// NO.25-034 | 07/2025

DISCUSSION PAPER

// REBECCA JANSSEN, SIMON REIF, AND SABRINA SCHUBERT

Market Entry of Digital Health Providers after the Introduction of a New Reimbursement Pathway

Leibniz Association

ZEW

Market Entry of Digital Health Providers after the Introduction of a New Reimbursement Pathway ♥

Rebecca Janßen
ZEW Mannheim

Simon Reif
ZEW Mannheim
FAU Erlangen-Nürnberg

Sabrina Schubert★
ZEW Mannheim
FAU Erlangen-Nürnberg

Version 1.0 #f30d2f4b
30 June 2025

Abstract

Digital therapeutics are increasingly used to complement traditional health care. In a pioneering move, Germany became the first country to introduce a structured regulatory framework — known as the DiGA scheme — that enables developers of digital therapeutics to be reimbursed in the statutory health insurance system. Our study evaluates the impact of this novel regulation on the development and market entry of patient-centered digital health applications. Using a panel dataset of app availability by language and month from the Apple App Store, covering the period from January 2018 to September 2021, we compare trends in health app availability in German to those in other languages. Applying event study designs and a set of synthetic control methods, we find that the DiGA regulation likely stimulated the development of German-language digital therapeutics in the app market. While the number of apps increased, our results suggest that neither the diversity of health conditions targeted nor the number of high-quality apps expanded significantly. To the contrary, the increase was almost exclusively driven by apps that sell patient data for advertisement. This suggests that the initial enthusiasm surrounding the new reimbursement pathway did not translate into a broad increase in high quality apps with strong data privacy protections. Further research is needed to assess the longer-term effects on innovation and quality, especially as other countries begin to adopt regulatory frameworks inspired by the German model.

Keywords: Digital Health, DiGA, Reimbursement, Digital Therapeutics
JEL codes: I11, I18, L52

♥ We would like to thank Stefan Boes, Markus Nagler, Daniel Polsky, Johannes Rincke, Ariel Stern, Harald Tauchmann and Achim Wambach as well as the participants of the Economics Seminar at the University of Gießen, the dggö Conference 2024 in Halle, the Workshop on Applied Economics in Digital Health in Potsdam 2024, the EuHEA Conference 2024 in Vienna, the EuHEA PhD Conference in Lucerne 2024, the DEASS Brown Bag Seminar at SUPSI in Lugano, and the Johns Hopkins Economic Brownbag seminar for valuable comments and suggestions. Benedikt Stelter and Kristina Kellersmann provided excellent research assistance. Funding from the Forum Gesundheitsstandort Baden-Württemberg (ID: 43-5400//141/1) and the Schöller Foundation is gratefully acknowledged.
★Corresponding author: ZEW - Centre for European Economic Research, L7 1, 68161 Mannheim, Germany. Email: sabrina.schubert@zew.de.

1 Introduction

In recent years, the digitization of healthcare has accelerated and brought a wave of innovations which aim to improve care quality and make healthcare systems more efficient. Among the most prominent examples is the growing availability of smartphone applications — often referred to as digital therapeutics — that support patients in their efforts to prevent, manage or cure health conditions. While there are numerous studies that show potential positive health effects of introducing digital therapeutics into healthcare systems (Han & Lee, 2018; Iribarren et al., 2021; Lu et al., 2018; Mira et al., 2014), it is less clear how health systems can systematically incentivize firms to develop digital therapeutics that provide medical benefits to users. The usual business models of app developers are based either on advertisement, selling data, reimbursement contracts with specific health insurers and healthcare providers, or patients who pay out of pocket. Integrating such digital therapeutics into clinical practice, allowing for reimbursement and prescription by physicians have therefore been identified as crucial factors for scaling their use (Gordon et al., 2020; Sim, 2019). In 2019, Germany was the first country worldwide to introduce such a structured pathway for reimbursement of digital therapeutics in a national health system: the DiGA scheme (Gerke et al., 2020).[1] The DiGA scheme defines clear rules for the quality of digital therapeutics, including the demonstration of effectiveness, and therein defines a pathway into reimbursement similar to the pharmaceutical or medical device sector. Once the process has been successfully completed, the digital therapeutic can be prescribed by physicians and psychotherapists to patients and the costs for the developers are reimbursed by the statutory health insurance (SHI). This new coverage by the SHI opened a market with more than 74 million individuals for developers. In this paper, we analyze the effect of the DiGA reimbursement scheme on market entry of health apps in general. We use the universe of available apps in the Apple App Store from 2018 to 2021 and exploit the exogenous introduction of the new reimbursement scheme for German digital therapeutics to estimate the effect of structured reimbursement on the number of available health apps and their characteristics.

Previous literature on digital therapeutics spans several disciplines, from medicine and public health to computer science and business studies as well as economics. First and foremost, RCTs from the medical field show that using apps in the treatment of several conditions can be beneficial, e.g., in the context of diabetes or heart failure (Bonoto et al., 2017; Cajita et al., 2016; Chong et al., 2023). Several studies also investigate the efficiency of such applications and their value for users' health (Ghose et al., 2022).

[1]DiGA stands for the German term *Digitale Gesundheits Anwendungen* meaning *Digital Health Applications*

From a developer perspective, accessibility of apps, improved knowledge design, and scientific signaling (Larsen et al., 2019) are discussed as key elements for the success of digital therapeutics. In addition, questions on how to regulate and approve new digital health services and ethical considerations such as equitable access are discussed across disciplines (Sim, 2019).

With respect to the related economics literature, there is a growing field of research on app markets focusing on regulation, developers' decisions, and innovation. Janßen et al. (2022) show that regulation affects market exit and entry decisions of developers by studying unintended consequences of General Data Protection Regulation (GDPR) in terms of impeded innovation. They find that one-third of apps exited the market after GDPR came into effect and a decrease in app entry by half. More recent studies have looked at changes in app developers' monetization strategies as well as market entry overall following the implementation of Apple's App Tracking Transparency (ATT) regulation that requires explicit consent from users for tracking outside of the app (Cheyre et al., 2023; Kesler, 2022; Kraft et al., 2023). Their overall findings are that Apple's ATT led to changes in the monetization strategy with a higher share of apps with upfront prices or in-app purchases, reduced app entry, and fewer updates. While this strand of literature shows that new regulations or shocks to the app market can negatively impact the entry and innovation behavior of developers, our paper investigates a positive shock in the form of a new reimbursement pathway for medical apps.

With respect to monetization, app developers across all genres can usually choose between different strategies, from free app provision to paid and freemium models with in-app purchases. Lately, freemium models have received increased attention in the literature particularly in relation to potential revenue optimization (Deng et al., 2022; Roma & Ragaglia, 2016). Kummer and Schulte (2019) relate data collection to monetization strategies of apps and find that free apps tend to collect more privacy-intrusive information about users than paid apps do, supporting the claim of "data is the new currency", as user information is used for lucrative targeted advertisement. The amount of data collection is a crucial issue particularly in the context of apps that collect sensitive health data. Kesler et al. (2019) look at competition and privacy choices of app developers and find a positive correlation between the amount of data collected and market power. In our paper, we analyze a unique setting where there is a clear monetization strategy (reimbursement by SHI) but also very strict data privacy requirements.

Another relevant strand of the literature examines the relationship between market size, insurance coverage, and research and development (R&D) in pharmaceutical markets. A large body of both theoretical and empirical work finds that increases in market size are important drivers of pharmaceutical innovation (Acemoglu & Linn, 2004;

Agarwal & Gaule, 2022; Dubois et al., 2015; Finkelstein, 2004). One frequently discussed driver of such an increase is insurance coverage. Frankovic and Kuhn (2023) use a model of the US healthcare system to quantify this effect and find that expansion of insurance coverage between 1965 and 2005 increased R&D spending growth by 57%. A prominent example of market size expansion from the United States is the introduction of Medicare Part D in 2006, which expanded prescription drug coverage for seniors. Higher coverage increased consumption, resulting in a larger market size (Duggan et al., 2008; Ketcham & Simon, 2008). This change increased R&D investment in drugs targeted at the elderly, reflected in a rise in clinical trials and Food and Drug Administration (FDA) approvals for such products (Blume-Kohout & Sood, 2013; Duggan & Morton, 2010). Similar relations between coverage, market size, and innovation can be expected for digital therapeutics; however, such effects have not yet been studied in the digital context.

Lastly, our work is related to the literature on how public funding increases innovation in a sector. A prominent example is the German feed-in tariff scheme for renewable energy over the last two decades. These guaranteed prices led to policy-induced innovation, e.g., more patents by inventors with residence in Germany (Böhringer et al., 2017; Lindman & Söderholm, 2016). In a broader sense, this can be attributed to public subsidies and their key role in supporting emerging industries (Sun et al., 2019). We contribute to this literature as the first to analyze the effects on market size and market entry decisions of digital therapeutics following the German DiGA regulation. As we are analyzing the first reimbursement scheme of its kind worldwide, our results can have implications for the design of future schemes in other countries.

Our results show that the number of available German digital therapeutics increased substantially after the introduction of the DiGA scheme. This trend is found in event study as well as synthetic control approaches. However, the statistical significance varies across methods. Additionally, we find that the number of diagnoses targeted with health apps did not increase, that the increase stems mostly from apps which are not based on scientific publications and use a monetization scheme that includes data collection for advertisement.

Our paper proceeds as follows: In section 2, we describe the institutional background, in particular the DiGA reimbursement path, in more detail. Section 3 presents our empirical analysis and results with additional robustness analyses, and heterogeneity analyses. Section 4 concludes.

2 Institutional Background

The number of health- and medical-related apps has grown in the last decade to 337,522 available digital therapeutics on the market in 2024 (Aitken & Nass, 2024). The majority of health-related apps are general fitness and wellness apps aiming for overall lifestyle improvements. In addition, there is also a growing market for digital therapeutics which focus on treatment, prevention, and management of specific clinical indications and diseases. Globally, countries follow different strategies on whether and how to include digital therapeutics into formal reimbursement structures (Tarricone et al., 2022). For market access, digital therapeutics require medical device certification, for example according to the European Medical Device Regulation.[2] Certification of "software as a medical device" has already been established for more than three decades (Forsström & Rigby, 1999; Murfitt, 1990). Conventionally, digital therapeutics are reimbursed based on contracts between individual insurances and developers (Van Den Berg et al., 2016). Germany has been the first country to introduce a structured reimbursement pathway for "apps on prescription" in a national healthcare system.

The DiGA reimbursement scheme was introduced in the Digital Healthcare Act (*Digitale-Versorgung-Gesetz*, DVG) in 2019 and allows physicians and psychotherapists to prescribe approved health applications (DiGAs) to their patients while the costs are fully covered by the statutory health insurance (Federal Institute for Drugs and Medical Devices (BfArM), 2020). Before a digital therapeutic can apply for approval, it must fulfill three basic conditions: It must be listed in an app store or be otherwise publicly available. It must be a certified low-risk medical device (classes I or IIa in the European Medical Device Regulation) and the main functionality must be digital. According to the official DiGA guide a DiGA "*supports the recognition, monitoring, treatment or alleviation of diseases or the recognition, treatment or alleviation or compensation of injuries or disabilities*" and should be used by either the patient or the patient and their healthcare provider (Federal Institute for Drugs and Medical Devices (BfArM), 2020, p. 12).

Developers must submit an application that demonstrates that the requirements for data security, quality, functionality and medical effect - in the sense of direct medical benefit and patient-relevant process improvements - are met. There are two approval pathways: First, given sufficient evidence on the clinical benefits, developers can apply for permanent approval. Second, the regulation allows for temporary approval with reduced initial requirements on the evidence for medical effectiveness. In the case of temporary approval, developers have 12 to 24 months to subsequently file additional evidence based on clinical studies. This temporary listing allows developers to conduct

[2] See https://www.ema.europa.eu/en/human-regulatory-overview/medical-devices for more details on medical devices in the EU.

large-scale trials on the effectiveness of their digital therapeutic while already generating revenue from prescriptions. From an economic perspective, this early reimbursement creates additional incentives to enter the market since the immediate entry costs are reduced. The regulatory agency (BfArM) is supposed to make a decision on accepting the digital therapeutic as a DiGA within three months after submission of the application documents in both cases (Federal Institute for Drugs and Medical Devices (BfArM), 2020).

After approval, price setting for DiGAs is similar to the price setting process for novel pharmaceuticals in Germany with free price setting within the first year (Gensorowsky et al., 2022). Each prescription is reimbursed by the statutory health insurance whereby one prescription usually lasts for a three-month period per user.[3] Considering the low marginal costs of providing an app to a patient, prices set by the developers for the first year are high: the average price per prescription was 406 Euros for the first 48 DiGAs (Brönneke et al., 2023). After the initial year, prices are negotiated between the developers and the association of statutory health insurers. In 2021, rules for these price negotiations were set between the association of digital health companies and the statutory health insurers in an arbitration process (Techniker Krankenkasse, 2022). These rules include disease-specific maximum prices. However, the majority of prices for permanently listed DiGAs are larger than 200 Euros per prescription (Schmidt et al., 2024). The combination of low development costs for digital therapeutics (e.g. expected cost of 500,000 Euros according to Khan et al. (2024)), marginal cost of zero per prescription and relatively high reimbursement rates in a scheme for 74 million insured individuals is likely to incentivize market entry by lowering initial financial barriers and enabling early-stage revenue generation.

The introduction of the new regulation followed the timeline illustrated in Figure 1. In May 2019, the first general draft for the scheme was made public, followed by approval of the scheme in the German Cabinet in July 2019. Then the legislation and implementation followed in November and December 2019. The regulator published a detailed guide for developers in April 2020 and the first available digital therapeutics via the DiGA scheme became available in October 2020.

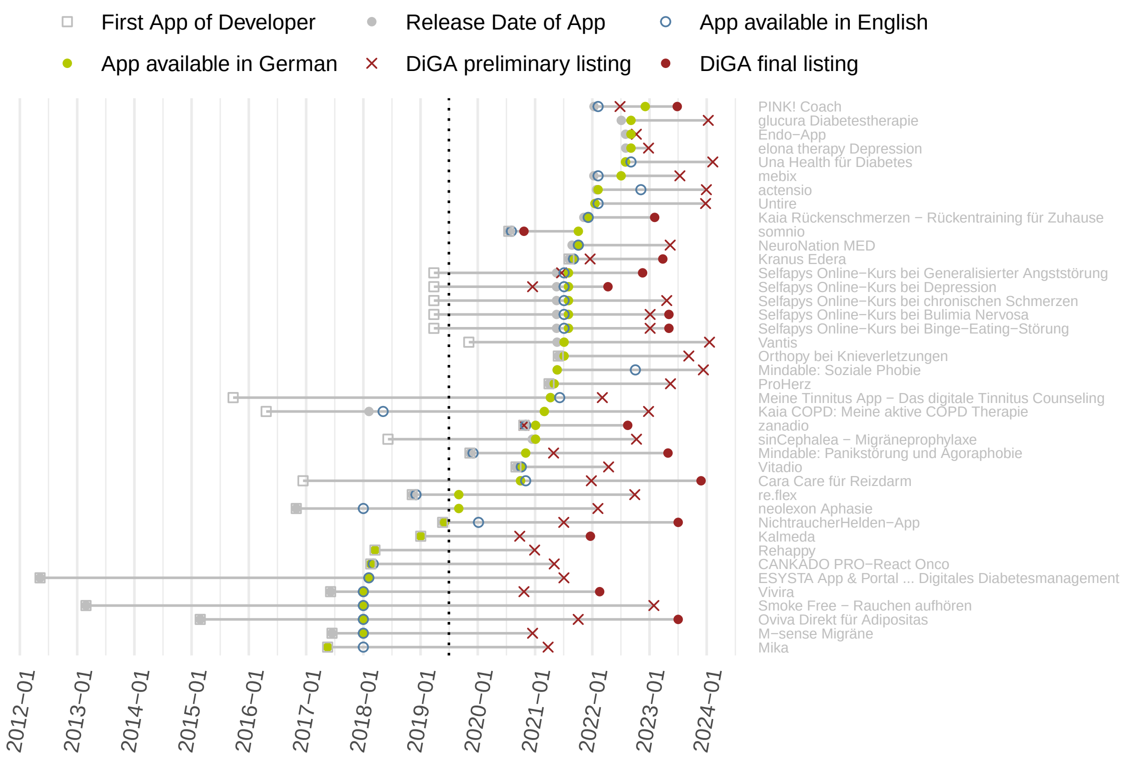
[3]This three-month time-frame is a quirk inherited from the patient-quarter-based reimbursement for ambulatory care physicians.

Figure 1: Timeline of the DiGA Legislation

As of May 2025, there have been 231 applications for inclusion in the DiGA scheme. Out of all applications, 61 percent have either been rejected (n=23) or withdrawn by the applicants (n=118). Currently, 21 applications are under review, and ten apps that had previously been approved were removed after their one-year temporary reimbursement period due to insufficient evidence (Federal Institute for Drugs and Medical Devices (BfArM), 2025a). To date, 58 applications are listed in the DiGA directory, 44 of which are available as smartphone apps in the Apple App Store and Google Play Store (Federal Institute for Drugs and Medical Devices (BfArM), 2025b).

In Figure 2, we illustrate the key milestones for approved DiGA apps as of March 2024 which were available in the Apple App Store in December 2023. While the majority of applications that were subsequently approved as DiGAs entered the App Store after July 2019, there are a number of apps that were already available before. Based on their early availability in German, it appears that many apps were initially intended for the German market.

Figure 2: Milestone Dates for DiGA Apps



Notes: Milestone dates for DiGA apps listed in the directory as of March 2024 and available in the Apple App Store as of December 2023. The black dotted line indicates July 2019, when the German government passed the law that would lead to the DiGA regulation. The Selfappy app contains separate modules for different diagnoses that underwent separate DiGA processes.

In other countries, similar reimbursement pathways for digital therapeutics were adopted later (MedTech Europe, 2021). In 2023, Belgium expanded its process for structured reimbursement mHealth and France introduced a fast-track pathway for reimbursing digital therapeutics PECAN (European Federation of Pharmaceutical Industries and Associations, 2023; Tarricone et al., 2024).

3 Empirical Analysis

We aim to identify whether the large financial incentives in the newly introduced DiGA reimbursement pathway increased the number of digital therapeutics developed for the German market. We measure availability of digital therapeutics in the form of apps since the majority of digital therapeutics are made available to patients via apps. Our hypothesis is that the number of apps developed for the German market increased after the DiGA regulation, compared to markets without structured reimbursement. We test our hypothesis using Apple App Store data and a variety of estimation strategies from classical event-study analysis to a synthetic control approach (Abadie, 2021; Abadie et al., 2010) and the recent synthetic difference-in-differences approach by Arkhangelsky et al. (2021).

3.1 Data and Sample Construction

For this project, we use data on all apps in the Apple App Store purchased from MightySignal, a commercial provider of app store analytics. We use monthly snapshots of the Apple App Store from January 2018 to September 2021 covering time periods before and after the introduction of the DiGA regulation. The data includes most of the information available to users at the App Store profile. For each app, we have the unique ID, the name, and the app's developer as identifying information. On the time dimension, we observe an app's release date, the date of its last update and whether the app is still available at any given monthly snapshot. We have information on the languages an app is offered in as well as the list of countries the developer made the app available in. The App Store also includes information on the number of written reviews for the current version as well as all past versions of each app. We also know the genre, whether there is a content rating in place, the app's description text, and privacy settings.

German language availability serves as an indicator for an app targeted at the German market since the DiGA regulation requires the app to be available in German. While developers can choose in which country an app can be downloaded, such country availability is only a weak indicator for the targeted market since there are no costs associated with making an app available in all countries. Instead, making an app available in a specific language is associated, at a minimum, with translation costs if the original language the app was developed in is not German. And even though German is also spoken in other countries, the largest market for German apps is in Germany.[4] We therefore assume that the majority of apps in German are intended to also serve the market in Germany - while apps not available in German are intended for other markets.

Next, we need to identify the apps that could fall under the DiGA regulation. Our aim is to identify patient-centered digital therapeutics which target a specific diagnosis or condition. The App Store genres that categorize apps are not sufficiently specific. The "Medical" category, for example, contains apps for veterinarians, study support for medical students, and programs for medical conferences - none of which fall within the scope of the DiGA regulation. The "Health and Fitness" category includes a large number of wellness apps and fitness trackers - also not targeted by the DiGA regulation. We therefore leverage additional information that developers have to provide when publishing to the Apple App Store. For the content rating section of the app, a developer must specify whether the app has a medical treatment focus.[5] Compared to the Google

[4]Switzerland, Austria, Belgium, Luxemburg and Liechtenstein with German as an official language did not have any regulation similar to DiGA during the time period of our analysis.

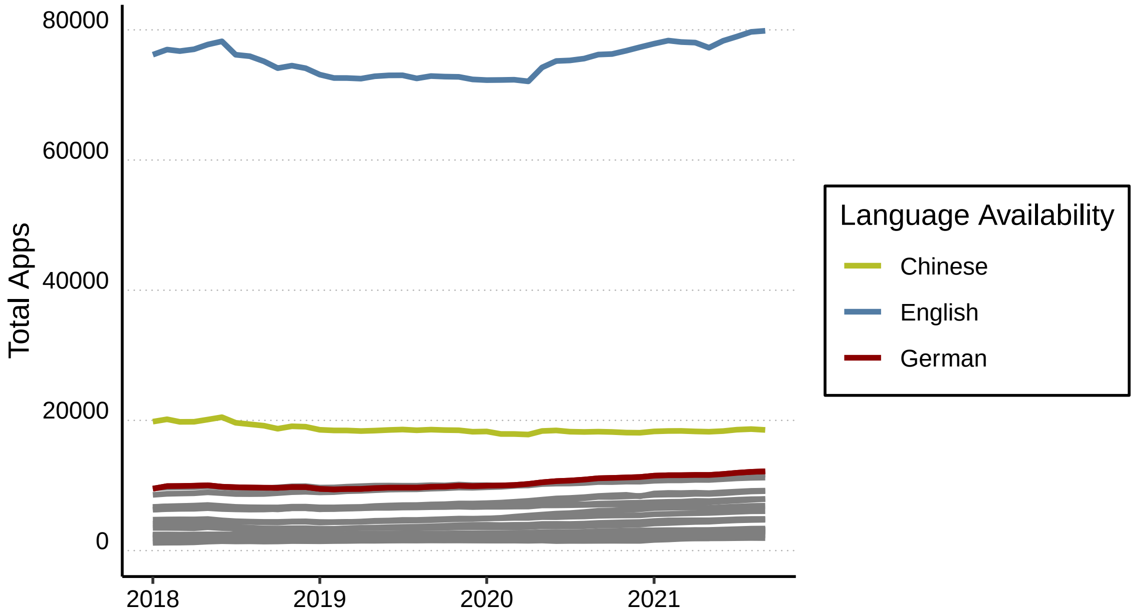
[5]More information on the Apple App Store rating scheme: https://developer.apple.com/help/app-store-connect/reference/age-ratings/.

Play Store, Apple has relatively strict rules for apps with medical focus. Developers agree to disclose any approval information, and if the information is not comprehensive or the medical safety is not ensured, the app can be removed from the store (Sadare et al., 2023). The majority of apps identified that way, however, do fall in the categories “Medical” and “Health and Fitness”. In the remainder of the paper, we refer to those apps as “medical treatment apps”. To expand our data set, we exploit the information in the description field. We use this text to identify the targeted diagnoses for each app. We use a language model, specifically ChatGPT, to match all apps to an ICD-10 (International Classification of Diseases) diagnosis.[6] We compared a subsample of all apps to a test data set with manual classification to assess the accuracy of the diagnosis assignment. Further, we enriched the data set with information on app-related publications using the PubMed API. For each app in our data, we search the PubMed database of medical peer-reviewed journal articles for entries containing both the name of the app and the term “app”.

For our empirical analysis, we group the apps on the month-language level. This gives us a month-language panel with the aggregated number of medical treatment apps. Figure 3 shows how the number of medical treatment apps evolves over time. By far the most apps are available in English (on average 76,184 apps), the second largest group are Chinese apps (on average 19,807) and after that German is the third most common language (on average 9,504 apps). Most European languages are, however, in a similar range. Table A.1 (in the Appendix), provides an overview of the other characteristics of the apps per language group. Not only does the number of apps vary between languages, but so do the average rating and the average number of reviews, which can be seen as indicators of the number of downloads. The number of apps that charge a price is very small, which is an indicator of the pricing challenges in this sector, whereas many apps include in-app purchases.

[6]We used ChatGPT (GPT-3.5-Turbo, OpenAI, May 2024, temperature und top_p set to 0) to support variable creation from raw text data in the app description in our app dataset. The prompt was “Determination of ICD Codes (“International Statistical Classification of Diseases and Related Health Problems”) for App Classification: Your task: Determine the ICD category for the following app description text. You know about all official ICD codes incl. number and name. Follow these rules: 1. Provide the corresponding ICD-10 code on category-level in response, i.e. “XXX - corresponding category name”. 2. If multiple categories seem applicable, always choose the ONE with highest priority. Only provide the following as final output “XXX - corresponding category name.”No explanation.”

Figure 3: Patient-Centered Digital Therapeutics per Language



Notes: Number of patient-centered digital therapeutics from January 2018 until September 2021. Languages are German, English, Chinese, Spanish, French, Italian, Portuguese, Russian, Japanese, Dutch, Korean, Turkish, Polish, Norwegian, Arabic, Greek, Swedish, Danish, Czech, Vietnamese, and Indonesian.

3.2 Estimation Strategy

Across all estimation strategies, we use the same basic setup. Our outcome variable is the number of medical treatment apps. We compare the number of these apps available in German to the number of apps available in other languages.[7] In this setting, one single app can appear in multiple groups if it is available in several languages. We define the government adoption of the draft in July 2019 as the date of policy intervention.[8]

3.2.1 Event Study

We start our analysis with a classical event-study framework. The total number of medical treatment apps in language i in month m, $Y_{i,m}$, is the dependent variable. Event-time indicators $D_{i,j}$ represent the months relative to treatment from August 2019 onwards. We use 20 leads and 25 lags of the treatment indicator to inspect parallel trends pre-treatment and dynamic effects after the treatment took effect. The model includes interactions of these event-time indicators with the treatment group indicator ,$German_i$, allowing us to trace the evolution of the treatment effect for German apps relative to

[7]We do not use the number of apps in other genres (like Games or Shopping) as (placebo) outcome since we cannot exclude that other shocks happened to other genres in any language, for example, an extremely popular game or regulation concerning other genres anywhere in the world.

[8]While early drafts of laws are often changed and sometimes scrapped, the approval in the Cabinet signaled that all parties of the ruling coalition - which had a large majority in parliament - were in favor of the law and it was highly likely to get implemented.

July 2019 (omitted reference). Other languages serve as control group. We control for language fixed effects, α_i, to control for time-invariant unobserved characteristics which influence the availability of apps per language, in particular the overall demand a language-specific app market faces. Additionally, we include month fixed effects, α_m, to control for unobserved time effects which affect the app supply in all languages.

$$Y_{i,m} = \sum_{j=-20}^{-2} \beta_j D_{i,j} \times German_i + \sum_{j=0}^{25} \beta_j D_{i,j} \times German_i + \alpha_i + \alpha_m + \epsilon_{i,m} \tag{1}$$

3.2.2 Synthetic Control Method

Next, we turn to the synthetic control (SC) method introduced by Abadie and Gardeazabal (2003) and later further developed in Abadie et al. (2010) and Abadie et al. (2015). SC methods are used to construct a comparison group which reflects the (counterfactual) development of the German medical treatment app market without the introduction of the DiGA scheme.

The synthetic control group is created by selecting the aggregated group of medical treatment apps in other languages that are similar to those in German in terms of key characteristics. This similarity is determined based on factors such as average price, average rating, and the number of apps in other genres by month and language to reflect the innovation process in the language-specific app market. We then compare the post-intervention outcomes of the German language apps with the synthetic control group. To assess the statistical significance of the estimated effects, classical inference methods are not suitable. The appropriate statistical tests typically conducted are placebo tests. Sensitivity analyses are performed as well to evaluate the robustness of the findings to different specifications and predictor variables. For technical implementation, we follow Abadie (2021).

Training Period. As proposed by Abadie (2021), we do not use the entire pre-period for identification of the synthetic control group. Instead, we use only data for half of the pre-period months from January 2018 to September 2018 to find the control group and use the remaining months until treatment starts in July 2019 to assess the fit.

Donor Pool. To ensure the efficacy of the synthetic control approach, we restrict the donor pool to widely spoken languages. It is crucial for the treatment group and the synthetic control to have similar donor characteristics. As German is the third most frequent language among the apps, the inclusion of language-app combinations with a substantial market size is important. We therefore only use language groups which have a minimum of 100 medical treatment apps at one point in time between January 2018 and September 2021. A list of the potential and actual donors for the synthetic control group is provided in Table A.2. As a rule of thumb, Abadie et al. (2010) suggest excluding units for which the prediction MSE is more than twice the MSE of the treated unit. Following this rule, our donor pool consists of 51 languages.

Predictors. We use predictors in the estimation for the synthetic control group to capture the characteristics of the language-specific app market. Our first set of predictors serves as a proxy for market size. Here, we use the average number of reviews and the share of apps with reviews since reviews increase with the number of people using an app. In a second set of predictors, we use the average rating of the apps and the share of apps with a rating as proxies for app quality. In addition, we proxy for revenue strategies in each market with the share of apps that have a positive price and the share of apps with in-app purchases. To proxy for the characteristics of developers, we use the average number of apps per developerer and the average number of apps in the medical genre per developer as predictors.

The estimation then follows the classical approach outlined in Equation 2:

$$\tau_t = Y_t^G - Y_t^C \quad \text{for all } t > T_0 \tag{2}$$

We want to estimate τ which is the effect of the DiGA regulation in month t. We differentiate the time periods in all months in the pre-treatment period $t = 1, ..., T_0$, and the post-treatment periods $t = T_0 + 1, ..., T$, where T_0 is July 2019. Y^G is the number of patient-centered digital therapeutics in the German language group, Y^C is the number of patient-centered digital therapeutics in the synthetic German language group, hence the counterfactual outcome of German apps without the DiGA regulation. Since the counterfactual is not observed, we build a synthetic control by weighting the outcomes of languages in the donor pool which consists of all language groups not affected by the DiGA introduction ($j = 1, ..., J$.). The weights of the donors are denoted by w.

$$Y_{t=1}^C = \sum_{i=1}^{J} w_j \cdot Y_{j,t=1} \tag{3}$$

The following minimization problem derives the vector of weights $W = w_1, ...w_J$ numerically.

$$\min \left(\sum_{m=1}^{N} v_p (X_{i,p} - X_{j,p} W)^2 \right) \tag{4}$$

where $\sum_{j=1}^{J} w_j = 1$, X_p is the set of covariates and v_p is a weight for each covariate.

There are assumptions which need to hold for this design. First, the provision of patient-centered digital therapeutics in other languages is not affected by the introduction of the German DiGA regulation. This implies there are no spillover effects in other languages, hence apps are not simultaneously provided in German and another language due to the regulation. This is unlikely to hold since developers who implement an app for the German market might also provide this app in other languages. However, it is reasonable to assume that such spillovers take place with a time lag. In addition, in this case, our results would underestimate the real effects. The second assumption, similar to traditional difference-in-difference models, is that no unobserved shocks should affect the outcome differently for treatment and control groups in periods after the reform, i.e., no further policy changes regarding digital therapeutics for patients. This assumption is likely to hold for the periods immediately after the introduction, since there were no further changes for the German system, and the first other countries to announce reimbursement were Belgium and France only after our analysis period. One large shock possibly affecting health app development - the Covid-19 pandemic - affected all countries equally.

3.2.3 Synthetic Difference-in-Differences

Our last empirical strategy combines classical event studies and synthetic controls. The synthetic control approach is based on the idea that a synthetic control is constructed by convex weighting of the underlying control units to get as close a match to the treated group as possible based on pre-treatment outcomes and predictors. This construction implies that the match of the synthetic control and the treatment group needs to be achieved in absolute levels. In our case, only two language groups have, in absolute terms, more apps than the German group. Therefore, even when the fit of their trend is bad, they need to be included with a positive weight in the synthetic control, to make the synthetic control similar to the treated unit in absolute levels. Recently, Arkhangelsky et al. (2021) developed an approach to mitigate this problem. The classical difference-in-differences (DID) approach only requires parallel trends, but no fit on absolute levels. Therefore, the new approach connects DID and synthetic controls into one approach, the so called synthetic difference-in-differences estimator (SDID). Similar to DID models, SDID permits treated and control groups to have a common trend on different absolute levels. The approach takes the benefit from SC of constructing a matched control group and at the same time reduces the strict reliance on the parallel trends assumption. Hence, the benefits of two methods are combined to mitigate the typical challenges encountered in traditional DID and SC methods. It is not possible to estimate causal relationships when parallel trends are not observed in aggregated data for DID, and it is required for the treated unit to be within the encompassing group of control units for SC. The SDID estimator estimates the generated average treatment effect on the treated from a two-way fixed effect regression. The unit (donor) fixed effects imply that treated and control units are matched on pre-trends but not on absolute levels, hence, as in a DID, the FE allows for a constant difference between the groups. In contrast to a standard DID, where this estimation would be conducted with equal weights on units and time periods, the SDID estimator puts different weights on units (donors) ω_c and time periods λ_t. The SC method outlined above uses unit weights, but no time period weights. In terms of parallel trends, the key idea is that the average of the treated group and the weighted average of the control groups show similar trends after treatment, and that this similarity holds primarily for the post-treatment period and the weighted average of the pre-treatment period, rather than assuming parallel trends for all groups and all time periods.

The estimator requires a balanced panel as we have in our case. We have N donor units, the language groups. As described above, we use 51 languages in the donor pool. The time periods T are all months from January 2018 to October 2021. The outcome is again the number of medical treatment apps per language (i) and month (t), Y_{it}.

First, the weights are determined by the following minimization problems as laid out by Arkhangelsky et al. (2021).

$$(\hat{\omega}_0, \hat{\omega}_{\text{sdid}}) = \arg\min_{\omega_0, \omega_{\text{sdid}}} \sum_{t=1}^{T_{pre}} (\omega_0 + \sum_{i=1}^{N_c} \omega_i Y_{it} - Y_{it})^2 + \zeta^2 T_{pre} ||\omega||_2^2 \tag{5}$$

ω_0 is an intercept allowing for constant difference between treatment and control units, $||\omega||_2$ stands for the Euclidean norm and ζ is a regularization parameter (for details see Arkhangelsky et al. (2021), pp. 4091-4092). We obtain a non-negative weight for all donors on the donor pool N_c and the sum of all weights is 1. The estimation for the time period weights λ follows a similar process.

$$(\hat{\lambda}_0, \hat{\lambda}_{sdid}) = \arg\min_{\lambda_0, \lambda_{sdid}} \sum_{n=1}^{N_c} (\lambda_0 + \sum_{t=1}^{T_{pre}} \lambda_i Y_{it} - Y_{it})^2 + \zeta^2 N_c ||\lambda||_2^2 \tag{6}$$

Second, the two-way fixed effects regression with unit (α_c) and time (β_t) fixed effects estimates the average treatment effect on the treated (τ) of the treatment (D_{ct}) with the identified weights.

$$(\hat{\tau}_{\text{sdid}}, \hat{\mu}, \hat{\alpha}, \hat{\beta}) = \arg\min_{\tau,\mu,\alpha,\beta} \left\{ \sum_{c=1}^{N} \sum_{t=1}^{T} (Y_{ct} - \mu - \alpha_c - \beta_t - D_{ct}\tau)^2 \, \hat{w}_{\text{csdid}} \hat{\lambda}_{\text{sdid}t} \right\} \tag{7}$$

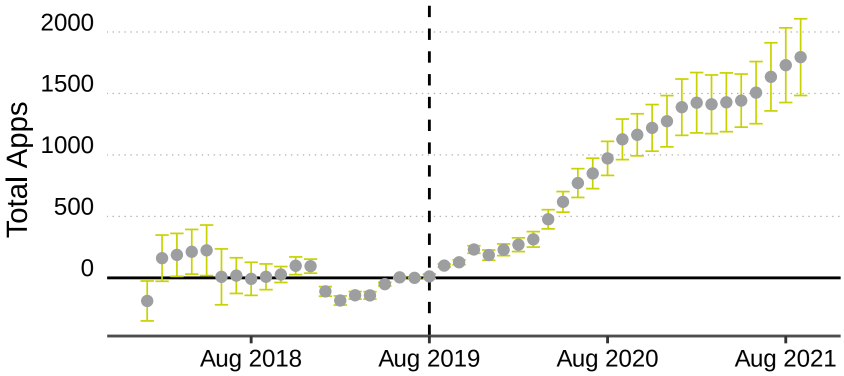
When using the SDID approach, we do not control for covariates. In the classical SC approach, the covariates serve to ensure they are as closely matched as possible between the treated unit and synthetic control unit. In the SDID approach, the heterogeneity explainable by covariates is removed before calculating the synthetic control. Arkhangelsky et al. (2021) argue that this way of using covariates causes several problems and should only be used when time varying covariates seem crucial. Since this is not the case in our setting, especially as we work with aggregated information from only the App Store, we do not include controls. We have one treated group only, the German language group, hence the corresponding standard errors can only be estimated via a placebo approach. This follows the same idea as in the classical SC approach (Abadie et al., 2010). The effect τ is estimated iteratively assigning the treatment randomly to one language in the donor pool. By repeating the process, we retrieve an entire vector of τ and calculate the placebo variance. Then, we rank the treated and all placebo estimates by their RMSPE (root of the mean squared prediction error) ratios ($RMSPE_{post}/RMSPE_{pre}$). The rank of the true treatment estimate compared to the N placebo estimates gives the p-value for the null hypothesis of no treatment effect ($RMSPE - Ratio_{Treat} \leq RMSPE - Ratio_{Placebo}$). Here, the p-value is the rank of the treated unit divided by the number of controls+1. A requirement for valid inference here is homoskedasticity. A Breusch-Pagan test gives no indication for violation of this assumption.

3.3 Empirical Results

3.3.1 Event Study

The results for the event study in Figure 4 show that the number of available medical treatment apps in German was similar to the other languages until the DiGA regulation was passed in the Cabinet. Thereafter, the number of apps available in German increased strongly compared to apps in other languages. Approximately one year after the policy intervention in August 2020, the number of medical treatment apps available in German had increased by around 970 relative to other languages. By the end of the analysis period, September 2021, this difference had grown to more than 1,700 additional German-language apps, compared to a baseline of 9,500 apps, suggesting substantial impact of the DiGA regulation (see Appendix Table A.3 for detailed quantitative results).

Figure 4: Event Study



Notes: The graph displays the result of the event study estimation where the total number of apps in German is compared to the total number of apps in other languages from January 2018 until September 2021 based on Equation 1. The (omitted) reference month is July 2019, the month when the German Cabinet passed the DiGA draft. The vertical dashed line indicates August 2019, the first month after the draft was passed. For each coefficient, we also plot 95% confidence intervals. For the estimation we use a language-month panel collapsed from all apps in the Apple App store where the content rating indicates a medical treatment focus.

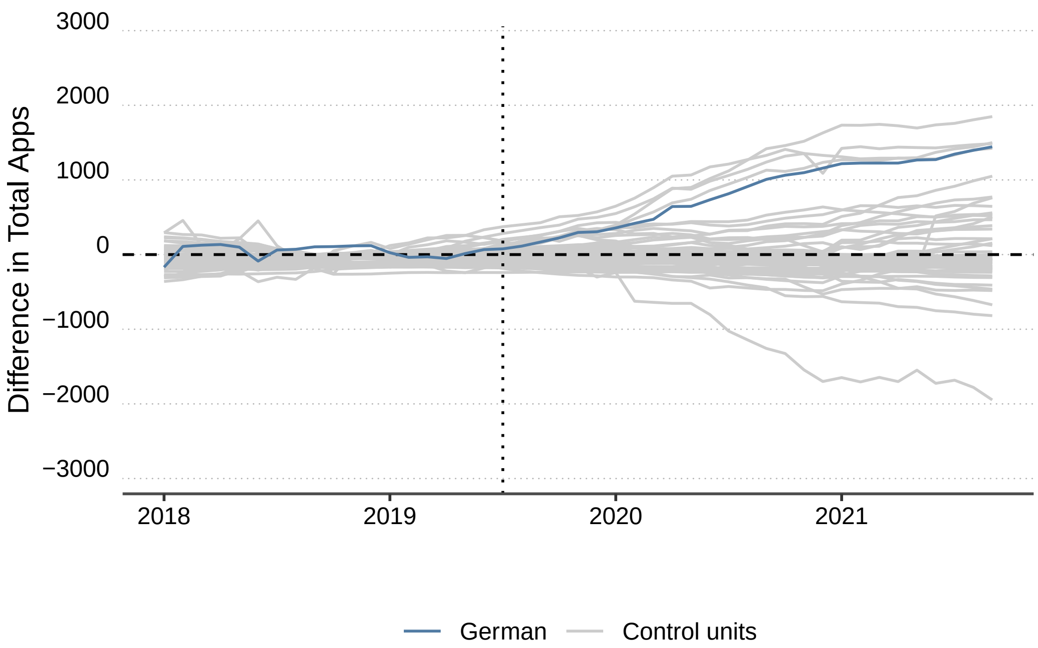
3.3.2 Synthetic Control Approach

Next, we estimate a standard SC model. The mean values of our covariates from the synthetic German group are much more similar to the actual German sample after matching (see Table A.4, in the Appendix). For the construction of the SC, the covariate with the highest weight (44 percent) is the share of apps with in-app purchases. The share of apps with any review receives the second highest weight (20 percent). In Table A.2 (in the Appendix), we present the weights assigned to the different languages from the donor pool to construct the synthetic German group. Spanish receives the highest weight, but in general the weights are split smoothly between the 47 languages. Four languages (Cambodian, Estonian, Malayalam, Swedish) receive no weight. This result shows that we do not depend on a single donor in our SC estimation.

The main results are summarized in Figure 5. The blue line represents the difference between the number of German apps and the number of apps for the synthetic German for each month in our data. This difference is close to zero before the DiGA regulation was passed in the Cabinet (for a comparison of the absolute numbers, see Figure A.1 in the Appendix). Since the synthetic control was constructed based only on the training period until October 2018 but the match remains good in the validation period from November 2018 onwards, we can be reassured about the effectiveness of the match. After the DiGA regulation was passed in the Cabinet, the number of patient-centered health care apps in German increased strongly compared to the synthetic control. In the first months of the intervention period, the difference between the German apps and the synthetic control is similar in magnitude to differences in the pre-period. Within the first year following the intervention—up to August 2020—the estimates indicate an increase of 911 additional medical treatment apps available in German (see Table A.5 in the Appendix for the specific numbers in each month). While this effect is somewhat smaller than the estimate from the event study, it remains of a similar order of magnitude, reinforcing the evidence of a substantial policy impact.

To assess the statistical significance of the increase in German language apps, classical inference methods cannot be used. We therefore use placebo inference. The grey lines in Figure 5 show a sample of placebo estimations, where a random language other than German is considered as treated. For most placebo tests, we find a substantially smaller or even negative effect. To compute p-values, we sort all placebo estimations by the treatment effect size. The share of languages that are more likely to show an effect than the German group is the p-value (see Figure A.2). Taking our estimations, the p-value for the increase in German language apps is 5/52 = 0.096, indicating a significant increase at the 10 percent significance level.

Figure 5: Synthetic Control Method



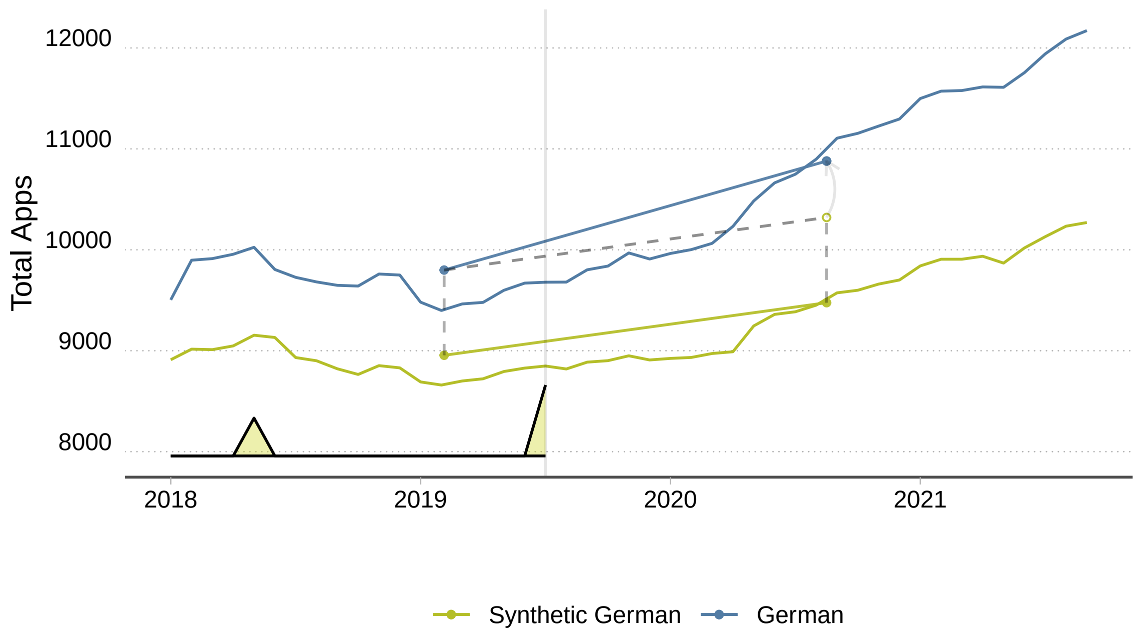
Notes: The graph displays the result of the SC estimation from January 2018 until September 2021. The blue line shows the difference between the number of German apps and the number of apps in the synthetic German. Additionally, the graph shows the differences for placebo estimations in grey, where a control group is constructed for other languages which were in fact not treated. The vertical dashed line indicates August 2019, the first month after the German Cabinet passed the DiGA draft. For the estimation we use a language-month panel collapsed from all apps in the Apple App store where the content rating indicates a medical treatment focus.

A common threat in SC estimations is single-donor dependence, where the results depend on the evolution of one single donor unit. To test whether such dependencce influences our result, we follow a leave-one-out approach, reestimating the model and omitting one donor at a time. Figure A.3 displays the most extreme results for these reestimations. Although there is clear variation in the results, none of these reestimations suggest a different pattern from our main estimation. However, it becomes clear that while the trend of the German language apps can be matched for all leave-one-out models, not all new donor pools can match the trend of the German apps in absolute terms.

3.3.3 Synthetic Difference-in-Differences Approach

The SDID approach tackles the issue of matching in absolute terms. In our estimation of the SDID, no language receives a higher weight than seven percent which further reduces the dependence of the results on single donors (see Table A.6 in the Appendix for the full list of weights). Figure 6 shows the main results of the SDID estimation (the corresponding values are displayed in Table A.7 in the Appendix). A feature which sets the method apart from conventional synthetic controls is the introduction of time period weights. The red triangles in Figure 6 display these weights. Only two periods (May 2018, $\lambda = 0.306$; July 2019, $\lambda = 0.694$) receive a positive weight. The parallelogram shows the DID incorporated in the method where the grey dashed line indicates the parallel trend to the synthetic control. The grey arrow shows the average treatment effect on the treated. According to our third estimation approach, approximately 600 additional medical treatment apps became available in German within the first year following the introduction of the new reimbursement scheme up to August 2020. Similar to the results from the synthetic control method, we observe a substantial increase in German-language apps, albeit of a smaller magnitude. When we overlay the two time series, (see Figure A.4 in the Appendix) we find that the pre-treatment match between the synthetic control and the German group worked well again. In this plot, we also indicate the confidence interval for the estimated treatment effect. Although the qualitative pattern from the SDID estimation is comparable to the event study and SC approaches, the difference is not significant in the SDID estimation.

Figure 6: Synthetic Difference-in-Differences



Notes: The graph displays the result of the SDID estimation from January 2018 until September 2021. The green triangles indicate the weights on pre-treatment periods and the light grey arrow indicates the average treatment effect on the treated. The parallelogram illustrates the part of the difference between German and synthetic German which can be explained by fixed effects. The vertical grey line indicates August 2019, the first month after the German Cabinet passed the DiGA draft. For the estimation we use a language-month panel collapsed from all apps in the Apple App store where the content rating indicates a medical treatment focus.

3.4 Robustness

In this section, we want to test whether our estimated treatment effect is stable in alternative specifications. Therefore, we reestimate the SC and SDID models using different predictor sets and sample compositions.

3.4.1 Alternative Predictors

First, we enlarge the predictor set of the SC model. We use the number of total apps in other genres in the App Store as additional predictors to proxy market size in a language.[9] In this estimation, some of the genres receive a substantial predictor weight (e.g., Food & Drink) while others show no relevance. The results of this specification are very similar to the original specification, however, the pre-trend match is worse in absolute terms (see Table A.8, Figure A.5, Table A.9, and Figure A.6 in the Appendix).

[9] We include the total number of apps in the genres: Business, Education, Entertainment, Finance, Food&Drink, Games, Health&Fitness, Lifestyle, Music, Navigation, Photo&Video, Productivity, Social&Networking, Sports, Travel, Utilities, Weather.

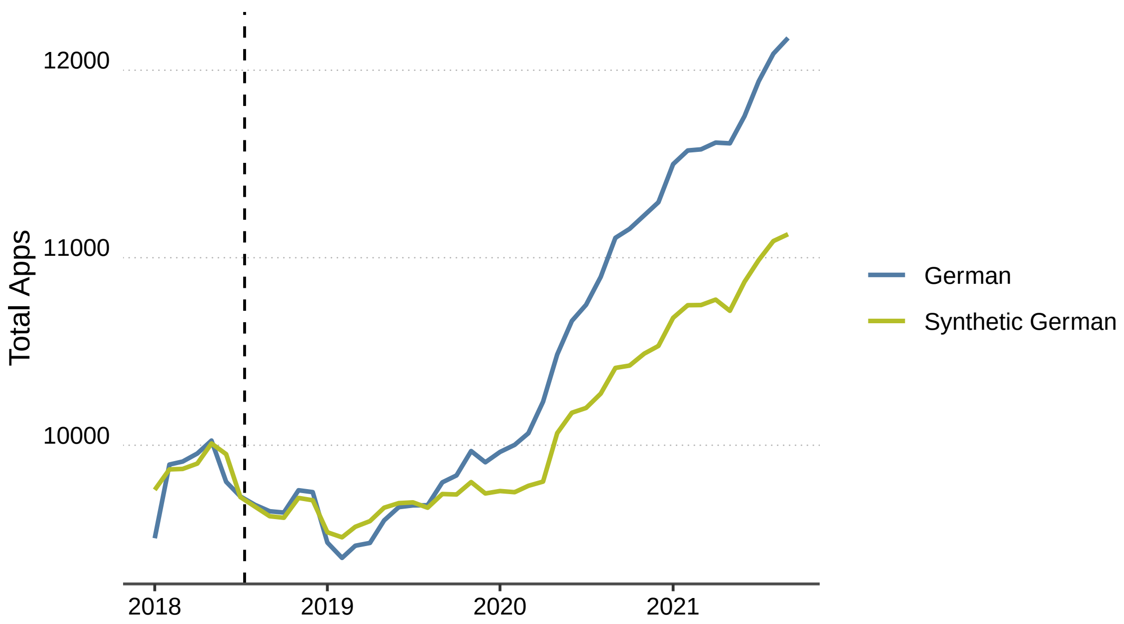
3.4.2 Alternative Donor Pool

In the main model, the selection of the donor pool is motivated by minimal market size. To increase the similarity of donors in the donor pool, we repeat the SC analyses with a donor pool of the top 20 languages in terms of number of medical treatment apps. This new donor group is more similar to the German language apps in terms of market size as well as other characteristics compared to the original donor pool. The results are again very similar to the main specification (see Figure A.7, Figure A.8, Figure A.9, Figure A.10, Figure A.11, Figure A.12, and Table A.10 in the Appendix).

3.4.3 Placebo in Time

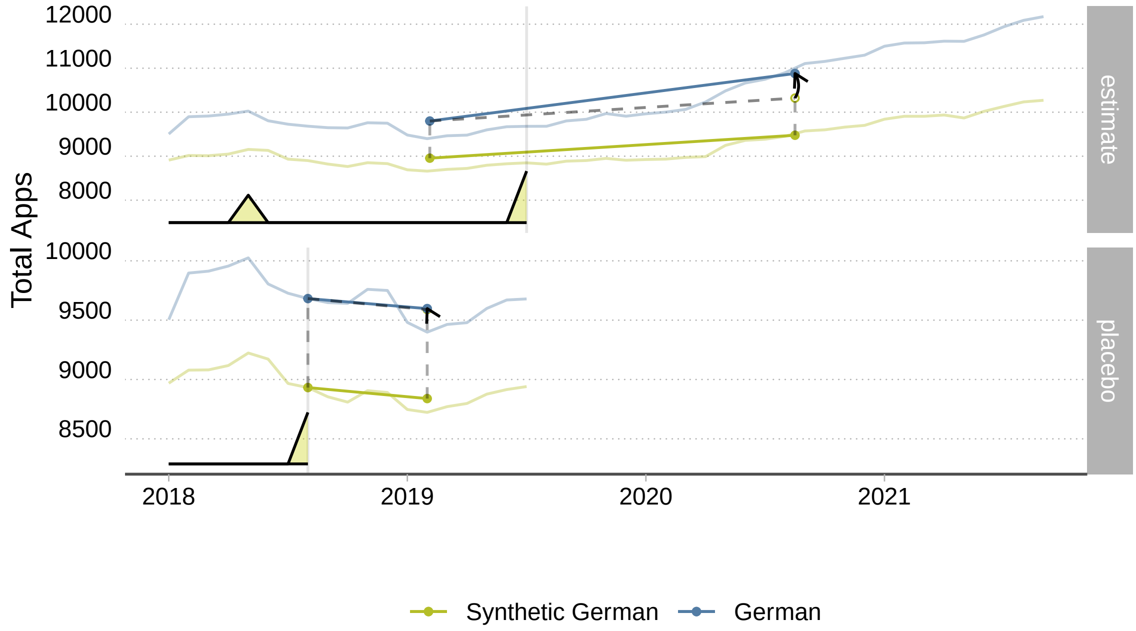
We perform a placebo in time analysis for the SC and the SDID method where we estimate the treatment effect for a placebo intervention in October 2018, ten months before the actual intervention. The null-effect for these placebo estimations in Figure 7 and Figure 8 is reassuring.

Figure 7: Synthetic Control - Placebo in Time



Notes: The graph displays a placebo in time analysis for the synthetic control method. The treatment is assigned in August 2018.

Figure 8: Synthetic Difference-in-Differences - Placebo in Time



Notes: The graph displays in the upper model the result of the main model of synthetic difference-in-differences estimation from January 2018 until December 2021 and in the lower panel a placebo in time estimation in the pre-treatment period (August 2018).

3.4.4 Covid Exclusion

The post period includes the time of the Covid-19 pandemic. One phenomenon of the pandemic was that a lot of apps were put on the market e.g., for contact tracing. Although most of these apps would not receive a medical treatment rating in the App Store and the shock affected all countries, we want to ensure that Covid-related apps do not inflate the results. We exclude all apps which mention the virus or the disease in their app description. The results remain unchanged in all specifications (see Figure A.13 for synthetic control, see Figure A.14 for synthetic difference-in-differences, see Figure A.15 for event study).

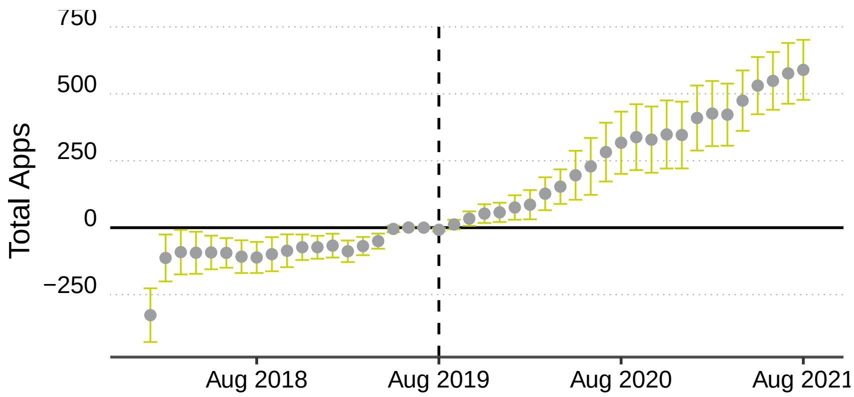
3.5 Heterogeneity Analyses

There is a consistent pattern in all our estimation strategies that more medical treatment apps are available in German after the introduction of the DiGA scheme. To gain more insights into the mechanisms behind this increase, we analyze subgroups of apps in the following sections.

3.5.1 App Usage

Not all apps available in the App Store are frequently used. To restrict our estimates to apps that do have a reasonable number of users, we restrict the analysis to apps that receive at least one review within their first year in the App Store. This serves as a proxy for apps that are not only available but also downloaded and used. As shown in Figure 9, the magnitude of the effect is significantly smaller for this group: the increase is only about a quarter of that observed in the unrestricted sample. This smaller effect suggests that while more providers enter the medical treatment app market, only a few apps are actively used. This may indicate that many of the additional apps offer no substantial benefits, lack real innovation, or are of low quality. However, it should be acknowledged that there are pre-trends in the pre-intervention period, which limits the interpretability of the results.

Figure 9: Event Study - Apps with Review within a Year



Notes: The graph displays the result of the event study estimation from January 2018 until September 2021. The sample is restricted to those apps which receive at least one review within the first year available.

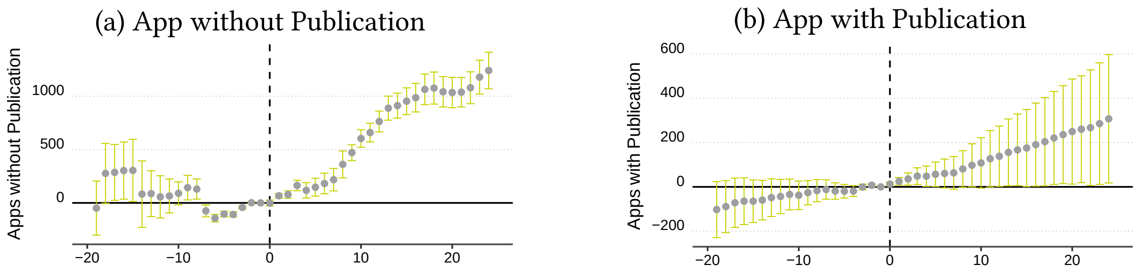
3.5.2 Quality of Apps

A quality signal for a health app is whether it has been evaluated in a scientific study. We assess which medical treatment apps have at least one related published paper in the PubMed database (one of the largest collections of peer-reviewed medical journal articles). We interpret the existence of a peer-reviewed published study as a quality indicator. As shown in Figure 10, the increase in the number of apps available in German is mainly driven by apps which do not have a corresponding scientific study. This suggests that the health benefits of these additional apps are unclear and may lack a scientific basis.

3.5.3 Data Privacy

Another important aspect of app quality for users is data privacy. If the introduction of the DiGA scheme incentivizes developers to enter the market, they should meet high privacy standards. The DiGA scheme allows a shift from a business model where apps are provided for free and monetized through advertising and data collection to a model

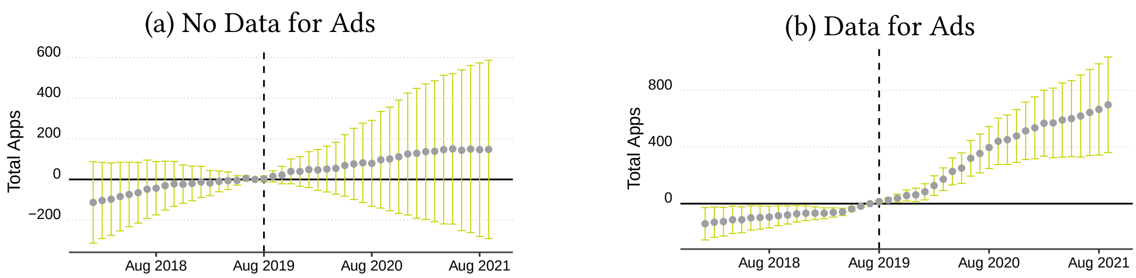
Figure 10: Event Study by Publication



Notes: The left graph displays the result of the event study estimation from January 2018 until September 2021 for apps with no publication listed in PubMed, the right graph for apps with publication.

with high reimbursement and strict data privacy requirements. In the Apple App Store, developers must disclose how they collect and use data. We classify the medical treatment apps into two groups: One that collects data for advertising and one that does not. The results in Figure 12 show that only a small share of the increase in medical treatment apps comes from those with high data privacy standards. Instead, most apps still collect data for advertising. This suggests that the potential for reimbursement alone may not be the main motivation for entering the market. Alternatively, developers may only transition away from an ad-based business model once their app is officially listed and generates revenue through the new reimbursement system.

Figure 12: Event Study by Privacy

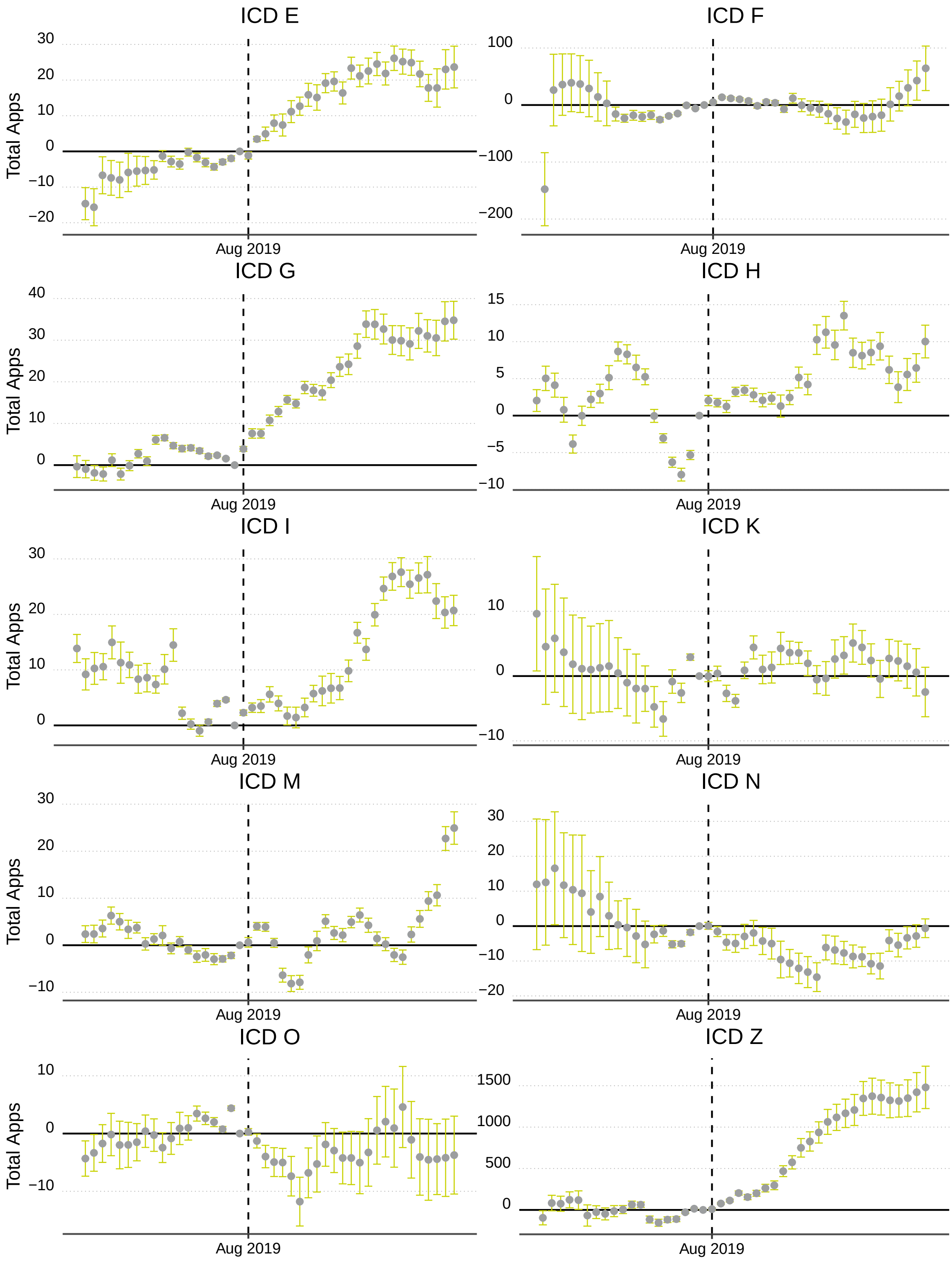


Notes: The left graph displays the result of the event study estimation from January 2018 until September 2021 for apps which do not collect data for advertisement, the right graph for apps which collect data for advertisement.

3.5.4 Targeted Diagnoses of Apps

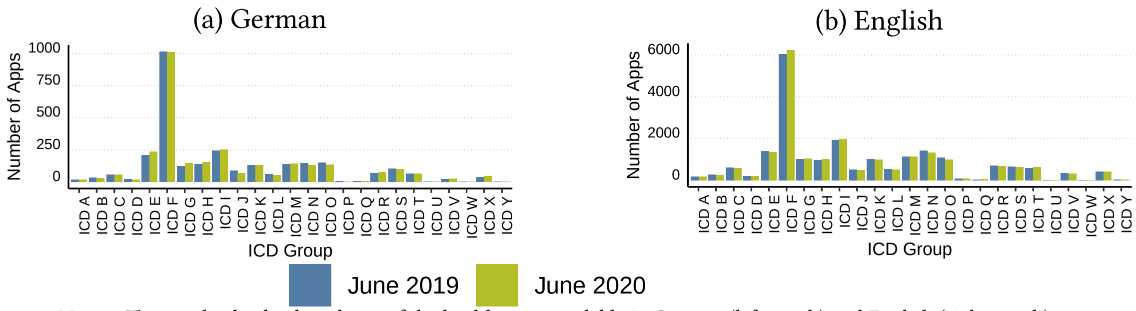
From a health system perspective, another important aspect is the variety of medical treatment apps in the medical conditions they target. To analyze this variety, we use the targeted diagnoses of apps based on their descriptions in the App Store. Using the first three digits of the assigned ICD-10 codes, we find that no newly added German app introduces a completely new diagnosis category at this level. When examining the distribution of newly introduced apps across disease categories, we observe that the vast majority fall into ICD Category Z (see Figure 14). This category includes factors influencing health status and contact with health services, such as preventive care, medical check-ups, rehabilitation, and counseling. Beyond this, there are only slight increases in a few other ICD groups, e.g., ICD Group E with endocrine, nutritional, and metabolic diseases (e.g., diabetes or obesity), ICD Group G with diseases of the nervous system (e.g., epilepsy or multiple sclerosis), ICD Group H with diseases of the eye and ear, or ICD Group I with diseases of the circulatory system (e.g., hypertension). This suggests that while more apps are entering the market, they mainly target general health and medical service-related topics, with only marginal expansions in specific medical fields. If we compare the distribution of medical treatment apps overall — not just newly introduced ones — in both English and German for June 2019 (before the introduction of the scheme) and June 2020 (after its introduction), we observe that the distribution across ICD groups is similar. There is a strong focus on apps targeting ICD group F, which includes mental and behavioral disorders.

Figure 14: Event Study - ICD Groups



Notes: The graphs display the result of the event study estimation from January 2018 until September 2021 separated by apps targeted towards diagnoses of ICD groups. ICD Group Descriptions (E: Endocrine, nutritional, and metabolic diseases; F: Mental and behavioral disorders; G: Diseases of the nervous system; H: Diseases of the eye, ear, and related structures; I: Diseases of the circulatory system; K: Diseases of the digestive system; M: Diseases of the musculoskeletal system and connective tissue; N: Diseases of the genitourinary system; O: Pregnancy, childbirth, and the puerperium; Z: Factors influencing health status and contact with health services). Omitted ICD groups show no effect.

Figure 15: Apps by ICD Groups



Notes: The graphs display bar charts of the health apps available in German (left graph) and English (right graph) in June 2019 and June 2020 by targeted ICD diagnoses of the app.

4 Conclusion

In this paper, we analyze how the introduction of reimbursement for digital therapeutics influences market entry decisions leveraging the German DiGA reimbursement scheme. Our results show an upward trend in German-language app-based digital therapeutics after the structured reimbursement pathway was announced. The estimated effect size varies across approaches, ranging from event studies to synthetic control methods. Qualitatively, the results are consistent across the different estimation strategies. The increase in German-language apps is statistically significant in the event study specification, our preferred model. However, the estimates are only marginally significant in the synthetic control estimations, and not statistically significant in the synthetic difference-in-differences analyses. One challenge for the statistical inference in the synthetic control setups is the single-treated unit in the analysis.

The expansion in German health apps appears to be largely quantitative rather than qualitative. We find no evidence of a shift toward improved data privacy, since the increase is driven primarily by apps with data collection for advertisement. We also find no broadening of therapeutic areas covered by apps, since the number of distinct three-digit ICD diagnoses associated with the app descriptions is unchanged. In addition, relatively few of the newly developed apps can be linked to peer-reviewed scientific publications. Such evidence would be necessary to shed light on the clinical effectiveness of the apps.

Taken together, our findings suggest that while the introduction of the DiGA scheme initially acted as a positive shock, accelerating the development of health apps, this early enthusiasm did not translate into a sustained wave of high-quality innovations. The relatively small number of apps that have successfully gone through the DiGA approval process underlines this interpretation. Even though the DiGA reimbursement scheme set strong monetary incentives for developers to enter the market, other aspects of the scheme might act as barriers, i.e., high evidence requirements or bureaucratic burdens in the highly regulated German health care system.

This result provides important insights for policymakers. If structured reimbursement is introduced to foster innovation and increasing quality, more attention may be needed on how to balance entry incentives with quality assurance mechanisms. The German DiGA scheme with its high reimbursement rates might have set regulatory requirements too high for most developers. Given that at the moment, medical treatment apps are exclusively low-risk medical devices, a scheme with lower requirements and lower reimbursement might be more beneficial for patients. Future research should examine which specific hurdles developers face in the approval process, and how similar policies might be adopted in other countries with different regulatory environments.

References

Abadie, A. (2021). Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects. *Journal of Economic Literature*, *59*(2), 391–425. https://doi.org/10.1257/jel.20191450

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association*, *105*(490), 493–505. https://doi.org/10.1198/jasa.2009.ap08746

Abadie, A., Diamond, A., & Hainmueller, J. (2015). Comparative Politics and the Synthetic Control Method. *American Journal of Political Science*, *59*(2), 495–510. https://doi.org/10.1111/ajps.12116

Abadie, A., & Gardeazabal, J. (2003). The Economic Costs of Conflict: A Case Study of the Basque Country. *American Economic Review*, *93*(1), 113–132. https://doi.org/10.1257/000282803321455188

Acemoglu, D., & Linn, J. (2004). Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry. *The Quarterly Journal of Economics*, *119*(3), 1049–1090. https://doi.org/10.1162/0033553041502144

Agarwal, R., & Gaule, P. (2022). What drives innovation? Lessons from COVID-19 R&D. *Journal of Health Economics*, *82*, 102591. https://doi.org/10.1016/j.jhealeco.2022.102591

Aitken, M., & Nass, D. (2024). DigitalHealth Trends 2024: Implications for reserach and patient Care. *IQVIA Report*.

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2021). Synthetic Difference-in-Differences. *American Economic Review*, *111*(12), 4088–4118. https://doi.org/10.1257/aer.20190159

Blume-Kohout, M. E., & Sood, N. (2013). Market size and innovation: Effects of Medicare Part D on pharmaceutical research and development. *Journal of Public Economics*, *97*, 327–336. https://doi.org/10.1016/j.jpubeco.2012.10.003

Böhringer, C., Cuntz, A., Harhoff, D., & Asane-Otoo, E. (2017). The impact of the German feed-in tariff scheme on innovation: Evidence based on patent filings in renewable energy technologies. *Energy Economics*, *67*, 545–553. https://doi.org/10.1016/j.eneco.2017.09.001

Bonoto, B. C., de Araújo, V. E., Godói, I. P., de Lemos, L. L. P., Godman, B., Bennie, M., Diniz, L. M., & Junior, A. A. G. (2017). Efficacy of Mobile Apps to Support the Care of Patients With Diabetes Mellitus: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *JMIR mHealth and uHealth*, *5*(3), e4. https://doi.org/10.2196/mhealth.6309

Brönneke, J. B., Herr, A., Reif, S., & Stern, A. D. (2023). Dynamic HTA for digital health solutions: Opportunities and challenges for patient-centered evaluation. *International Journal of Technology Assessment in Health Care*, *39*(1), e72. https://doi.org/10.1017/S0266462323002726

Cajita, M. I., Gleason, K. T., & Han, H.-R. (2016). A Systematic Review of mHealth-Based Heart Failure Interventions. *Journal of Cardiovascular Nursing*, *31*(3), E10–E22. https://doi.org/10.1097/JCN.0000000000000305

Cheyre, C., Leyden, B. T., Baviskar, S., & Acquisti, A. (2023). The Impact of Apple's App Tracking Transparency Framework on the App Ecosystem. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4453463

Chong, S. O. K., Pedron, S., Abdelmalak, N., Laxy, M., & Stephan, A.-J. (2023). An umbrella review of effectiveness and efficacy trials for app-based health interventions. *npj Digital Medicine*, *6*(1), 233. https://doi.org/10.1038/s41746-023-00981-x

Deng, Y., Lambrecht, A., & Liu, Y. (2022, April). Spillover Effects and Freemium Strategy in the Mobile App Market. https://doi.org/10.2139/ssrn.3149550

Dubois, P., De Mouzon, O., Scott-Morton, F., & Seabright, P. (2015). Market size and pharmaceutical innovation. *The RAND Journal of Economics*, *46*(4), 844–871. https://doi.org/10.1111/1756-2171.12113

Duggan, M., Healy, P., & Morton, F. S. (2008). Providing Prescription Drug Coverage to the Elderly: America's Experiment with Medicare Part D. *Journal of Economic Perspectives*, *22*(4), 69–92. https://doi.org/10.1257/jep.22.4.69

Duggan, M., & Morton, F. S. (2010). The Effect of Medicare Part D on Pharmaceutical Prices and Utilization. *American Economic Review*, *100*(1), 590–607. https://doi.org/10.1257/aer.100.1.590

European Federation of Pharmaceutical Industries and Associations. (2023). *Improving access to digital therapeutics in Europe* (tech. rep.).

Federal Institute for Drugs and Medical Devices (BfArM). (2020). *The Fast-Track Process for Digital Health Applications (DiGA) according to Section 139e SGB V. A Guide for Manufacturers, Service Providers and Users* (tech. rep.). Federal Institute for Drugs and Medical Devices (BfArM). Bonn.

Federal Institute for Drugs and Medical Devices (BfArM). (2025a). BfArM - Wissenswertes zu DiGA.

Federal Institute for Drugs and Medical Devices (BfArM). (2025b). DiGA-Verzeichnis.

Finkelstein, A. (2004). Static and Dynamic Effects of Health Policy: Evidence from the Vaccine Industry. *The Quarterly Journal of Economics*, *119*(2), 527–564. https://doi.org/10.1162/0033553041382166

Forsström, J. J., & Rigby, M. (1999). Considerations on the quality of medical software and information services. *International Journal of Medical Informatics*, *56*(1-3), 169–176. https://doi.org/10.1016/S1386-5056(99)00044-1

Frankovic, I., & Kuhn, M. (2023). Health insurance, endogenous medical progress, health expenditure growth, and welfare. *Journal of Health Economics*, *87*, 102717. https://doi.org/10.1016/j.jhealeco.2022.102717

Gensorowsky, D., Witte, J., Batram, M., & Greiner, W. (2022). Market access and value-based pricing of digital health applications in Germany. *Cost Effectiveness and Resource Allocation*, *20*(1), 25. https://doi.org/10.1186/s12962-022-00359-y

Gerke, S., Stern, A. D., & Minssen, T. (2020). Germany's digital health reforms in the COVID-19 era: Lessons and opportunities for other countries. *npj Digital Medicine*, *3*(1), 94. https://doi.org/10.1038/s41746-020-0306-7

Ghose, A., Guo, X., Harbin Institute of Technology, Li, B., Carnegie Mellon University, Dang, Y., & South China University of Technology. (2022). Empowering Patients Using Smart Mobile Health Platforms: Evidence of a Randomized Field Experiment. *MIS Quarterly*, *46*(1), 151–192. https://doi.org/10.25300/MISQ/2022/16201

Gordon, W. J., Landman, A., Zhang, H., & Bates, D. W. (2020). Beyond validation: Getting health apps into clinical practice. *npj Digital Medicine*, *3*(1), 14. https://doi.org/10.1038/s41746-019-0212-z

Han, M., & Lee, E. (2018). Effectiveness of Mobile Health Application Use to Improve Health Behavior Changes: A Systematic Review of Randomized Controlled Trials. *Healthcare Informatics Research*, *24*(3), 207. https://doi.org/10.4258/hir.2018.24.3.207

Iribarren, S. J., Akande, T. O., Kamp, K. J., Barry, D., Kader, Y. G., & Suelzer, E. (2021). Effectiveness of Mobile Apps to Promote Health and Manage Disease: Systematic Review and Meta-analysis of Randomized Controlled Trials. *JMIR mHealth and uHealth*, *9*(1), e21563. https://doi.org/10.2196/21563

Janßen, R., Kesler, R., Kummer, M., & Waldfogel, J. (2022). *GDPR and the Lost Generation of Innovative Apps* (tech. rep. No. w30028). National Bureau of Economic Research. Cambridge, MA. https://doi.org/10.3386/w30028

Kesler, R. (2022). The Impact of Apple's App Tracking Transparency on App Monetization. https://doi.org/10.2139/ssrn.4090786

Kesler, R., Kummer, M., & Schulte, P. (2019). Competition and Privacy in Online Markets: Evidence from the Mobile App Industry. https://doi.org/10.2139/ssrn.3542203

Ketcham, J. D., & Simon, K. I. (2008). Medicare Part D's effects on elderly patients' drug costs and utilization. *The American Journal of Managed Care*, *14*(11 Suppl), SP14–22.

Khan, Z. A., Kidholm, K., Pedersen, S. A., Haga, S. M., Drozd, F., Sundrehagen, T., Olavesen, E., & Halsteinli, V. (2024). Developing a Program Costs Checklist of Digital Health Interventions: A Scoping Review and Empirical Case Study. *PharmacoEconomics*, *42*(6), 663–678. https://doi.org/10.1007/s40273-024-01366-y

Kraft, L., Skiera, B., & Koschella, T. (2023). Economic Impact of Opt-in versus Opt-out Requirements for Personal Data Usage: The Case of Apple's App Tracking Transparency (ATT). *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4598472

Kummer, M., & Schulte, P. (2019). When Private Information Settles the Bill: Money and Privacy in Google's Market for Smartphone Applications. *Management Science*, *65*(8), 3470–3494. https://doi.org/10.1287/mnsc.2018.3132

Larsen, M. E., Huckvale, K., Nicholas, J., Torous, J., Birrell, L., Li, E., & Reda, B. (2019). Using science to sell apps: Evaluation of mental health app store quality claims. *npj Digital Medicine*, *2*(1), 18. https://doi.org/10.1038/s41746-019-0093-1

Lindman, Å., & Söderholm, P. (2016). Wind energy and green economy in Europe: Measuring policy-induced innovation using patent data. *Applied Energy*, *179*, 1351–1359. https://doi.org/10.1016/j.apenergy.2015.10.128

Lu, C., Hu, Y., Xie, J., Fu, Q., Leigh, I., Governor, S., & Wang, G. (2018). The Use of Mobile Health Applications to Improve Patient Experience: Cross-Sectional Study in Chinese Public Hospitals. *JMIR mHealth and uHealth*, *6*(5), e126. https://doi.org/10.2196/mhealth.9145

MedTech Europe. (2021). *Recognising the value of digital health apps: An assessment of five European healthcare systems* (tech. rep.).

Mira, J. J., Navarro, I., Botella, F., Borrás, F., Nuño-Solinís, R., Orozco, D., Iglesias-Alonso, F., Pérez-Pérez, P., Lorenzo, S., & Toro, N. (2014). A Spanish Pillbox App for Elderly Patients Taking Multiple Medications: Randomized Controlled Trial. *Journal of Medical Internet Research*, *16*(4), e99. https://doi.org/10.2196/jmir.3269

Murfitt, R. R. (1990). United States Government regulation of medical device software: A review. *Journal of Medical Engineering & Technology*, *14*(3), 111–113. https://doi.org/10.3109/03091909009015422

Roma, P., & Ragaglia, D. (2016). Revenue models, in-app purchase, and the app performance: Evidence from Apple's App Store and Google Play. *Electronic Commerce Research and Applications*, *17*, 173–190. https://doi.org/10.1016/j.elerap.2016.04.007

Sadare, O., Melvin, T., Harvey, H., Vollebregt, E., & Gilbert, S. (2023). Can Apple and Google continue as health app gatekeepers as well as distributors and developers? *npj Digital Medicine*, *6*(1), 8. https://doi.org/10.1038/s41746-023-00754-6

Schmidt, L., Pawlitzki, M., Renard, B. Y., Meuth, S. G., & Masanneck, L. (2024). The three-year evolution of Germany's Digital Therapeutics reimbursement program and its path forward. *npj Digital Medicine*, *7*(1), 139. https://doi.org/10.1038/s41746-024-01137-1

Sim, I. (2019). Mobile Devices and Health. *New England Journal of Medicine*, *381*(10), 956–968. https://doi.org/10.1056/NEJMra1806949

Sun, X., Liu, X., Wang, Y., & Yuan, F. (2019). The effects of public subsidies on emerging industry: An agent-based model of the electric vehicle industry. *Technological Forecasting and Social Change*, *140*, 281–295. https://doi.org/10.1016/j.techfore.2018.12.013

Tarricone, R., Petracca, F., Cucciniello, M., & Ciani, O. (2022). Recommendations for developing a lifecycle, multidimensional assessment framework for mobile medical apps. *Health Economics*, hec.4505. https://doi.org/10.1002/hec.4505

Tarricone, R., Petracca, F., & Weller, H.-M. (2024). Towards harmonizing assessment and reimbursement of digital medical devices in the EU through mutual learning. *npj Digital Medicine*, *7*(1), 268. https://doi.org/10.1038/s41746-024-01263-w

Techniker Krankenkasse. (2022). *DiGA-Report 2022* (tech. rep.). Techniker Krankenkasse.

Van Den Berg, M. J., Van Loenen, T., & Westert, G. P. (2016). Accessible and continuous primary care may help reduce rates of emergency department use. An international survey in 34 countries. *Family Practice*, *33*(1), 42–50. https://doi.org/10.1093/fampra/cmv082

Appendix

Table A.1: Descriptive Statistics

Language	Total Apps	Average Rating	Share Apps with Rating	Share Apps with Reviews	Share Apps with Price	Share Apps with In-App-Purchases
Arabic	2249	1.00	0.23	0.35	0.05	0.95
Chinese	19807	0.32	0.08	0.14	0.03	0.98
Czech	2007	0.56	0.15	0.24	0.09	0.92
Danish	2438	0.65	0.17	0.25	0.11	0.92
Dutch	4743	0.45	0.12	0.20	0.08	0.92
English	76184	0.41	0.11	0.20	0.10	0.94
French	8562	0.68	0.17	0.27	0.09	0.91
German	9504	0.56	0.15	0.24	0.10	0.91
Greek	2416	0.39	0.10	0.18	0.05	0.95
Indonesian	1422	0.57	0.15	0.24	0.08	0.94
Italian	6667	0.51	0.14	0.22	0.11	0.92
Japanese	6297	0.63	0.17	0.24	0.09	0.93
Korean	4125	0.75	0.18	0.27	0.09	0.95
Norwegian	1876	0.59	0.16	0.25	0.09	0.93
Polish	3860	0.42	0.11	0.19	0.06	0.95
Portuguese	6279	0.66	0.16	0.26	0.07	0.95
Russian	6703	0.65	0.16	0.26	0.10	0.94
Spanish	9562	0.65	0.17	0.27	0.09	0.91
Swedish	2458	0.60	0.16	0.25	0.09	0.91
Turkish	3486	0.44	0.12	0.20	0.06	0.96
Vietnamese	1186	0.65	0.17	0.27	0.08	0.94

Notes: Descriptive statistics for January 2018. The table displays the number of medical treatment apps and other characteristics per language group. One app can be attributed to several languages if the app is available in several languages.

Table A.2: Donor Weights

Donors (Languages)	Donor Weights	Donors (Languages)	Donor Weights
Spanish	0.37	Lithuanian	0.004
Catalan	0.149	Malay	0.004
Danish	0.124	Thai	0.004
Japanese	0.087	Ukrainian	0.004
English	0.054	Urdu	0.004
French	0.02	Azerbaijani	0.003
Italian	0.016	Bulgarian	0.003
Polish	0.013	Croatian	0.002
Chinese	0.011	Northern Sami	0.002
Indonesian	0.011	Persian	0.002
Dutch	0.01	Tibetan	0.002
Turkish	0.008	Bengali	0.001
Czech	0.007	Burmese	0.001
Hungarian	0.007	Filipino	0.001
Norwegian	0.007	Hindi	0.001
Portuguese	0.007	Kazakh	0.001
Greek	0.006	Marathi	0.001
Hebrew	0.006	Serbian	0.001
Korean	0.006	Slovenian	0.001
Russian	0.006	Tamil	0.001
Slovak	0.006	Telugu	0.001
Vietnamese	0.006	Cambodian	0
Finnish	0.005	Estonian	0
Romanian	0.005	Malayalam	0
Arabic	0.004	Swedish	0
Latvian	0.004		

Notes: The table displays all donors (languages) in the donor pool and their respective weight assigned in the SC estimation. This weight refelcts the contribution share of a language to the synthetic German comparison group.

Table A.3: Event Study

Relative Time to Treatment	Coefficient	Standard Error	P-value	CI 95 Percent	
-19:German	-187.79	80.89	0.02	-350.10	-25.48
-18:German	160.25	93.74	0.09	-27.85	348.35
-17:German	186.81	87.05	0.04	12.12	361.49
-16:German	212.29	90.55	0.02	30.59	393.99
-15:German	223.90	102.55	0.03	18.11	429.69
-14:German	8.23	113.23	0.94	-218.99	235.45
-13:German	18.46	72.36	0.80	-126.74	163.66
-12:German	-7.88	66.99	0.91	-142.32	126.55
-11:German	8.85	52.13	0.87	-95.77	113.46
-10:German	27.33	32.34	0.40	-37.57	92.22
-9:German	98.42	35.88	0.01	26.42	170.42
-8:German	95.38	28.55	0.00	38.09	152.68
-7:German	-110.00	19.53	0.00	-149.20	-70.80
-6:German	-183.87	18.62	0.00	-221.22	-146.51
-5:German	-141.71	15.42	0.00	-172.65	-110.77
-4:German	-142.44	15.15	0.00	-172.83	-112.05
-3:German	-51.17	7.30	0.00	-65.82	-36.53
-2:German	4.19	3.62	0.25	-3.08	11.46
-1:German	0.00	0.00	0.00	0.00	0.00
0:German	11.87	9.95	0.24	-8.11	31.84
1:German	99.77	6.34	0.00	87.06	112.48
2:German	126.81	10.14	0.00	106.47	147.15
3:German	230.88	15.25	0.00	200.29	261.48
4:German	185.19	20.73	0.00	143.60	226.79
5:German	228.27	23.71	0.00	180.70	275.84
6:German	269.77	27.88	0.00	213.83	325.71
7:German	313.35	31.34	0.00	250.45	376.24
8:German	476.58	38.95	0.00	398.41	554.74
9:German	618.23	41.72	0.00	534.51	701.95
10:German	771.17	58.41	0.00	653.97	888.37
11:German	849.48	61.61	0.00	725.85	973.11
12:German	971.85	68.96	0.00	833.46	1110.23
13:German	1126.81	82.22	0.00	961.82	1291.80
14:German	1163.77	85.25	0.00	992.71	1334.83
15:German	1220.06	94.51	0.00	1030.41	1409.71
16:German	1274.29	103.91	0.00	1065.78	1482.80
17:German	1388.37	114.15	0.00	1159.31	1617.42
18:German	1424.79	122.35	0.00	1179.28	1670.29
19:German	1411.90	118.90	0.00	1173.32	1650.49
20:German	1428.00	119.13	0.00	1188.95	1667.05
21:German	1441.88	107.64	0.00	1225.88	1657.89
22:German	1506.71	126.04	0.00	1253.80	1759.63
23:German	1635.13	138.13	0.00	1357.97	1912.30
24:German	1730.02	151.43	0.00	1426.15	2033.89
25:German	1795.48	155.50	0.00	1483.45	2107.51

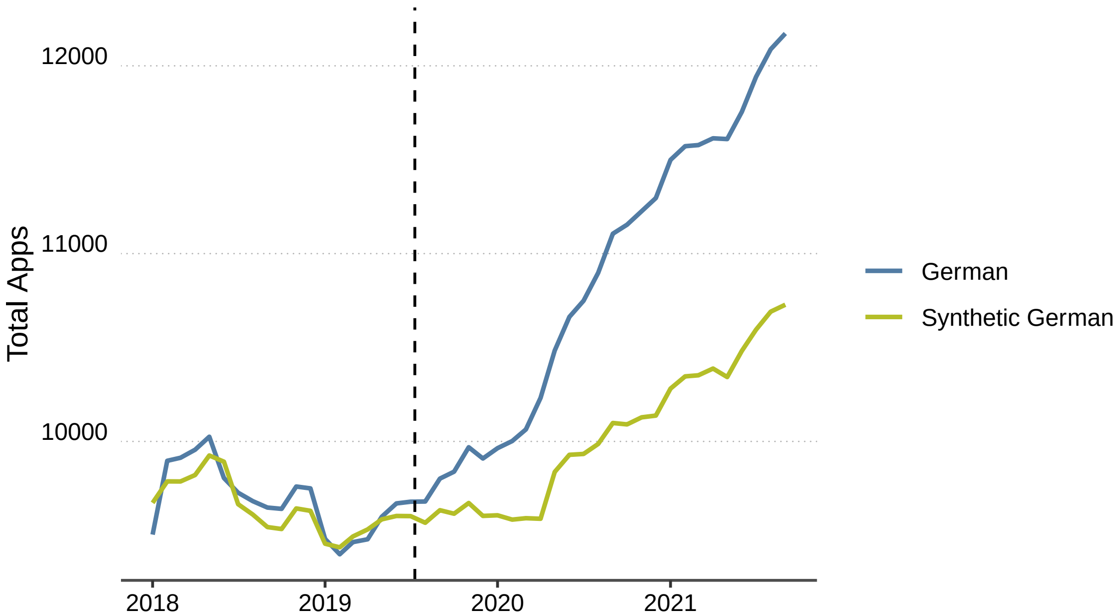
Notes: This table displays event study results. The dependent variable is the number of medical treatment apps. The time period 0 is August 2019, the treatment is in July 2019.

Table A.4: Balance Table

	German	Synthetic German	Sample Mean	Predictor Weight
Average Rating	0.76	0.75	0.90	0.02
Share Apps with Rating	0.19	0.19	0.23	0.11
Average Number of Reviews	121.62	275.10	1388.56	0.01
Share Apps with Reviews	0.27	0.27	0.30	0.20
Share Apps with Price	0.08	0.08	0.07	0.07
Av. Number of Apps per Developer	113.54	113.69	97.61	0.07
Share Apps with In-App-Purchases	0.94	0.94	0.95	0.44
Av. Number of Apps per Developer Medical	3.45	3.46	3.93	0.07
Total Apps	9597.27	9550.83	3630.30	0.01

Notes: Balance table for the synthetic control group and German for all predictors in the SC estimation. The third column displays the sample mean without the weighting of the SC estimation. The last column displays the predictor weight in the SC estimation.

Figure A.1: Synthetic Control Method - German & Synthetic



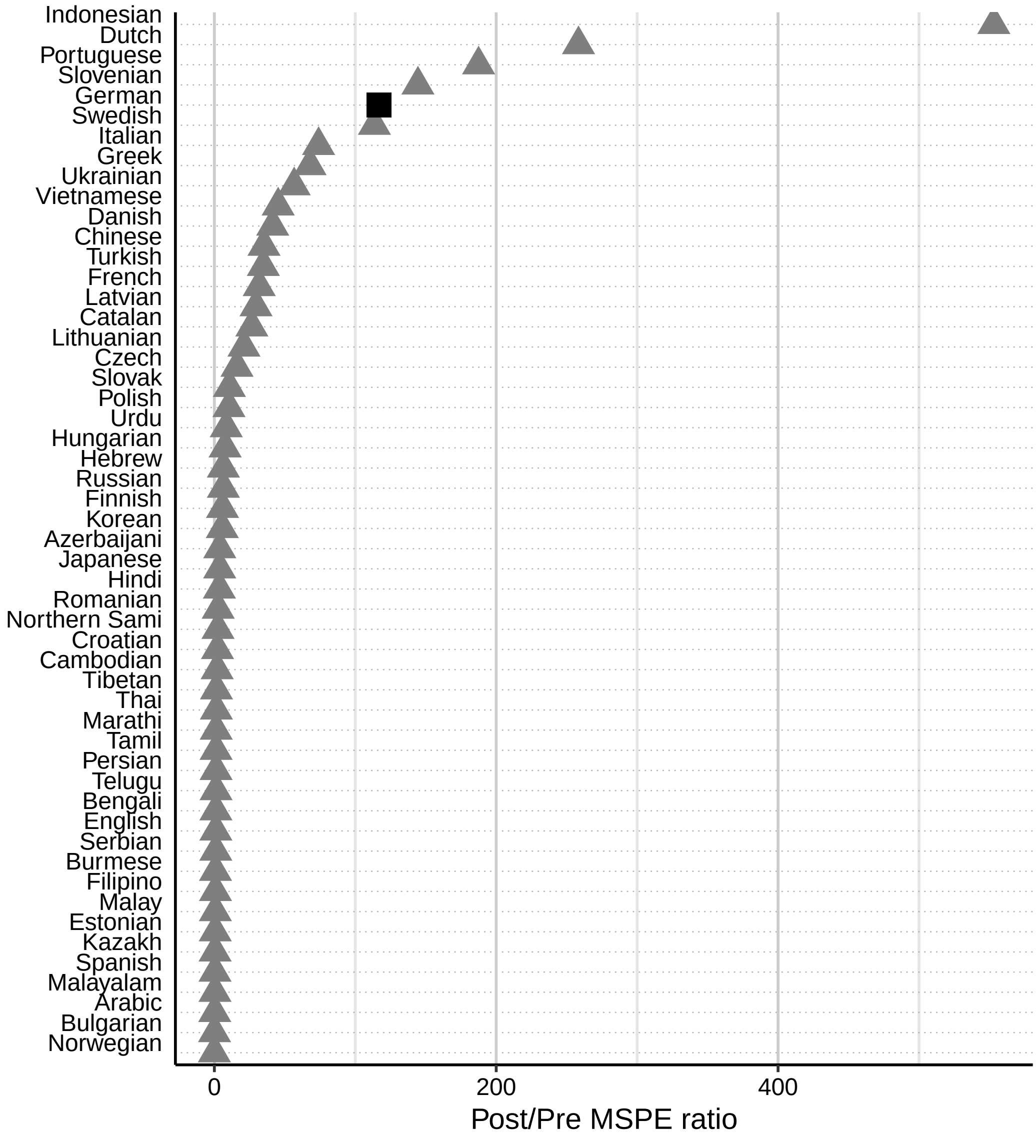
Notes: The graph displays the result of the SC estimation from January 2018 until September 2021. The pre-intervention period is divided into a training and a validation period. The pre-treatment period up to 2018 serves as training period.

Table A.5: Synthetic Control Method - Treatment Effects by Month

Month	Treatment Effect	Total Apps German	Total Apps Synthetic German
August 2019	113.07	9680	9566.93
September 2019	168.10	9802	9633.90
Oktober 2019	223.87	9839	9615.13
November 2019	296.78	9969	9672.22
December 2019	305.62	9909	9603.38
January 2020	357.28	9964	9606.72
February 2020	418.31	10002	9583.69
March 2020	472.71	10064	9591.29
April 2020	643.83	10232	9588.17
May 2020	646.43	10484	9837.57
June 2020	734.17	10663	9928.83
July 2020	815.73	10749	9933.27
August 2020	911.39	10898	9986.61
September 2020	1007.73	11106	10098.27
October 2020	1063.35	11154	10090.65
November 2020	1097.64	11226	10128.36
December 2020	1158.38	11296	10137.62
January 2021	1217.73	11499	10281.27
February 2021	1225.64	11572	10346.36
March 2021	1225.81	11578	10352.19
April 2021	1225.95	11614	10388.05
May 2021	1267.11	11610	10342.89
June 2021	1272.89	11755	10482.11
July 2021	1346.08	11941	10594.92
August 2021	1397.07	12088	10690.93
September 2021	1444.05	12172	10727.95

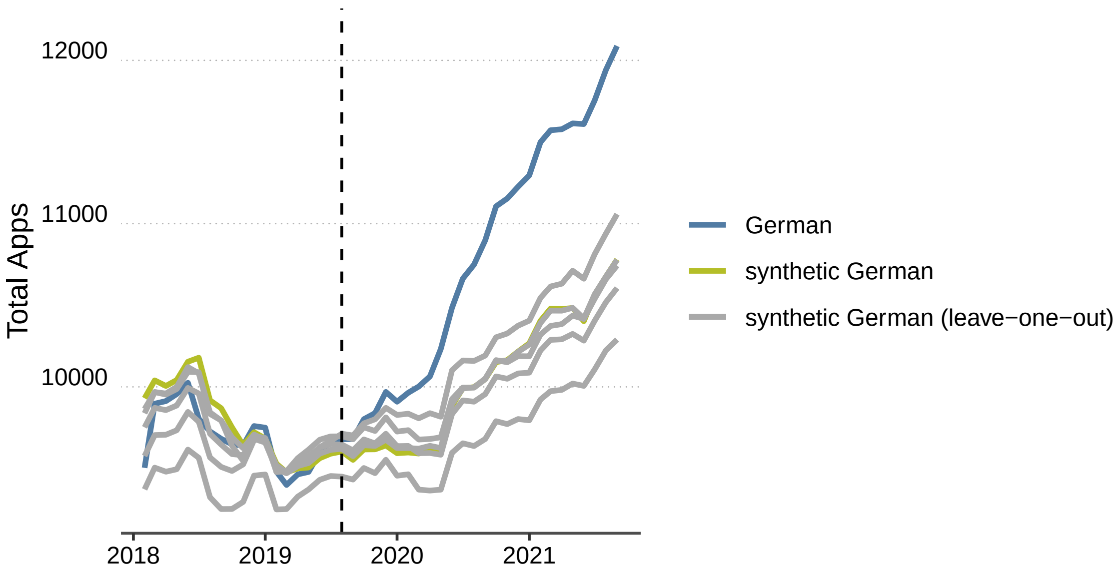
Notes: This table displays the unaveraged monthly treatment effects. The last column displays the total number of apps from the synthetic control group, hence the predicted number of German apps in the market without the introduction of the DiGA scheme, the counterfactual German number of medical treatment apps.

Figure A.2: Synthetic Control Method - Placebo Estimation MSPE Ratio



Notes: The graph displays the result of the SC estimation from January 2018 until September 2021. Additionally, the graph shows placebo estimations assuming groups other than German were treated. The graph shows the ratio of the pre/post mean squared predicted error of placebo runs assigning the treatment to a different group. The larger the ratio, the larger the identified effect.

Figure A.3: Synthetic Control Method - Leave-One-Out



Notes: The graph displays the result of the SC estimation from January 2018 until September 2021. The estimation for the SC from the main model is displayed in green. The grey lines show repetitions of the main model whereby the donor pool is reduced by one language.

Table A.6: Donor Weights

Donors (Languages)	Donor Weights
English	0.07
Spanish	0.07
French	0.06
Italian	0.06
Portuguese	0.06
Russian	0.04
Dutch	0.04
Latvian	0.04
Lithuanian	0.03
Greek	0.03
Korean	0.03
Norwegian	0.03
Japanese	0.03
Serbian	0.02
Turkish	0.02
Hindi	0.01
Bulgarian	0.01
Burmese	0.01
Northern Sami	0.01
Cambodian	0.01
Estonian	0.01
Persian	0.01
Filipino	0.01
Slovenian	0.01
Polish	0.01
Malayalam	0.01
Croatian	0.01
Kazakh	0.01
Urdu	0.01
Telugu	0.01
Marathi	0.01
Bengali	0.01
Tamil	0.01
Vietnamese	0.01
Azerbaijani	0.01
Thai	0.01

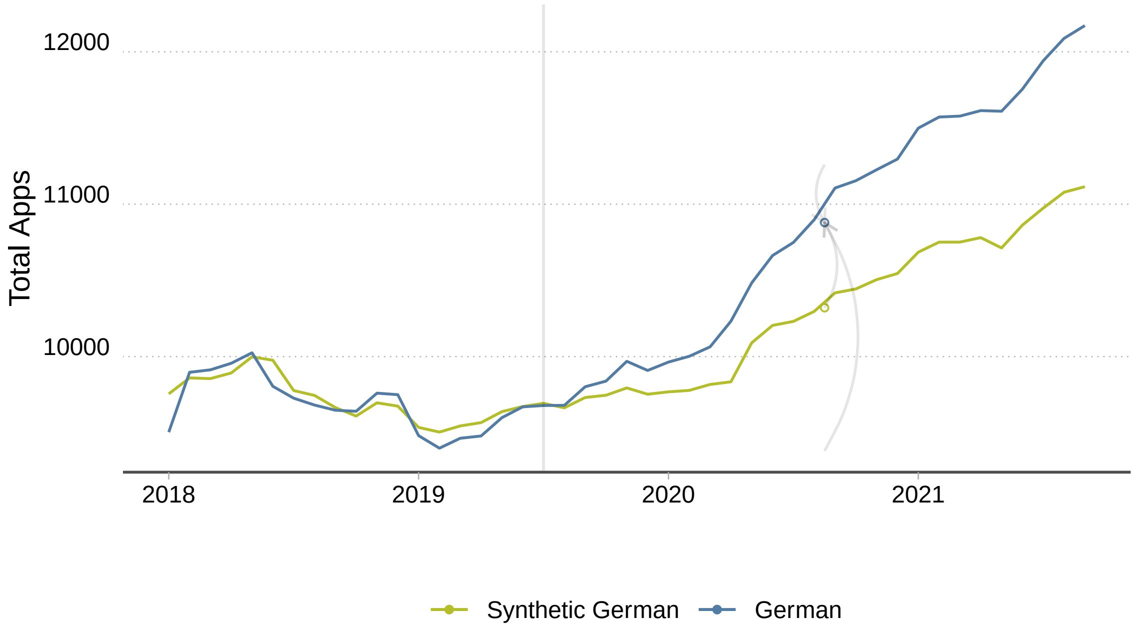
Notes: This table displays the weights of the language groups in the donor pool in the SDID estimations.

Table A.7: Synthetic Difference-in-Differences - Treatment Effects by Month

Month	Treatment Effect	Total Apps German	Total Apps Synthetic German
August 2019	16.22	9680	9663.78
September 2019	70.42	9802	9731.58
Oktober 2019	92.69	9839	9746.31
November 2019	174.66	9969	9794.34
December 2019	155.95	9909	9753.05
January 2020	195.45	9964	9768.55
February 2020	223.94	10002	9778.06
March 2020	246.88	10064	9817.12
April 2020	397.51	10232	9834.49
May 2020	393.54	10484	10090.46
June 2020	458.41	10663	10204.59
July 2020	518.20	10749	10230.80
August 2020	601.74	10898	10296.26
September 2020	688.08	11106	10417.92
October 2020	710.08	11154	10443.92
November 2020	721.11	11226	10504.89
December 2020	750.72	11296	10545.28
January 2021	813.98	11499	10685.02
February 2021	820.93	11572	10751.07
March 2021	826.78	11578	10751.22
April 2021	833.41	11614	10780.59
May 2021	897.16	11610	10712.84
June 2021	892.32	11755	10862.68
July 2021	966.81	11941	10974.19
August 2021	1009.67	12088	11078.33
September 2021	1057.02	12172	11114.98

Notes: This table displays the unaveraged monthly treatment effects. The last column displays the total number of apps from the synthetic control group, hence the predicted number of German apps without introduction of the DiGA scheme.

Figure A.4: Synthetic Difference-in-Differences - Placebo Inference



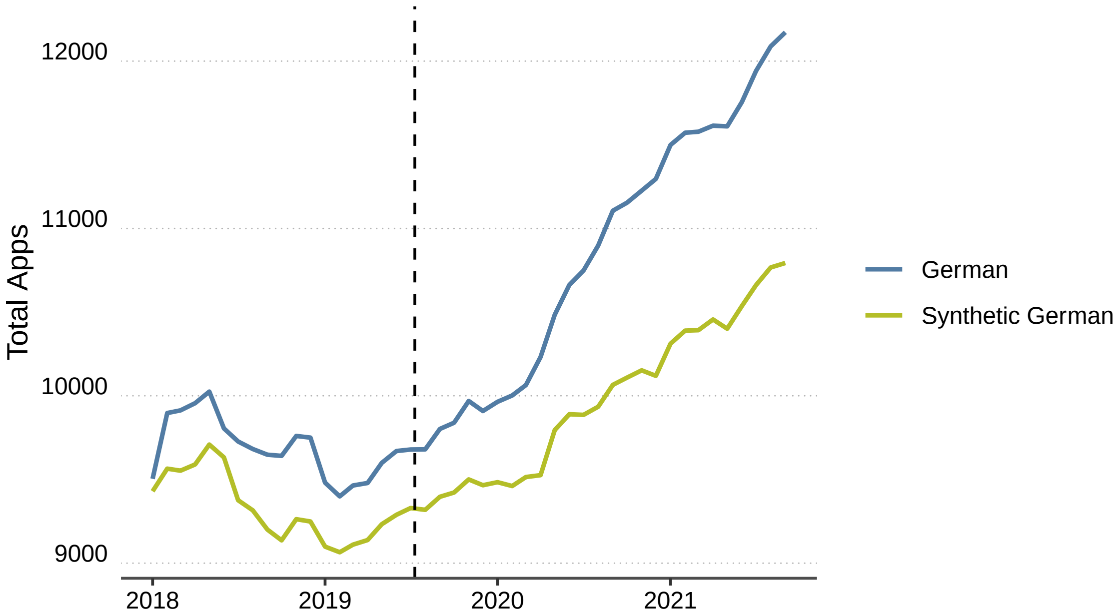
Notes: The graph displays the result of the SDID estimation from January 2018 until September 2021. The effects displayed are the same as in the main model. For assessing the pre-treatment match, the two lines are overlaid by adding the constant to the synthetic control. The small grey arrow indicates the point estimate of the average treatment effect on the treated. The two large grey arrows indicate the 95% confidence interval.

Table A.8: Balance Table - Predictor Set 2

	German	Synthetic German	Sample Mean	Predictor Weight
Business	893.91	787.07	700.10	0.01
Education	419.64	436.83	436.43	0.00
Entertainment	214.82	219.25	211.21	0.01
Finance	112.46	139.60	136.81	0.00
FoodDrink	222.64	232.44	221.26	0.22
Games	513.91	636.93	629.40	0.00
HealthFitness	351.54	319.71	270.89	0.01
Lifestyle	541.46	521.35	509.30	0.09
Music	92.18	104.36	98.79	0.08
Navigation	60.36	54.04	48.47	0.04
PhotoVideo	121.46	117.38	103.64	0.01
Productivity	262.54	218.28	205.29	0.00
SocialNetworking	183.00	169.53	158.06	0.05
Sports	127.36	94.53	90.00	0.01
Travel	269.46	232.39	198.34	0.00
Utilities	522.82	549.17	514.26	0.00
Weather	9.27	10.12	9.85	0.00
Average Rating	0.76	0.76	0.73	0.02
Share Apps with Rating	0.19	0.19	0.18	0.13
Average Number of Reviews	121.62	170.77	218.95	0.00
Share Apps with Reviews	0.27	0.27	0.26	0.00
Share Apps with Price	0.08	0.08	0.06	0.16
Av. Number of Apps per Developer	113.54	121.54	128.27	0.00
Share Apps with In-App-Purchases	0.94	0.95	0.95	0.00
Av. Number of Apps per Developer Medical	3.45	3.46	3.56	0.11
Total Apps	9597.27	9191.94	8472.98	0.04

Notes: Balance table for the synthetic control group and German for all predictors in the SC estimation. The third column displays the sample mean without the weighting of the SC estimation. The last column displays the predictor weight in the SC estimation. This model uses an alternative set of predictors including those from the main model plus the total number of apps in other genres in the App Store.

Figure A.5: Synthetic Control Method - Predictor Set 2



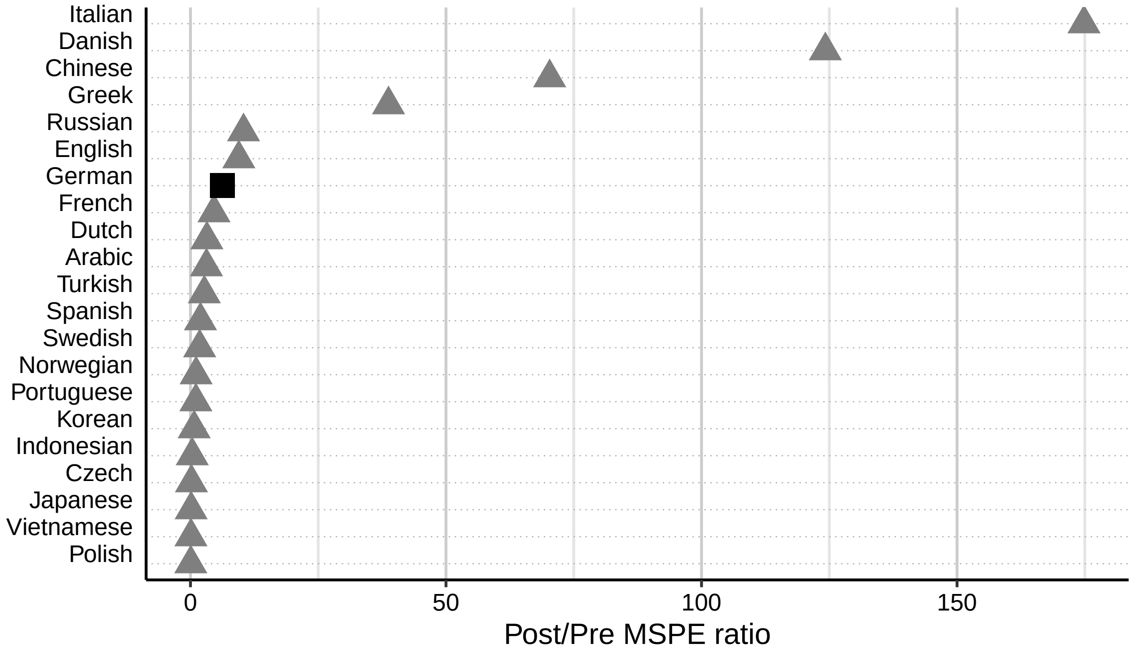
Notes: The graph displays the result of the SC estimation from January 2018 until September 2021. The pre-intervention period is divided into a training and a validation period.The pre-treatment period up to 2018 serves as a training period. The model is estimated with an alternative set of covariates including those from the main model plus the total number of apps in other genres in the App Store.

Table A.9: Synthetic Control Method - Treatment Effects by Month - Predictor Set 2

Month	Treatment Effect	Total Apps German	Total Apps Synthetic German
August 2019	361.41	9680	9318.59
September 2019	405.58	9802	9396.42
Oktober 2019	416.85	9839	9422.15
November 2019	468.71	9969	9500.29
December 2019	443.45	9909	9465.55
January 2020	480.22	9964	9483.78
February 2020	541.13	10002	9460.87
March 2020	549.56	10064	9514.44
April 2020	706.25	10232	9525.75
May 2020	689.09	10484	9794.91
June 2020	773.09	10663	9889.91
July 2020	862.91	10749	9886.09
August 2020	963.19	10898	9934.81
September 2020	1040.59	11106	10065.41
October 2020	1045.30	11154	10108.70
November 2020	1073.84	11226	10152.16
December 2020	1176.73	11296	10119.27
January 2021	1187.38	11499	10311.62
February 2021	1182.54	11572	10389.46
March 2021	1185.79	11578	10392.21
April 2021	1157.79	11614	10456.21
May 2021	1209.05	11610	10400.95
June 2021	1218.89	11755	10536.11
July 2021	1279.52	11941	10661.48
August 2021	1321.18	12088	10766.82
September 2021	1378.34	12172	10793.66

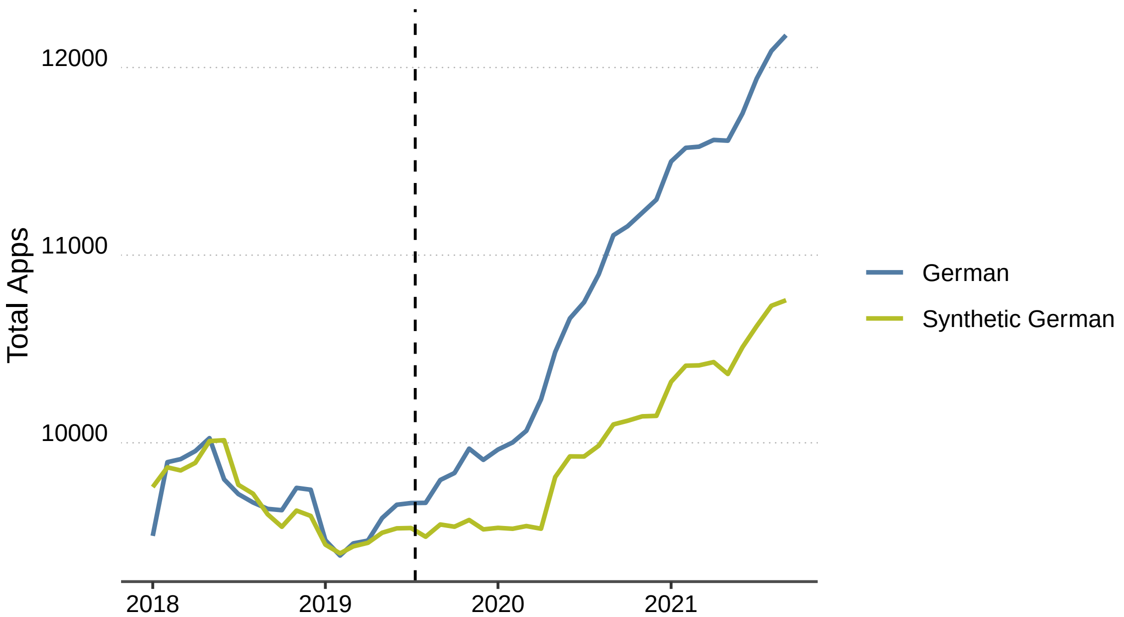
Notes: This table displays the unaveraged monthly treatment effects. The last column displays the total number of medical treatment apps from the synthetic control group, hence the predicted number of German apps without introduction of the DiGA scheme. The model is estimated with an alternative set of covariates.

Figure A.6: Synthetic Control Method - Placebo Estimation - Predictor Set 2



Notes: The graph displays the result of the SC estimation from January 2018 until September 2021. The model is estimated with an alternative set of covariates. The graph shows the ratio of the pre/post mean squared predicted error of placebo runs assigning the treatment to a different group. The larger the ratio, the larger the identified effect.

Figure A.7: Synthetic Control Method - Top 20 Donor Pool



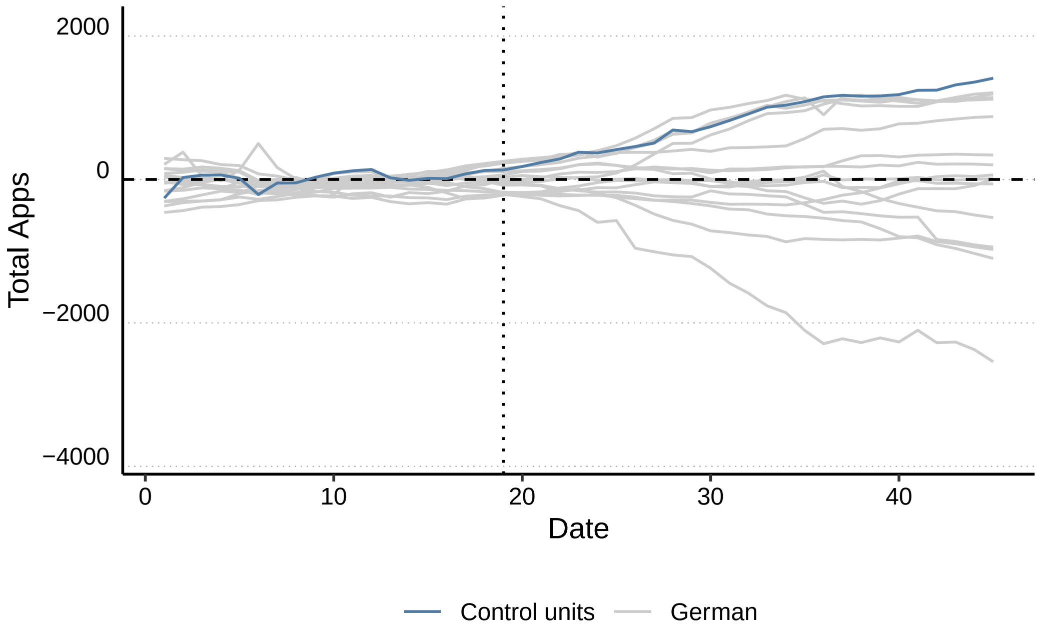
Notes: The graph displays the result of the SC estimation from January 2018 until September 2021. The pre-intervention period is divided into a training and a validation period. The pre-treatment period up to 2018 serves as training period.

Table A.10: Synthetic Control Method - Treatment Effects by Month - Top 20 Donor Pool

Month	Treatment Effect	Total Apps German	Total Apps Synthetic German
August 2019	180.62	9680	9499.38
September 2019	237.25	9802	9564.75
Oktober 2019	286.13	9839	9552.87
November 2019	380.19	9969	9588.81
December 2019	370.27	9909	9538.73
January 2020	417.15	9964	9546.85
February 2020	459.93	10002	9542.07
March 2020	507.63	10064	9556.37
April 2020	689.88	10232	9542.12
May 2020	666.45	10484	9817.55
June 2020	735.36	10663	9927.64
July 2020	822.09	10749	9926.91
August 2020	913.29	10898	9984.71
September 2020	1008.10	11106	10097.90
October 2020	1036.63	11154	10117.37
November 2020	1084.96	11226	10141.04
December 2020	1152.50	11296	10143.50
January 2021	1173.94	11499	10325.06
February 2021	1161.11	11572	10410.89
March 2021	1165.34	11578	10412.66
April 2021	1183.81	11614	10430.19
May 2021	1243.30	11610	10366.70
June 2021	1245.27	11755	10509.73
July 2021	1319.15	11941	10621.85
August 2021	1357.72	12088	10730.28
September 2021	1412.12	12172	10759.88

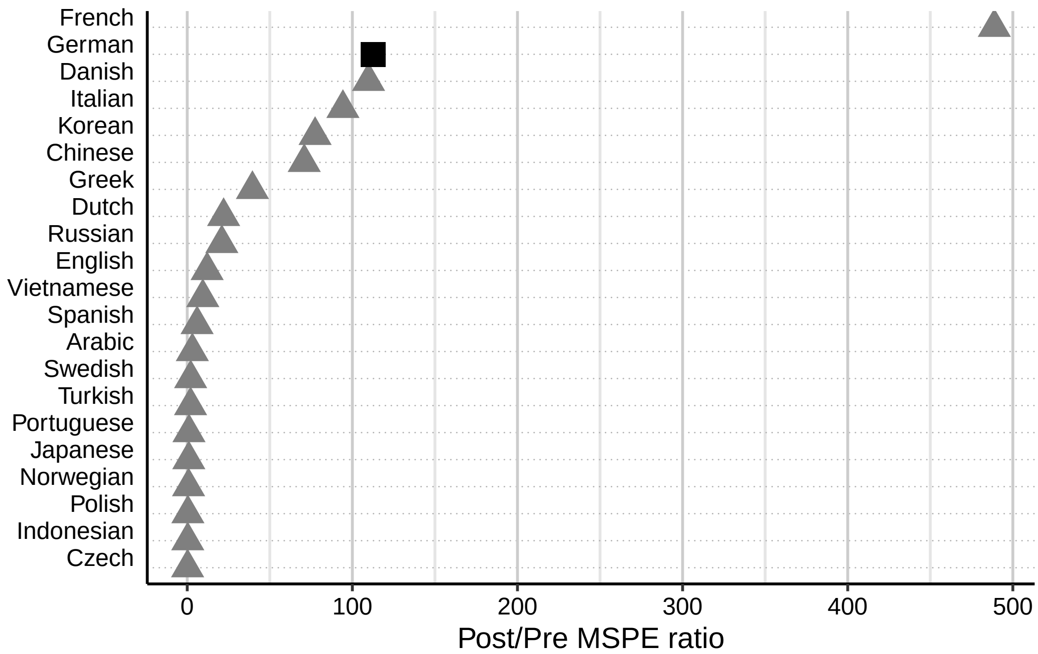
Notes: This table displays the unaveraged monthly treatment effects. The last column displays the total number of apps from the synthetic control group, hence the predicted number of German apps in the market without the introduction of the DiGA scheme, the counterfactual German number of apps.

Figure A.8: Synthetic Control Method - Placebo Estimation Sample - Top 20 Donor Pool



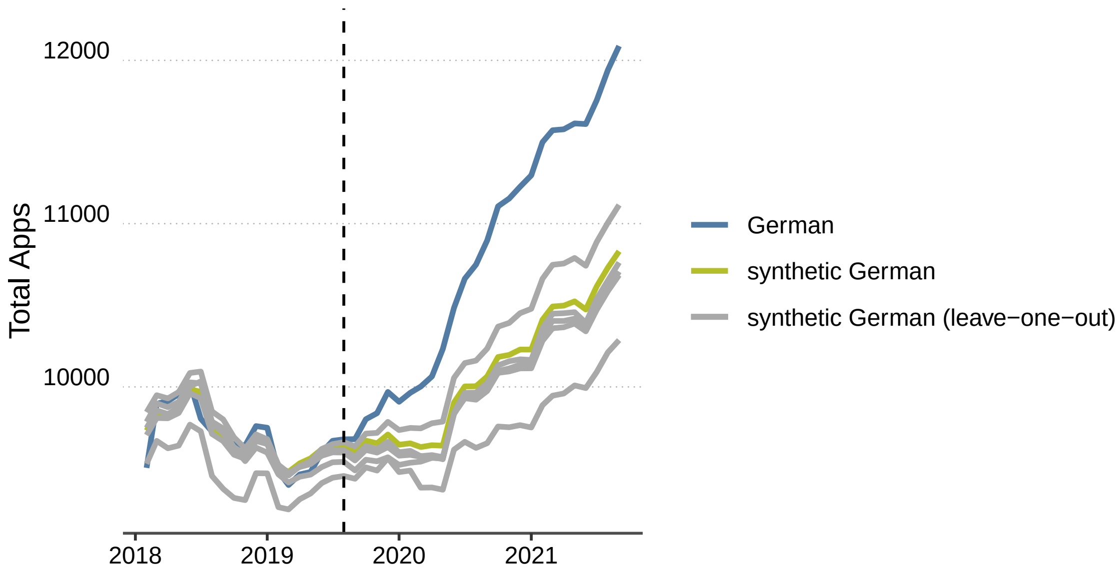
Notes: The graph displays the result of the SC method estimation from January 2018 until September 2021. Additionally, the graph shows placebo estimation results assuming other groups than German were treated.

Figure A.9: Synthetic Control Method - Placebo Estimation MSPE Ratio - Top 20 Donor Pool



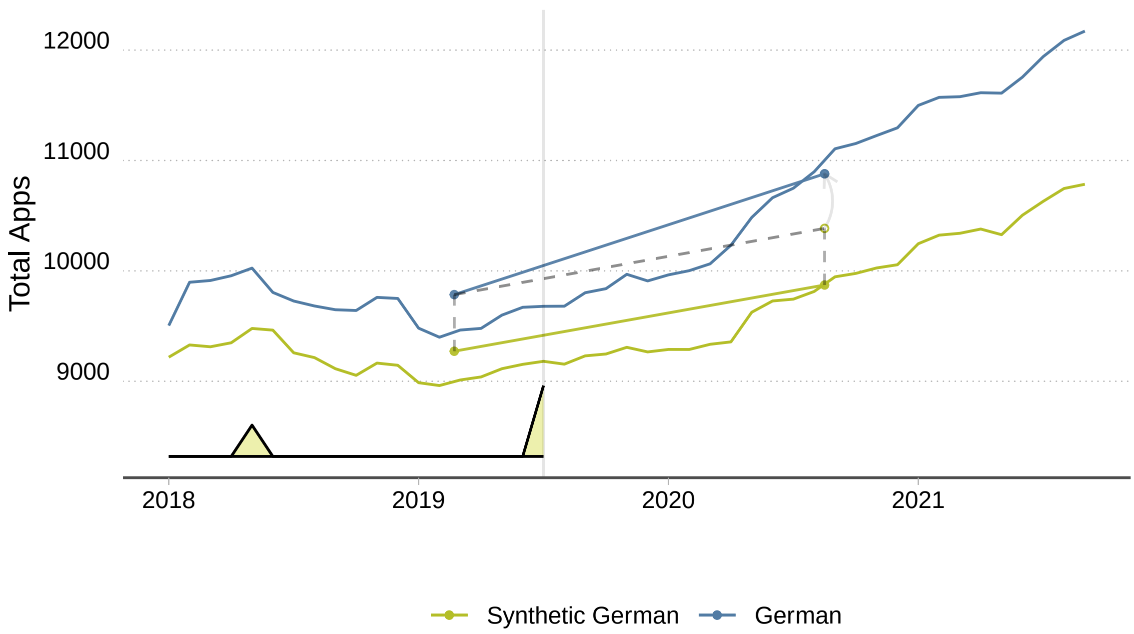
Notes: The graph displays the result of a synthetic control method estimation from January 2018 until September 2021. Additionally, the graph shows placebo estimation results assuming other groups than German were treated. The graph shows the ratio of the pre/post mean squared predicted error of placebo runs assigning the treatment to a different group. The larger the ratio, the larger the identified effect.

Figure A.10: Synthetic Control Method - Leave-One-Out - Top 20 Donor Pool



Notes: The graph displays the result of the SC method estimation from January 2018 until September 2021. The estimation for the synthetic control from the main model is displayed in green. The grey lines show repetitions of the main model whereby the donor pool is reduced by one language.

Figure A.11: Synthetic Difference-in-Differences - Top 20 Donor Pool



Notes: The graph displays the result of the SDID estimation from January 2018 until September 2021. The red triangles indicate the weights on pre-treatment periods, the light grey arrow indicates the average treatment effect on the treated. The parallelogram illustrates the part of the difference between German and synthetic German which can be explained by fixed effects.

Figure A.12: Event Study - Top 20 Donor Pool

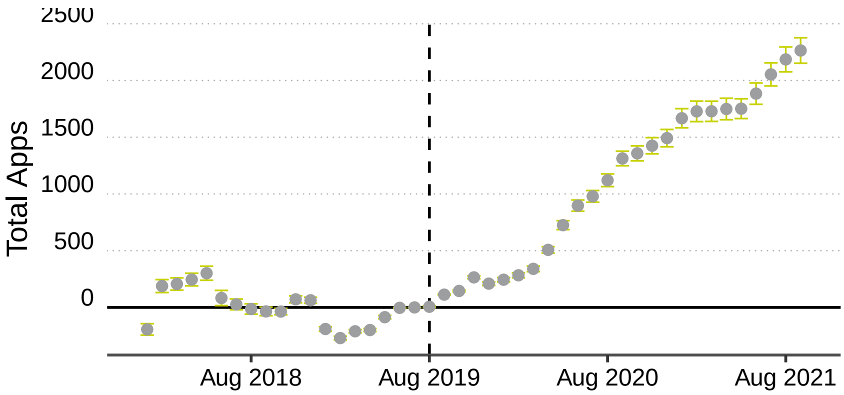
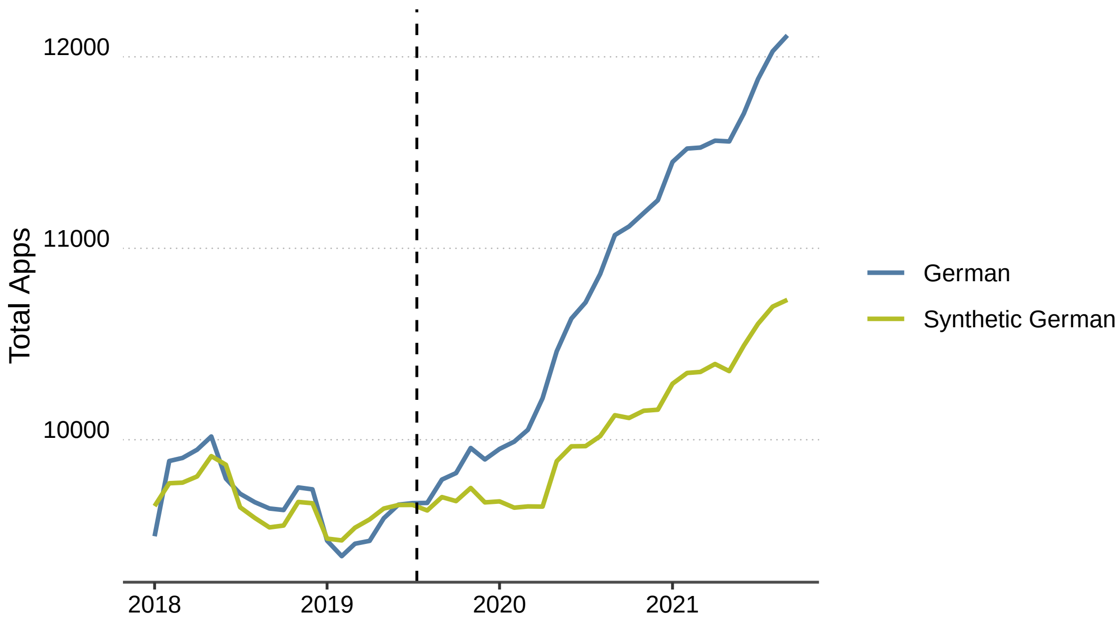
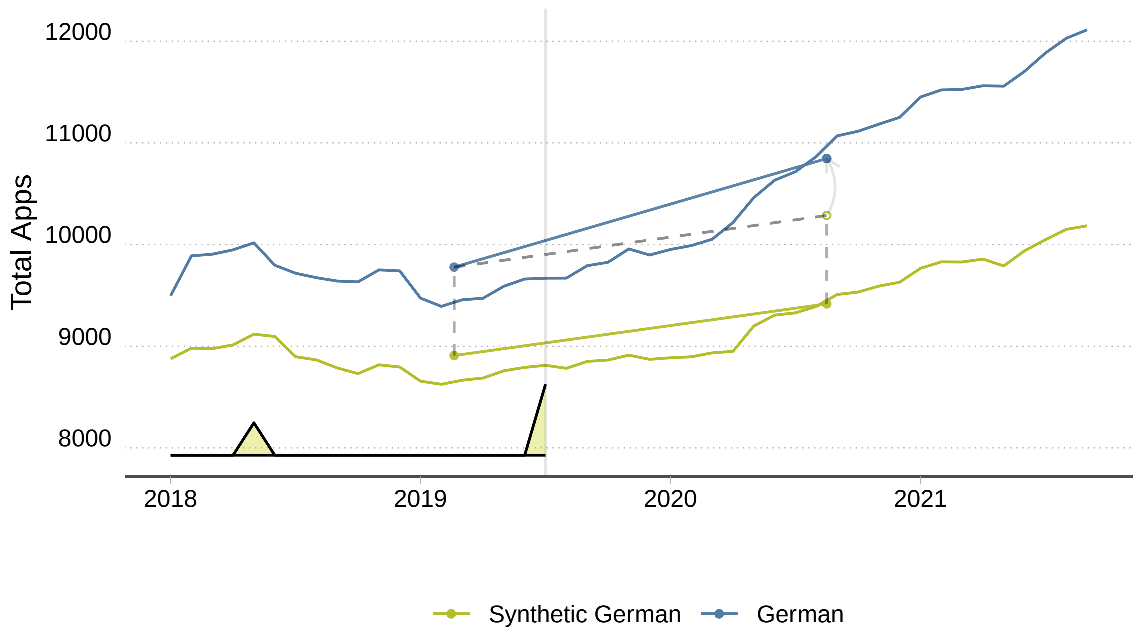


Figure A.13: Synthetic Control Method - Excluding Covid Apps



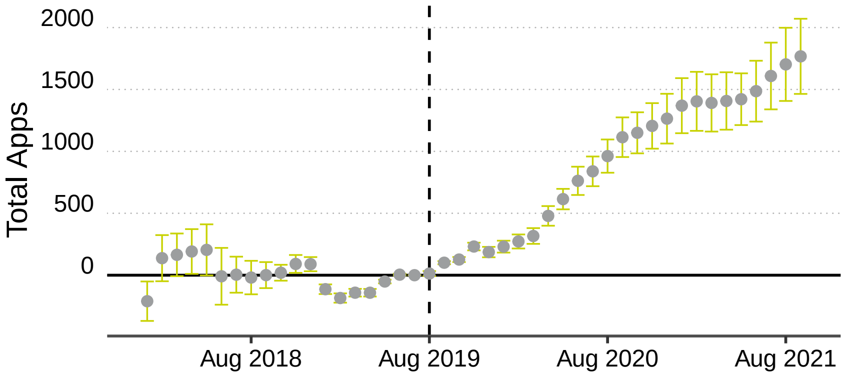
Notes: The graph displays the result of the SC estimation from January 2018 until September 2021. The pre-intervention period is divided into a training and a validation period. The pre-treatment period up to 2018 serves as training period. All apps related to Covid are excluded from the sample.

Figure A.14: Synthetic Difference-in-Differences - Excluding Covid Apps



Notes: The graph displays the result of the SDID estimation from January 2018 until September 2021. The red triangles indicate the weights on pre-treatment periods, the light grey arrow indicates the average treatment effect on the treated. The parallelogram illustrates the part of the difference between German and synthetic German which can be explained by fixed effects. All apps related to Covid are excluded from the sample.

Figure A.15: Event Study - Excluding Covid Apps



Notes: The graph displays the result of the event study estimation from January 2018 until September 2021. All apps related to Covid are excluded from the sample.

Download ZEW Discussion Papers:

https://www.zew.de/en/publications/zew-discussion-papers

or see:

https://www.ssrn.com/link/ZEW-Ctr-Euro-Econ-Research.html
https://ideas.repec.org/s/zbw/zewdip.html

// IMPRINT

ZEW – Leibniz-Zentrum für Europäische Wirtschaftsforschung GmbH Mannheim

ZEW – Leibniz Centre for European Economic Research

L 7,1 · 68161 Mannheim · Germany
Phone +49 621 1235-01
info@zew.de · zew.de