.0564)
k = 3 × Urban Clinic	-0.234*	0.328***	0.365	0.469***
	(0.125)	(0.0644)	(0.322)	(0.0702)
k = 4 × Urban Clinic	-0.184	0.360***	0.0424	0.532***
	(0.132)	(0.0799)	(0.271)	(0.0831)
k = 5 × Urban Clinic	-0.315**	0.382***	-0.0413	0.578***
	(0.135)	(0.0994)	(0.331)	(0.100)
k ≥ 6 × Urban Clinic	-0.580***	0.401***	0.294	0.617***
	(0.159)	(0.122)	(0.497)	(0.125)
k = -1 × Rural Clinic	0	0	0	0
k = 0 × Rural Clinic	-0.108	0.136***	-0.262*	0.200***
	(0.0797)	(0.0231)	(0.154)	(0.0245)
k = 1 × Rural Clinic	-0.235**	0.159***	-0.306	0.245***
	(0.0947)	(0.0347)	(0.210)	(0.0373)
k = 2 × Rural Clinic	-0.187*	0.194***	-0.0652	0.299***
	(0.109)	(0.0511)	(0.352)	(0.0593)
k = 3 × Rural Clinic	-0.236*	0.224***	0.154	0.347***
	(0.126)	(0.0685)	(0.514)	(0.0796)
k = 4 × Rural Clinic	-0.332***	0.224***	0.539	0.349***
	(0.114)	(0.0751)	(0.804)	(0.0874)
k = 5 × Rural Clinic	-0.519**	0.186**	1.678	0.316***
	(0.209)	(0.0837)	(1.156)	(0.0978)
k ≥ 6 × Rural Clinic	-0.731***	0.173*	0.518	0.358***
	(0.258)	(0.0976)	(1.647)	(0.100)
Observations	1,147	407	1,147	407
R-squared	0.796	0.986	0.796	0.984

Notes: *Patient Deaths* is the number of ART patients that died, measured at the clinic-year level. *Patients In Care* is the number of ART patients that visited the clinic at least once, measured at the clinic-year level. *New Patients* is the number of new ART patients that initiated ART for the first time. *Patient Retention* is the number of returning (did not initiate for the first time that year) ART patients that visited the clinic, measured at the clinic-year level. We add 1 to the outcome measure. $k = 0$ is the year immediately following EMR implementation. Each regression includes year and clinic fixed effects following (2). Heteroskedasticity-robust standard errors clustered at the clinic level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A17

Dynamic Effect of EMR Adoption, Year FE Interacted with Clinic Size

	(1)	(2)	(3)	(4)
Event Year	log(Patient Deaths)	log(Patients In Care)	log(New Patients)	log(Retention)
k = -6	0.0382	N/A	-0.0879	N/A
	(0.0520)		(0.0718)	
k = -5	0.0268	N/A	-0.102	N/A
	(0.0557)		(0.0903)	
k = -4	0.0323	N/A	-0.0513	N/A
	(0.0656)		(0.0876)	
k = -3	0.00314	N/A	0.0594	N/A
	(0.0696)		(0.0804)	
k = -2	0.00581	N/A	0.0876	N/A
	(0.0678)		(0.0804)	
k = -1	0	0	0	0
k = 0	-0.0625	0.177***	0.0675	0.247***
	(0.0578)	(0.0181)	(0.114)	(0.0211)
k = 1	-0.123	0.240***	0.107	0.339***
	(0.0814)	(0.0320)	(0.165)	(0.0344)
k = 2	-0.170*	0.292***	0.232	0.414***
	(0.0936)	(0.0476)	(0.244)	(0.0487)
k = 3	-0.261**	0.334***	0.224	0.483***
	(0.111)	(0.0619)	(0.320)	(0.0621)
k = 4	-0.249**	0.368***	0.0673	0.544***
	(0.125)	(0.0777)	(0.406)	(0.0752)
k = 5	-0.379***	0.389***	0.103	0.589***
	(0.144)	(0.0942)	(0.531)	(0.0902)
k ≥ 6	-0.576***	0.412***	0.212	0.639***
	(0.164)	(0.118)	(0.676)	(0.113)
Year FE by Clinic Size	Y	Y	Y	Y
Observations	1,166	411	1,166	411
R-squared	0.799	0.986	0.792	0.984

Notes: Patient Deaths is the number of ART patients that died, measured at the clinic-year level. *Patients In Care* is the number of ART patients that visited the clinic at least once, measured at the clinic-year level. *New Patients* is the number of new ART patients that initiated ART for the first time. *Patient Retention* is the number of returning (did not initiate for the first time that year) ART patients that visited the clinic, measured at the clinic-year level. We add 1 to the outcome measure. $k = 0$ is the year immediately following EMR implementation. Each regression includes year and clinic fixed effects following (2). Year fixed effects are interacted with an indicator for large clinic. Large clinics have above-median number of initiating patients as of 2013. Heteroskedasticity-robust standard errors clustered at the clinic level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A18

Dynamic Effect of EMR Adoption, Excluding Late Adopters

Event Year	(1) log(Patient Deaths)	(2) log(Patients In Care)	(3) log(New Patients)	(4) log(Retention)
k = -6	0.0297	N/A	0.0222	N/A
	(0.0873)		(0.0966)	
k = -5	0.0731	N/A	0.0661	N/A
	(0.0785)		(0.0775)	
k = -4	0.0618	N/A	-0.0270	N/A
	(0.0690)		(0.0849)	
k = -3	0.0886	N/A	-0.0976	N/A
	(0.0579)		(0.0902)	
k = -2	0.0634	N/A	-0.106	N/A
	(0.0559)		(0.0769)	
k = -1	0	0	0	0
k = 0	-0.0355	0.160***	0.0813	0.226***
	(0.0650)	(0.0179)	(0.111)	(0.0207)
k = 1	-0.104	0.206***	0.0924	0.297***
	(0.0905)	(0.0287)	(0.150)	(0.0328)
k = 2	-0.143	0.241***	0.202	0.354***
	(0.109)	(0.0423)	(0.202)	(0.0462)
k = 3	-0.217	0.272***	0.155	0.410***
	(0.135)	(0.0554)	(0.248)	(0.0598)
k = 4	-0.207	0.300***	-0.0366	0.464***
	(0.152)	(0.0690)	(0.322)	(0.0715)
k = 5	-0.348**	0.316***	-0.0191	0.504***
	(0.166)	(0.0874)	(0.408)	(0.0879)
k ≥ 6	-0.592***	0.335***	0.0937	0.548***
	(0.195)	(0.110)	(0.546)	(0.111)
F-test on pre-trends	0.571	N/A	0.976	N/A
p-value of F-test	0.722	N/A	0.438	N/A
Observations	943	411	943	411
R-squared	0.766	0.985	0.779	0.983

Notes: Patient Deaths is the number of ART patients that died, measured at the clinic-year level. *Patients In Care* is the number of ART patients that visited the clinic at least once, measured at the clinic-year level. *New Patients* is the number of new ART patients that initiated ART for the first time. *Patient Retention* is the number of returning (did not initiate for the first time that year) ART patients that visited the clinic, measured at the clinic-year level. We add 1 to the outcome measure. $k = 0$ is the year immediately following EMR implementation. Each regression includes year and clinic fixed effects following (2). Clinics adopting EMRs after 2017 are excluded from the analysis. Heteroskedasticity-robust standard errors clustered at the clinic level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A19

Dynamic Effect of EMR Adoption, Inverse Hyperbolic Sine Transformation

Event Year	(1) arcsinh(Patient Deaths)	(2) arcsinh(Patients In Care)	(3) arcsinh(New Patients)	(4) arcsinh(Retention)
k = -6	-0.00587	N/A	0.0488	N/A
	(0.0764)		(0.0755)	
k = -5	-0.00311	N/A	0.0413	N/A
	(0.0742)		(0.0609)	
k = -4	0.0176	N/A	-0.0821	N/A
	(0.0629)		(0.0736)	
k = -3	-0.00426	N/A	-0.106	N/A
	(0.0572)		(0.0802)	
k = -2	0.0268	N/A	-0.0650	N/A
	(0.0557)		(0.0713)	
k = -1	0	0	0	0
k = 0	-0.0727	0.160***	0.118	0.226***
	(0.0679)	(0.0180)	(0.125)	(0.0208)
k = 1	-0.143	0.206***	0.161	0.297***
	(0.0909)	(0.0288)	(0.163)	(0.0329)
k = 2	-0.178*	0.241***	0.274	0.355***
	(0.105)	(0.0423)	(0.216)	(0.0462)
k = 3	-0.255*	0.272***	0.240	0.410***
	(0.129)	(0.0554)	(0.258)	(0.0599)
k = 4	-0.238*	0.300***	0.0366	0.465***
	(0.140)	(0.0690)	(0.300)	(0.0715)
k = 5	-0.371**	0.316***	0.0597	0.505***
	(0.152)	(0.0874)	(0.387)	(0.0880)
k ≥ 6	-0.635***	0.335***	0.192	0.549***
	(0.173)	(0.110)	(0.499)	(0.111)
F-test on pre-trends	0.120	N/A	1.207	N/A
p-value of F-test	0.988	N/A	0.311	N/A
Observations	1,166	411	1,166	411
R-squared	0.780	0.985	0.783	0.983

Notes: *Patient Deaths* is the number of ART patients that died, measured at the clinic-year level. *Patients In Care* is the number of ART patients that visited the clinic at least once, measured at the clinic-year level. *New Patients* is the number of new ART patients that initiated ART for the first time. *Patient Retention* is the number of returning (did not initiate for the first time that year) ART patients that visited the clinic, measured at the clinic-year level. $k = 0$ is the year immediately following EMR implementation. Each regression includes year and clinic fixed effects following (2). Heteroskedasticity-robust standard errors clustered at the clinic level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A20

Dynamic Effect of EMR Adoption, Alternative Log Transformation

Event Year	(1) altlog(Patient Deaths)	(2) altlog(Patients In Care)	(3) altlog(New Patients)	(4) altlog(Retention)
k = -6	0.0244	N/A	0.0783	N/A
	(0.111)		(0.110)	
k = -5	-0.0356	N/A	0.0406	N/A
	(0.119)		(0.102)	
k = -4	0.0496	N/A	-0.235*	N/A
	(0.0733)		(0.135)	
k = -3	-0.0226	N/A	-0.228*	N/A
	(0.0784)		(0.135)	
k = -2	0.0198	N/A	-0.182	N/A
	(0.0758)		(0.131)	
k = -1	0	0	0	0
k = 0	-0.126	0.160***	0.1000	0.226***
	(0.102)	(0.0180)	(0.148)	(0.0208)
k = 1	-0.213	0.206***	0.158	0.297***
	(0.136)	(0.0288)	(0.199)	(0.0329)
k = 2	-0.280*	0.241***	0.268	0.355***
	(0.159)	(0.0423)	(0.263)	(0.0462)
k = 3	-0.374**	0.272***	0.189	0.410***
	(0.184)	(0.0554)	(0.315)	(0.0599)
k = 4	-0.375*	0.300***	-0.134	0.465***
	(0.205)	(0.0690)	(0.383)	(0.0715)
k = 5	-0.521**	0.316***	-0.112	0.505***
	(0.226)	(0.0874)	(0.498)	(0.0880)
k ≥ 6	-0.820***	0.335***	0.0248	0.549***
	(0.263)	(0.110)	(0.617)	(0.111)
F-test on pre-trends	0.317	N/A	1.436	N/A
p-value of F-test	0.902	N/A	0.217	N/A
Observations	1,166	411	1,166	411
R-squared	0.673	0.985	0.749	0.983

Notes: Patient Deaths is the number of ART patients that died, measured at the clinic-year level. *Patients In Care* is the number of ART patients that visited the clinic at least once, measured at the clinic-year level. *New Patients* is the number of new ART patients that initiated ART for the first time. *Patient Retention* is the number of returning (did not initiate for the first time that year) ART patients that visited the clinic, measured at the clinic-year level. We transform outcomes using "altlog" defined as $\log(y + 0.01)$. $k = 0$ is the year immediately following EMR implementation. Each regression includes year and clinic fixed effects following (2). Heteroskedasticity-robust standard errors clustered at the clinic level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A21

Dynamic Effect of EMR Adoption, Denominator Transformation

Event Year	(1) denom(Patient Deaths)	(2) denom(Patients In Care)	(3) denom(New Patients)	(4) denom(Retention)
k = -6	-0.000574	N/A	-0.00191	N/A
	(0.00126)		(0.00867)	
k = -5	-0.000287	N/A	0.000707	N/A
	(0.000955)		(0.00702)	
k = -4	-0.000225	N/A	0.00234	N/A
	(0.000902)		(0.00644)	
k = -3	0.000501	N/A	0.00257	N/A
	(0.000749)		(0.00603)	
k = -2	0.000462	N/A	4.74e-05	N/A
	(0.000692)		(0.00416)	
k = -1	0	0	0	0
k = 0	0.000159	0.106***	-0.00571	0.116***
	(0.000703)	(0.0115)	(0.00693)	(0.00923)
k = 1	-0.000708	0.130***	-0.00864	0.148***
	(0.000928)	(0.0188)	(0.00817)	(0.0146)
k = 2	-0.00124	0.150***	-0.0107	0.175***
	(0.00114)	(0.0287)	(0.0104)	(0.0219)
k = 3	-0.00195	0.170***	-0.0152	0.205***
	(0.00138)	(0.0383)	(0.0132)	(0.0296)
k = 4	-0.00121	0.189***	-0.0193	0.235***
	(0.00153)	(0.0475)	(0.0157)	(0.0347)
k = 5	-0.00279	0.197***	-0.0284	0.256***
	(0.00180)	(0.0608)	(0.0195)	(0.0432)
k ≥ 6	-0.00453**	0.206***	-0.0300	0.275***
	(0.00216)	(0.0776)	(0.0221)	(0.0562)
F-test on pre-trends	0.367	N/A	0.137	N/A
p-value of F-test	0.870	N/A	0.983	N/A
Observations	943	411	943	411
R-squared	0.604	0.884	0.503	0.917

Notes: *Patient Deaths* is the number of ART patients that died, measured at the clinic-year level. *Patients In Care* is the number of ART patients that visited the clinic at least once, measured at the clinic-year level. *New Patients* is the number of new ART patients that initiated ART for the first time. *Patient Retention* is the number of returning (did not initiate for the first time that year) ART patients that visited the clinic, measured at the clinic-year level. We consider the ratio of each outcome by the number of patients in care as of 2017, and we only include clinics that have adopted EMRs by 2017. $k = 0$ is the year immediately following EMR implementation. Each regression includes year and clinic fixed effects following (2). Heteroskedasticity-robust standard errors clustered at the clinic level. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A22

Pre-trends F-Test Using Borusyak (2021) Estimator

Dependent Variables		Borusyak, Jaravel & Spiess	
		F-test on pre-trends	p-value of F-test
log(Patient Deaths)	All	0.8	0.475
	Male	2.1	0.107
	Female	0.3	0.798
	Age ≤ 9	1.2	0.298
	Age 10 to 17	1.1	0.358
	Age 18 to 49	0.5	0.705
	Age ≥ 50	1.8	0.152
log(New Patients)	All	1.0	0.398
	Male	1.0	0.380
	Female	0.7	0.550
	Age ≤ 9	0.0	0.992
	Age 10 to 17	0.3	0.831
	Age 18 to 49	0.8	0.495
	Age ≥ 50	0.4	0.763

Notes: Alternative test for existence of pre-trends using Borusyak et al. (2021) estimator.

Appendix Table A23

Average Appointment Intervals and Clinic-Level Outcomes

	(1)	(2)	(3)	(4)
	Annual Tests per Patient		Share Lapsed 2 Years After Initiation	
Average Interval in Months	0.0487	-0.145	-0.116***	-0.110***
	(0.114)	(0.118)	(0.0135)	(0.0145)
Controls	N	Y	N	Y
Observations	313	313	313	313
R-squared	0.110	0.211	0.620	0.638

Notes: Clinic-year level regression. *Tests* includes TB testing and treatment, CD4 count orders, viral load orders, and other referrals for services. *Lapsed* indicates share of patients with no visit in the 6 month period before their 2 year anniversary of start of care. Controls include share of male patients, patients in each age group, new patients with advanced HIV, and new patients underweight. Robust standard errors. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A24

Patient-Level Predictors of Appointment Intervals

	(1)	(2)
	Interval in Months	
	At Initiation	Later Visits
Age 0 to 9	-0.0562***	-0.590***
	(0.00572)	(0.00276)
Age 10 to 17	0.0104	-0.551***
	(0.00643)	(0.00272)
Age 18 to 49	-0.0487***	-0.116***
	(0.00375)	(0.00154)
Male	-0.0222***	0.000394
	(0.00218)	(0.00102)
Advanced HIV at Initiation	-0.0951***	0.0138***
	(0.00254)	(0.00108)
Underweight at Initiation	-0.0865***	-0.0605***
	(0.00309)	(0.00147)
Missed Last Appointment		-0.157***
		(0.00273)
Years Since Initiation		0.208***
		(0.000308)
Year FE	Y	Y
Clinic FE	Y	Y
Observations	328,692	3,447,753
R-squared	0.092	0.231

Notes: Patient-visit level regression. Robust standard errors. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A25

Appointment Intervals and Patient Outcomes

	(1)	(2)	(3)	(4)
	At Initiation		One Year After Initiation	
	Missed Next Appointment	Lapsed 1 Year After Initiation	Missed Next Appointment	Lapsed 2 Years After Initiation
2 Weeks	-0.105***	-0.0422***	0.0144	0.0130
	(0.00322)	(0.00420)	(0.0111)	(0.0167)
1 Month	0.00164	0.0300***	-0.0110***	0.0716***
	(0.00249)	(0.00338)	(0.00360)	(0.00569)
2 Months	-0.0397***	-0.000284	-0.0343***	0.0183***
	(0.00317)	(0.00432)	(0.00329)	(0.00515)
3 Months			-0.0242***	-0.0198***
			(0.00327)	(0.00509)
4 Months			-0.0181***	-0.0312***
			(0.00334)	(0.00526)
Year FE	Y	Y	Y	Y
Controls	Y	Y	Y	Y
Observations	333,281	281,993	205,433	157,424
R-squared	0.017	0.016	0.037	0.026

Notes: Patient-visit level regression. *Lapsed* indicates share of patients with no visit in the 6 month period before their 1 or 2 year anniversary of start of care. *Interval* is the dispensing interval. Controls include gender, age group, advanced HIV at initiation, underweight at initiation, indicator for missed last appointment. Robust standard errors. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Appendix Table A26

Disability-Adjusted Life Years (DALYs) and Cost-Effectiveness Analysis

DALYs Avoided Due to EMR per Individual (in years of life)

		Health State Weight		
		0.034	*0.053*	*0.079*
Average Effect Size of EMR on Additional Years of Life (in years)	*0.5*	0.48	0.47	0.46
	1.0	0.97	0.95	0.92
	1.412	**1.36**	**1.34**	**1.30**
	2.0	1.93	1.89	1.84
	3.0	2.90	2.84	2.76
	4.0	3.86	3.79	3.68
	5.0	4.83	4.74	4.61
	6.0	5.80	5.68	5.53
	7.0	6.76	6.63	6.45
	8.0	7.73	7.58	7.37

EMR Cost-Effectiveness (in U$ per DALY)

		Health State Weight		
		0.034	*0.053*	*0.079*
Average Effect Size of EMR on Additional Years of Life (in years)	*0.5*	$ 20.41	$ 20.82	$ 21.40
	1.0	$ 10.20	$ 10.41	$ 10.70
	1.412	**$ 7.23**	**$ 7.37**	**$ 7.58**
	2.0	$ 5.10	$ 5.20	$ 5.35
	3.0	$ 3.40	$ 3.47	$ 3.57
	4.0	$ 2.55	$ 2.60	$ 2.68
	5.0	$ 2.04	$ 2.08	$ 2.14
	6.0	$ 1.70	$ 1.73	$ 1.78
	7.0	$ 1.46	$ 1.49	$ 1.53
	8.0	$ 1.28	$ 1.30	$ 1.34

Notes: Disability-Adjusted Life Years (DALYs) calculated considering additional 1.412 years of life after. EMR implementation per patient, and disability-adjustment of Health State Weight of 0.053 (0.034-0.079) for patients under "HIV/AIDS: receiving antiretroviral treatment" according to Salomon et al. (2012) for these additional years of life.

Appendix References

Abraham, S. and Sun, L. (2020). Estimating Dynamic Treatment Effects in Event Studies With Heterogeneous Treatment Effects. *Journal of Econometrics*.

Borusyak, K., Jaravel, X., and Spiess, J. (2021). Revisiting Event Study Designs: Robust and Efficient Estimation. *Work in Progress*, pages 1–48.

Callaway, B. and Sant'Anna, P. H. (2020). Difference-in-Differences with Multiple Time Periods. *Journal of Econometrics*.

de Chaisemartin, C. and D'Haultfœuille, X. (2020). Difference-in-differences estimators of intertemporal treatment effects. *Available at SSRN 3731856*.

de Chaisemartin, C. and D'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–2996.

Driessen, J., Cioffi, M., Alide, N., Landis-Lewis, Z., Gamadzi, G., Gadabu, O. J., and Douglas, G. (2013). Modeling return on investment for an electronic medical record system in Lilongwe, Malawi. *Journal of the American Medical Informatics Association*, 20(4):743–748.

Salomon, J. (2013). Health benefits, costs, and cost-effectiveness of earlier eligibility for adult antiretroviral therapy and expanded treatment coverage: a combined analysis of 12 mathematical models.

Salomon, J. A., Vos, T., Hogan, D. R., Gagnon, M., Naghavi, M., Mokdad, A., Begum, N., Shah, R., Karyana, M., Kosen, S., Farje, M. R., Moncada, G., Dutta, A., Sazawal, S., Dyer, A., Seiler, J., Aboyans, V., Baker, L., Baxter, A., Benjamin, E. J., Bhalla, K., Abdulhak, A. B., Blyth, F., Bourne, R., Braithwaite, T., Brooks, P., Brugha, T. S., Bryan-Hancock, C., Buchbinder, R., Burney, P., Calabria, B., Chen, H., Chugh, S. S., Cooley, R., Criqui, M. H., Cross, M., Dabhadkar, K. C., Dahodwala, N., Davis, A., Degenhardt, L., Díaz-Torné, C., Dorsey, E. R., Driscoll, T., Edmond, K., Elbaz, A., Ezzati, M., Feigin, V., Ferri, C. P., Flaxman, A. D., Flood, L., Fransen, M., Fuse, K., Gabbe, B. J., Gillum, R. F., Haagsma, J., Harrison, J. E., Havmoeller, R., Hay, R. J., Hel-Baqui, A., Hoek, H. W., Hoffman, H., Hogeland, E., Hoy, D., Jarvis, D., Karthikeyan, G., Knowlton, L. M., Lathlean, T., Leasher, J. L., Lim, S. S., Lipshultz, S. E., Lopez, A. D., Lozano, R., Lyons, R., Malekzadeh, R., Marcenes, W., March, L., Margolis, D. J., McGill, N., McGrath, J., Mensah, G. A., Meyer, A. C., Michaud, C., Moran, A., Mori, R., Murdoch, M. E., Naldi, L., Newton, C. R., Norman, R., Omer, S. B., Osborne, R., Pearce, N., Perez-Ruiz, F., Perico, N., Pesudovs, K., Phillips, D., Pourmalek, F., Prince, M., Rehm, J. T., Remuzzi, G., Richardson, K., Room, R., Saha, S., Sampson, U., Sanchez-Riera, L., Segui-Gomez, M., Shahraz, S., Shibuya, K., Singh, D., Sliwa, K., Smith, E., Soerjomataram, I., Steiner, T., Stolk, W. A., Stovner, L. J., Sudfeld, C., Taylor, H. R., Tleyjeh, I. M., Van Der Werf, M. J., Watson, W. L., Weatherall, D. J., Weintraub, R., Weisskopf, M. G., Whiteford, H., Wilkinson, J. D., Woolf, A. D., Zheng, Z. J., and Murray, C. J. (2012). Common values in assessing health outcomes from disease and injury: Disability weights measurement study for the Global Burden of Disease Study 2010. *The Lancet*, 380(9859):2129–2143.

Tanser, F., Bärnighausen, T., Grapsa, E., Zaidi, J., and Newell, M.-L. (2013). High coverage of ART associated with decline in risk of HIV acquisition in rural KwaZulu-Natal, South Africa. *Science*, 339(6122):966–971.

UNAIDS (2022). UNAIDS data 2022. *Programme on HIV/AIDS*.