

NBER WORKING PAPER SERIES

THE SHORT-RUN EFFECTS OF GDPR ON TECHNOLOGY VENTURE INVESTMENT

Jian Jia
Ginger Zhe Jin
Liad Wagman

Working Paper 25248
<http://www.nber.org/papers/w25248>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2018

Jia gratefully acknowledges support from the NET Institute. Wagman gratefully acknowledges support from Data Catalyst and from the NET Institute. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2018 by Jian Jia, Ginger Zhe Jin, and Liad Wagman. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

The Short-Run Effects of GDPR on Technology Venture Investment
Jian Jia, Ginger Zhe Jin, and Liad Wagman
NBER Working Paper No. 25248
November 2018
JEL No. D43,D8,L13,L15,L5

ABSTRACT

The General Data Protection Regulation (GDPR) came into effect in the European Union in May 2018. We study its short-run impact on investment in new and emerging technology firms. Our findings indicate negative post-GDPR effects on EU ventures, relative to their US counterparts. The negative effects manifest in the overall dollar amounts raised across funding deals, the number of deals, and the dollar amount raised per individual deal.

Jian Jia
Illinois Institute of Technology
565 W Adams, 4th Flr
Chicago, IL 60661
jjia5@hawk.iit.edu

Liad Wagman
Illinois Institute of Technology
Main Campus, IIT Tower, 18C6-2
Downtown Campus, Room 452
lwagman@stuart.iit.edu

Ginger Zhe Jin
University of Maryland
Department of Economics
3115F Tydings Hall
College Park, MD 20742-7211
and NBER
jin@econ.umd.edu

1 Introduction

The rise of big data in the global economy has transformed marketplaces, altering the ways in which firms and consumers interact. Individuals are no longer mere consumers of goods, information and services, but public producers of often valuable data. These data have become key inputs in technology-driven innovation, spanning industry sectors from health, advertising, and security, to e-commerce, transportation, and banking. These data are also key inputs in the matching processes among consumers and firms, as well as between firms and other firms. For individual firms, data is a valuable asset to monetize, especially if the data is personally-identifiable, real-time and of high-frequency. Individuals' traits and attributes, their behaviors and online footprints, their comments and photos, their work and leisure habits, and more, are increasingly regarded as business assets that can be used to target services or offers, to provide relevant advertising, financial offerings, and healthcare, or to trade with other parties.

In an effort to leverage the value inherent in the data created by individuals, new services, companies, and markets are emerging. The services, tools, and products being made possible by the increased availability of data are bearing benefits for data subjects and data holders alike. These benefits include tailor-made recommendations, digital personal assistants, new products and offerings, and easy access to previously thin or unavailable markets. The Federal Trade Commission's 2016 report on big data (FTC, 2016) highlights a number of benefits to underserved populations, including increased educational attainment, access to credit through non-traditional methods, specialized health care for underserved communities, and better access to employment.

Despite those benefits, public concerns over the use of personal data have increased. Recent Pew surveys find that 91% of respondents believe they have lost control over how personal information is collected and used, 61% would like to do more to protect their privacy, and 66% said current laws are insufficient for protecting their privacy and would support

more regulation.¹ These concerns are amplified by recent incidences of data breaches and data misuses, and a lack of regulatory actions after these scandals.

Those concerns are not without merit. The Commission’s earlier report (FTC, 2014) indicates that data brokers collect and store billions of data elements covering nearly every U.S. consumer, with one data broker holding information on more than 1.4 billion consumer transactions and 700 billion data elements, while another broker added more than 3 billion new data points to its database each month. Another report found that 95% of the top 200 free iOS and Android apps exhibit at least one risky behavior including location tracking, access to social networks, and disclosing the user’s personally identifiable information.² The FTC’s 2016 report also highlights possible risks that could result from biases or inaccuracies about certain groups, including more individuals mistakenly denied opportunities based on the actions of others, sensitive information being exposed, existing disparities being reinforced, increased targeting of vulnerable consumers for reasons such as fraud, an increase in prices for goods and services in lower-income communities, and the weakening of consumer choice.

Against this backdrop, the General Data Protection Regulation was adopted in the European Union on April 14, 2016, becoming enforceable two years later on May 25, 2018. The regulation aims to protect data by ‘design and default,’ wherein firms are obligated to handle data according to a set of principles and safeguards. GDPR mandates a higher degree of privacy, data management, and control, requires informed opt-in consent for data collection, and assigns substantial liability risks and penalties for data flow and data processing violations. Under the regulation, firms that process personal information must develop protocols to respond to individual data requests within a month, appoint a data protection officer to oversee compliance activities, audit internal data processes, and take proactive

¹<http://www.pewresearch.org/fact-tank/2018/03/27/americans-complicated-feelings-about-social-media-in-an-era-of-privacy-concerns/>

²<https://www.appthority.com/company/press/press-releases/appthority-exposes-security-and-privacy-risks-behind-top-400-mobile-apps/>

steps to anonymize and secure personal data and minimize its collection. In the event of a data breach, organizations must promptly notify the regulator and affected individuals. The regulation requires that users have the rights to access, correct, and erase their personal data, and imposes fines up to 4% of global revenue for any violation.³

The enactment of GDPR is particularly likely to influence technology firms, given an ever increasing need for the use of data as a core product input. This study presents an analysis of the impact of the rollout of GDPR on new technology venture investment in the EU. Our findings indicate a negative differential effect on EU ventures after the rollout of GDPR relative to their US counterparts. These negative effects manifest in the overall number of financing rounds, the overall dollar amount raised across rounds, and in the dollar amount raised per individual round. Specifically, our findings suggest a \$3.38 million decrease in the aggregate dollars raised by EU ventures per state per crude industry category per week, a 17.6% reduction in the number of weekly venture deals, and a 39.6% decrease in the amount raised in an average deal following the rollout of GDPR.

We then proceed to break down these effects by two crude industry categories and four firm age groups. Somewhat surprisingly, we do not identify particularly different effects for ventures that fall under the healthcare-financial category, despite the existence of arguably stringent regulations that govern data flows in those industries in the US. However, we do find that the negative effects of GDPR on technology investment appear particularly pervasive for nascent, 0-3 year old ventures. We use our results to provide a back-of-the-envelope calculation of a range of job losses that may be incurred by these ventures, which we estimate to be between 3,604 to 29,819 jobs, corresponding to 4.09 to 11.20% jobs created

³On June 28, 2018, California adopted a data regulation law (the California Consumer Privacy Act of 2018, A.B. 375) that is set to take effect on January 1 2020. The law, as currently written, would require that firms provide consumers with the ability to port their profiles to other providers, be informed about what personal data are stored about them, why they are collected, request their deletion, and opt out of their sale. The legislation is still subject to amendments, and it is widely understood that the version that was passed is highly likely to change based on input from stakeholders before its implementation in 2020. On November 1, 2018, Senator Ron Wyden's office began circulating a discussion draft of a bill tentatively-named the "Consumer Data Protection Act," which aims to introduce similar and other protections at the federal level.

by 0-3 year old ventures in our sample.

1.1 Literature Review

The literature that examines the implications of data policies and data regulation is growing (Acquisti et al., 2016, offer a recent survey). Goldfarb and Tucker (2011, 2012) examine the effects of the implementation of the EU Privacy Directive and find some evidence that after the Privacy Directive was passed, advertising effectiveness decreased significantly. They argue that digitization means that privacy policy is now a part of innovation policy, an assertion that is in line with our analysis. More recently, Goldberg et al. (2018) examine the impact of the GDPR on online web traffic, sales, and third-party tracking. Using proprietary datasets, they show that for EU firms, recorded pageviews fall by 7.5%, recorded conversions fall 12.5%, and third-party tracking falls 6.2% after the rollout of GDPR.⁴ They demonstrate that beliefs about local regulatory strictness plays a factor in firms' reactions. Their results are complementary to our findings. In particular, while their focus is on the health and regulatory compliance of both online publishers and e-commerce sites, our focus is on the broader tradeoff between innovation and data regulation, for which specific considerations such as ad tracking are one of many data monetization components. Moreover, rather than focus on existing firms, our study focuses on emerging ventures that are typically much smaller.

In this regard, our analysis more closely maps the theoretical works of Krasteva et al. (2015) and Campbell et al. (2015), who show that compliance costs and data regulation, respectively, can create barriers to entry and may thus hurt innovation. In particular, Campbell et al. (2015) show that though privacy regulation imposes costs on all firms, it is small firms and new firms that are most adversely affected, particularly for goods where

⁴In a related paper, Tucker (2014) uses a randomized field experiment to show that users are more likely to click on personalized ads once they are given more control over their personally identifiable information, a change that was only driven by the change of user perception of privacy control, as the website did not change how advertisers used data to target and personalize ads.

the price mechanism does not mediate the effect, such as the advertising-supported internet. Krasteva et al. (2015) show that as the costs of compliance by small firms increase, more innovations will be developed within established firms.⁵

As far as health-related data regulation, using variations in state medical privacy laws, Miller and Tucker (2009, 2011) show that privacy regulations restricting a hospital's release of patient information significantly reduced the adoption of electronic medical records. This reduction is costly, as a 10 percent increase in the adoption of such systems can reduce infant mortality by 16 deaths per 100,000 births. A related issue is an insurer's access to information about a person's genetic test results and subsequent price discrimination. In the US, most states have banned the use of genetic information by health insurers; however, some theoretical results show that inefficiencies may arise when test information is private relative to when it is public (Hoel and Iversen, 2002), and some empirical findings identify relationships among how consumers are informed, the control they have over their private information, and whether consumers elect to partake in genetic testing and share their information with providers (Miller and Tucker, 2017).

Data regulation has also been studied in financial markets. Kim and Wagman (2015) theoretically show that an opt-in approach for information trade in financial markets can lead to higher prices, and empirically demonstrate their results in the market for mortgages, with indications of higher mortgage rates, lower mortgage underwriting standards, and potential downstream foreclosures. Pagano and Jappelli (1993, 2002) predict that if banks share information about their customers, they would increase lending to safe borrowers, thereby decreasing default rates. Other studies focus on the effects of credit bureaus and creditor rights using data from a cross-section of countries (see, e.g., Djankov et al. 2007; Qian and Strahan 2007). Hertzberg et al. (2011) and Doblas-Madrid and Minetti (2013) analyze micro data to show that the effect of lenders' information sharing is to reduce incidence of

⁵While one may argue that higher compliance costs may have a positive effect on innovation (e.g., Porter, 1991), Jaffe and Palmer (1997) find little evidence that industries' inventive output (as measured by patent applications) is related to compliance costs.

delinquencies and defaults, but lenders may reduce credit in anticipation of other lenders' reaction to negative news.

The aforementioned works largely examine a specific aspect or implication of data regulation, such as advertising, pricing, defaults in financial markets, and the impact on medical effectiveness. Our work is complementary, in that our analysis centers around the effect of data regulation on technology ventures and the nascent firms that data regulation is most likely to affect. We demonstrate that younger firms are particularly susceptible to the consequences of data regulation. Our findings are thus in line with Kortum and Lerner (2000) who show that the industrial innovations that venture capitalists help facilitate is a multiple of the ratio of venture capital to R&D expenditures, as well as with Kerr et al. (2014), who suggest that the bundle of inputs that angel investors provide have a large and significant impact on the success and survival of new ventures.

The remainder of the paper proceeds as follows. Section 2 describes the data and Section 3 presents the overall empirical approach. Section 4 provides results at the aggregate level, and Section 5 gives sub-sample results broken down by crude industry category and firm age. Section 6 discusses implications for employment and Section 7 concludes.

2 Data

Our primary data source is Crunchbase, a platform for tracking information about technology businesses, particularly nascent ventures.⁶ We collect data on all technology-venture related activity in the EU and US from July 2017 to September 2018, including the parameters of venture financing rounds, such as venture information (name, location, operating category, founding date, financing dates, and a range on the number of employees) and funding information (the size of the funding round in USD, the date each round was announced, the type of financing round such as seed, Series A, Series B, and the number, names, locations, and

⁶For recent activity in the academic literature that pertains to this data source, see Hochberg (2016), Kaplan and Lerner (2017), Lerner et al. (2018), and Chatterji et al. (2018).

types of the participating investors). Each venture in the dataset is also tagged with a few relevant product keywords (e.g., ‘software’, ‘e-commerce’, ‘finance’, etc).

We treat each funding round observed as a ‘deal’ event, tallying deals per week in each crude industry category and in each US state or EU member state (henceforth, state).⁷ Since each deal is tagged with product keywords, we further group ventures into either healthcare-financial or other technology. We choose to group the data into these two crude categories, partly because healthcare and finance are subject to industry-specific regulations in the US,⁸ thus comparing them against other technology firms allows us to detect a potentially differential effect of GDPR on healthcare-financial firms. Another reason is that the industry mix of ventures varies greatly across states, but every state has at least one deal in each of the two crude categories throughout the sample. Hence, the two-category grouping helps us construct a balanced sample at the aggregate level by week, state and category. We further collect local macroeconomic controls such as unemployment rate, CPI, interest rate, and GDP, for each state in which a venture is located.⁹

Table 1 reports the summary statistics at the aggregate level in the EU and US. Panel A indicates that our sample comprises 24 EU member states. On average, both the weekly dollar amount raised (in millions) and the number of deals per state per category are similar between the EU and US. Panel B reports summary statistics for each of the two categories we track. Figure 1 depicts weekly trends for the number of deals and dollar amount raised for the EU and US. From both the aggregate and average aspects, there are no noticeable differential trends in the EU and US.

⁷Despite Brexit, we include Great Britain as part of the treatment group due to its adoption of a GDPR-like regulation in the same time frame as the rest of the EU, and due to the fact that it is still bound by GDPR during our sample. In addition, the few observations we have for Bulgaria, Cyprus, Malta, and Lithuania are removed because some macroeconomic variables were not available for those member states at monthly frequencies.

⁸In the US, the Health Insurance Portability and Accountability Act (HIPPA) governs data collection, data use and data security for health care, and the Gramm-Leach-Bliley (GLB) Act governs similar issues for finance.

⁹For a few months in 2018 for which macroeconomic data was yet available (August and September for CPI and unemployment; July through September for GDP), we extrapolated macroeconomic variables by using their corresponding growth averages from 2017 and 2016.

At the individual-deal level, we remove observations with missing dollar amounts or missing funding types (e.g., ‘angel’, ‘seed’, ‘Series A’, etc).¹⁰ We calculate a venture’s time-variant age based on its founding date.¹¹ We consider four different age categories: new firms (0-3 years old), young firms (3-6 years old), established firms (6-9 years old), and mature firms (9+ years old). Firms may consequently switch between age categories in our sample. We also group deals, based on their funding type group, into three unique funding stages of pre-stage, main-stage, and late stage.¹²

Summary statistics at the deal level are also in Table 1. Panel C of the table indicates that the average dollar amount raised per deal is similar (about \$22 million) in the EU and US in our sample. Its distribution is highly skewed in both the EU and US, with the median dollar amount raised per deal (\$1.42 million for the EU and \$3 million for the US) much lower than the average. The average firm age is about 3 years in both the EU and US when excluding mature (9+ year old) firms. Category-specific deal-level summary statistics are presented in Panel D, where it can be seen that the average funding size for a healthcare-financial venture is larger in the EU than in the US. Panel E suggests that most funding deals take place in either the pre-stage or in the main stage. Panel F provides information about the distribution of firm ages in our dataset. While they are similar, the US has a larger proportion of 9+ year old firms. The EU, in contrast, has a larger proportion of firms in the 0-3 and 3-6 age groups, firms which may be particularly susceptible to an increase in the costs of compliance. Of particular interest is the fact that close to 70% of technology ventures in the EU and US in our sample are relatively young, 0-6 year old ventures. It is

¹⁰There are 21,726 deal observations in the overall sample. Of those, 4,358 observations are missing dollar amounts and an additional 175 observations are missing funding types, which together amount to about 20% of the overall sample. Dollar-amount specifications thus use 17,192 observations; number-of-deals specifications use the full sample comprising 21,726 observations.

¹¹There are some cases where a founding date is unavailable or when a venture’s first financing round predates its founding; in those cases, we use the venture’s first financing round as its founding date.

¹²Specifically, we group angel, seed, pre-seed, convertible note, and product crowdfunding into pre-stage, we group series A, B, C, bridge series A-B, initial coin offering, and equity crowdfunding into the main stage, and we group series D and later, private equity, debt financing, and other post IPO activities into the late stage. The precise grouping of funding types does not change the nature of the results.

also apparent that the older the firm, the higher the average dollar amount raised per deal.

3 Empirical Approach

We aim to study the effects of the rollout of GDPR in May 2018 on venture financing in the EU. We do so by contrasting venture activity in the EU with the US before and after the rollout of GDPR. While GDPR was enacted in April 2016, its enforceability began to take hold in May 2018, with mandatory implementation by EU member states and mandatory compliance by firms. Our hypothesis is that as GDPR’s enforceability came into place, entrepreneurs and investors both realized the actual compliance and implementation costs, as well as the ex-post implications of GDPR. This is particularly evident in the month immediately before the GDPR effective date, as major platforms like Google, Facebook, Amazon, and Apple, on which a vast number of technology ventures rely, began to reveal the ways in which they were tightening their platforms and app stores with new data sharing, data portability, and data liability rules.¹³

We test the effect of GDPR using a difference-in-differences methodology (DID). We carry out the empirical analysis at two levels. At the aggregate level, each observation is defined by week-state-category, while the dependent variables are either the total dollar amount raised across all deals in that week-state-category, or the number of deals reached in that week-state-category. Both could be zero if no venture in the state had any deal in that week-category. At the deal level, the sample includes every deal that has non-missing amount raised and

¹³Examples include SafeDK in 1/25/18 documenting that more than half of mobile applications are not GDPR ready (<https://www.mobilemarketer.com/news/study-55-of-apps-may-not-meet-gdpr-privacy-standards/515546/>), and numerous examples from May 2018. Those include Apple reportedly removing apps that share location data (<https://www.idownloadblog.com/2018/05/09/apple-removing-apps-location-data/>) and updating its privacy terms (<https://techcrunch.com/2018/05/23/apple-introduces-new-privacy-portal-to-comply-with-gdpr/>), Facebook announcing that “Businesses may want to implement code that creates a banner and requires affirmative consent? Each company is responsible for ensuring their own compliance” (<https://developers.facebook.com/ads/blog/post/2018/05/10/compliance-protections-gdpr/>), Shopify updating its app permissions for merchants and developers (<https://www.shopify.com/partners/blog/gdpr-compliance>), Google releasing new consent requirements to developers (<https://bit.ly/2ziUgJA>), all shortly before GDPR took effect on May 25, 2018.

non-missing funding types. By definition, the dependent variable (amount raised per deal) is always positive. Since both levels of analysis use samples derived from the same raw data, we consider the total dollar amount raised per week-state-category an overall metric of venture investment, which encompasses an extensive margin (total number of deals per week-state-category) and an intensive margin (amount raised per deal conditional on having reached a deal).

Our treatment group comprises EU ventures and our control group comprises US ventures. While the treatment group does have lower levels of venture activity than the control group, there does not appear to be a differential pre-trend that would violate the common trend assumption in our DID analysis. At the aggregate level, Figure 1 depicts trend lines of the weekly total amount raised in the EU and the US, the weekly total number of deals in the EU and the US, the average amount raised per week-state-category, and the average total number of deals per week-state-category from July 2017 to September 2018. All four subfigures suggest that some sustained divergence took place between EU and US ventures around the time that GDPR came into effect. Both EU and US trends also track each other closely otherwise, and particularly so up until May 2018. At the deal level, Figure 2 confirms that there are no differential trends between the EU and US in the frequency of deals or in the average dollar amount raised per deal.

For aggregate-level analysis, the specification is given by:

$$y_{sct} = \alpha_s + \alpha_c + \alpha_t + \delta X_{st} + \beta GDPR_{sct} + \varepsilon_{sct},$$

where s denotes state, c denotes category, t indexes week, $GDPR_{sct}$ indicates whether the state s is located in the treatment group (EU) and subject to GDPR at time t . Y_{sct} is the dependent variable of interest, which is either the total dollar amount raised or the number of financing deals in each week-state-category. Week, state and category fixed effects are denoted by α_t , α_s and α_c , respectively, whereas X_{st} are state-specific macroeconomic control

variables (monthly unemployment rate, CPI, interest rate, and quarterly GDP), and ε_{sct} is an error term. We use a Tobit specification censored at 0 for dollar amount regressions because we only observe deals that go through. We use a Poisson specification for the number of deals regressions due to a relatively large number of zeroes at the week-category-state observation level. In all cases, we obtain similar results with OLS.

At the aggregate level, the coefficient β captures the effect of GDPR across both categories. Standard errors are clustered at the state level, because the GDPR requires state-specific enforcement and the heterogeneity is confirmed in market perception.¹⁴

Figure 3 depicts coefficient plots of the monthly pre-treatment tests for the number of deals and dollar amount raised using Poisson and Tobit specifications, respectively. To perform pre-treatment tests, we run the same specification for the pre-GDPR data, including a full set of interactions between the dummy of EU and each monthly dummy. The coefficients of these interactions are shown in the figures, along with their confidence intervals. Figure 3(a) shows that there is no pre-existing differential trend between the EU and US in the number of deals before May 25, 2018, confirming the observable trends in Figures 1(a) and 1(c). Figure 3(b) suggests there is no pre-existing differential trend between the EU and US in the aggregate dollar amount raised per week before April 30, 2018. It is possible that due to some of the major platforms announcing their data-related policy changes in early May 2018, some market movement may have taken place earlier than May 25. As robustness checks, we report the results excluding May 2018 and they are largely unchanged. For heterogeneous effects by category, we apply the same specification to healthcare-financial

¹⁴Despite GDPR applying to all EU countries, the policy change is at the state level. This follows from the definition of the ‘lead supervisory authority,’ which has the “primary responsibility for dealing with a cross-border data processing activity, for example when a data subject makes a complaint about the processing of his or her personal data.” The location of the lead supervisory authority is based on a firm’s main establishment location. Recital 127 further states that: “Each supervisory authority not acting as the lead supervisory authority should be competent to handle local cases where the controller or processor is established in more than one Member State, but the subject matter of the specific processing concerns only processing carried out in a single Member State and involves only data subjects in that single Member State.” Goldberg et al. (2018) additionally demonstrate that GDPR suffers from implementation heterogeneity across EU countries, heterogeneity that lines up with traditional member state enforcement behaviors.

and other technology separately. For heterogeneous effects by firm age, we reorganize the aggregate data by week, state and firm age (0-3, 3-6, 6-9, 9+). We then apply the same specification to each group of firm age separately.

For each aggregate-level regression, we perform pre-treatment tests by quarter. To conduct such tests, we run the same specification using pre-GDPR data only, include interactions between the EU dummy and quarterly dummies (2017Q4, 2018Q1, 2018Q2 while 2017Q3 is the default), and test whether the coefficients of these pre-treatment interactions are jointly zero.¹⁵ The test-statistics is never significant above the 90% confidence level. In few cases, the pre-treatment test gets close to be significant at the 85-90% level, mostly because the negative effect of GDPR started to appear a few weeks before May 25, similar to what we have shown in Figure 3(b). For this reason, the reported estimate of β is likely a conservative estimate of the true effect.

At the deal level, we use the specification:

$$\ln(y_{jsct}) = \alpha_s + \alpha_c + \alpha_t + \delta X_{jsct} + \beta GDP R_{jsct} + \varepsilon_{jsct},$$

where j identifies deals according to their assigned unique identifier, the dependent variable $\ln(y_{jsct})$ is log of the dollar amount raised in deal j , X_{jsct} denotes deal-level variables such as funding type, investor type, and firm age, α_t , α_s , and α_c are week, state and category fixed effects, $GDP R_{jsct}$ is a dummy variable that equals one after May 25, 2018 if applicable to deal j in state s and zero otherwise, and ε_{jsct} is an error term. We use the log of the amount raised per deal because the amount raised is always positive but its distribution is highly skewed. As a robustness check, we also report results that top-code the amount raised at the 95 percentile of the sample. The same specifications apply to the subsamples by category and by firm age. In all regressions, we cluster the error term by state.

¹⁵We test quarterly instead of monthly or weekly interactions because the more we zoom in a particular subsample, the more spontaneity there is in a short time window. Quarterly gives us a reasonable time frame to average over these idiosyncrasies, across all types of samples.

4 Overall Effects of GDPR

We begin by examining how the aggregate weekly dollar amount for each state in each category changes from the pre to the post period of the implementation of GDPR. Column 1 of Table 2 reports the results of our baseline specification. We focus on the *marginal* effects computed from the estimated coefficients of GDPR according to the Tobit specification. It can be seen that on average, each category in each EU state incurs a \$13.90 million decrease after the rollout of GDPR. In Column 2, we right censor the weekly aggregate dollar amount at the 95-percentile value (\$175 million) to reduce the influence of outliers. Column 2 suggests that each category in each EU state, after top coding, experiences a \$3.38 million decrease following the rollout of GDPR. The effects do not change when adding a linear trend (column 3 of Table 2).

As previously indicated, our pre-treatment tests suggest that the effect of GDPR’s rollout on EU ventures may have started earlier in May. Column 4 demonstrates that when May is excluded, the estimated effect is even stronger, suggesting a \$4.49 million decrease in the aggregate dollar amount per state per category after the rollout of GDPR. Using an OLS specification, Column 5 indicates that the aggregate dollar amount for each state in each category faces a 26.5% decrease after the rollout of GDPR. This marginal effect on dollar amount is computed from the estimated coefficient of GDPR, accounting for the fact that the dependent variable is log of one plus the dollar amount.

The aggregate dollar amount estimation combines extensive margin (number of deals) and intensive margin (dollars raised per deal) effects. To decompose them, we estimate the effect of GDPR’s rollout on the number of deals per week-state-category, and on the dollar amount per deal. Since the number of deals per week-state-category is an integer count (including zero) but the dollar amount per deal is always positive, we use a Poisson specification for the former (at the aggregate level) and the log linear specification for the latter (at the deal level).

Table 3 reports the effect of GDPR’s rollout on the aggregate number of deals. We focus on the *marginal* effects computed from the estimated coefficients of GDPR according to the Poisson specification. Our baseline model suggests a 17.6% decrease in the number of EU venture deals. Column 2 adds a linear trend to the baseline specification, and Column 3 excludes May observations. Both specifications give similar results (in the case of Column 3, excluding May gives a greater decrease of 22.82%). An OLS specification with the dependent variable of $\ln(1 + \# \text{ of deals})$ in Column 4 indicates a 9.07% decrease. The marginal effect on the number of deals is computed from the estimated coefficient of GDPR, accounting for the fact that the dependent variable is log of one plus the number of deals.

Table 4 provides the results of the deal-level log linear specification. Our baseline model in Column 1 suggests a 39.6% decrease in the dollar amount per deal after the rollout of GDPR. Column 2 summarizes a similar estimation when right censoring the sample at the 95-percentile level, and suggests a similar reduction of 38.0%. Column 3 adds a linear trend and Column 4 excludes May from the sample; the results are similar in both cases.

5 Heterogeneous Effects of GDPR

While the overall effects we measure may be negative and statistically significant, there may exist heterogeneity both at the aggregate level and at the deal level in the effects across firm technology category types and firm age groups. This section applies the baseline specification to some of these subsamples. An example of such heterogeneity is presented in Figures 4 and 5, where the weekly trends of the number of deals and dollar amounts for the EU and US are plotted separately for healthcare-financial and other technology. There is no apparent differential trend between the EU and US prior to May 2018, and GDPR seems to have had effects in both category groups. Figure 4 in particular demonstrates that the sample comprises some significant outliers and, as such, right-censoring the data to mitigate the impact of outliers on our dollar amount estimations is likely necessary.

Table 5 reports the baseline model specification for the number of deals, the aggregate dollar amount raised (with right censoring at the 95-percentile level), and the dollar amount raised per deal for healthcare-financial and other technology, respectively.

Columns 1, 3 and 5 of Table 5 indicate that GDPR had negative effects on the number of deals, the aggregate dollar amount raised, and the dollar amount raised per deal by EU ventures in the healthcare-financial category. The three columns suggest a decline in the number of deals of 18.86% and reductions of \$5.22 million and 56.6% in the aggregate dollar amount per week and in the dollar amount per deal, respectively. For the group comprising all other technology, Columns 2 and 6 suggest large negative effects for the number of deals and dollar amount per deal; however, the effect on the aggregate dollar amount per week is insignificant in Column 4. We believe this is possibly because of a large standard deviation on the aggregate dollar amount per week in this rather broad category.

The somewhat comparable negative effects on ventures in healthcare-financial versus other technology may be perceived as surprising, in light of the seemingly stringent healthcare and financial data privacy laws in the US (HIPPA and GLB). However, those laws are arguably older and systems to comply with them have been in place for a number of years. Moreover, HIPPA, for instance, allows providers to require consent prior to providing services, a requirement that GDPR explicitly prohibits. GLB, for instance, adopts an opt-out approach, where information is collected by default and consumers have a limited ability to opt out. GDPR, in contrast, mandates an informed opt-in consent for each type of data collected, and further requires data management, data auditing and classification, data risk identification and mitigation, and data interfaces for users to easily obtain their own data and request that it be deleted. GDPR also imposes substantially larger penalties of up to 4% of a firm's global revenues. Hence, our results potentially indicate that GDPR is widely transformational across the technology sector in a rather broad way, even when compared to existing strict data regulations of US firms.

Another example of subsample heterogeneity is presented in Figures 6 and 7, where the

average weekly number of deals and aggregate dollar amount raised, respectively, are depicted for four different age groups of EU and US ventures. Similarly to the preceding analysis, we also examine the heterogeneous effects of GDPR across these firm age groups. Under the aggregate dollar amount per week specification, when examining subsamples according to firm age groups, Column 1 of Table 6 suggests that GDPR had a negative effect on the aggregate dollar amount invested per week in the new (0-3 year old) firm subsample, with a reduction of \$0.9 million per week. Columns 2 and 3 indicate insignificant effects for firms in the 3-9 year old age group, and Column 4 indicates a larger negative effect of \$7.1 million per week for firms in the 9+ year-old age group. Under the number of deals per week specification, Table 7 presents similar findings. Columns 1 and 4 indicate significant reductions in the number of deals for firms in the 0-3 and 9+ year-old subsamples of 19.02% and 29.53%, respectively, and Columns 2 and 3 indicate insignificant effects for firms in the intermediate (3-9 year old) subsamples.

At the deal level, Column 1 of Table 8 suggests a negative effect under the deal-level dollar amount specification for firms in the 0-3 year old subsample, with a reduction of 27.1% per deal. Columns 2 and 3 suggest negative effects for firms in the 3-6 and 6-9 age groups, with reductions of 31.4% and 77.3%, respectively. Due to the absence of significant aggregate effects for these age groups in terms of the weekly number of deals and total dollar amounts invested, we believe these reductions may be the result of the more sporadic nature of deals for firms in these age groups.¹⁶ Column 4 indicates an insignificant deal-level effect for the subsample comprising 9+ year-old firms.

Ventures in the 0-3 year old group in our sample are those ventures that primarily seek seed, series A and series A-B bridge rounds, as indicated in Figure 8 — rounds where angel investment and venture capital begin to overlap for the first time, with venture capital replacing funding that was previously raised from angel investors. In the figure, the circles

¹⁶Firms in the 0-3 age group comprise 46% of our observations, whereas 3-6 year-old firms comprise 28%, and 6-9 year-old firms comprise 10.5%.

depict the relative numbers of observations, with larger circles indicating more deals. The numbers of seed, Series A, and Series A-B deals in Figure 8(a) for 0-3 year old firms are significantly larger than their corresponding numbers in Figures 8(b)-8(d) for 3+ year-old firms. The combined estimates of Tables 6 through 8 thus suggest that those nascent firms that most critically depend on angel investment as well as those firms that are in the process of making the transition from angel investment to venture capital are particularly susceptible to a negative effect from GDPR.

6 Implications for Employment

For each venture, our dataset provides a range on the number of employees (e.g., 1-10, 10-50, 50-100, etc). However, this range is time-invariant as of the composition time of our dataset, October 1, 2018; thus, we do not have historical ranges. In other words, as of October 1, 2018, we have the total dollar amount raised by each firm and a range on its current number of employees. We can use this information to provide a back-of-the-envelope measure of the average dollar amount raised per ‘current’ employee as a function of the firm’s age. We focus on new (0-3 year old) ventures, because the literature has demonstrated that they tend to be the primary job creators,¹⁷ and we focus our analysis on EU ventures to assess the potential for EU technology job losses as a result of GDPR.

To provide a back-of-the-envelope calculation of the effect on jobs, we calculate the average dollar raised per deal and the total number of deals in the post-GDPR period by 0-3 year old ventures. Columns 4 and 5 of Table 9 show our calculations. In the third quarter of 2018, 0-3 year old ventures raised on average \$3.32 million per deal, with 690 deals made by 0-3 year old ventures. Once the estimated reduction in the dollar raised per deal and the reduction in the number of deals are applied to those firms’ totals, the predicted overall dollar amount suggests a \$1589.77 million decrease for 0-3 year old firms. This is corroborated at

¹⁷See, e.g., Haltiwanger et al. (2013).

the aggregate level when using estimates of the change in the total weekly dollar amount invested per state per category to estimate the decrease in the aggregate amount invested in 0-3 year-old EU firms after GDPR, suggesting \$1217.70 million decrease for 0-3 year old firms, in line with the first approach. To extrapolate these losses to an entire calendar year, we examine the ratio of EU venture activity in June through September 2017 to the rest of calendar year 2017 for new ventures, and apply the same multiplier to these totals.

To identify the average bounds on the annualized dollar amount raised per employee, we focus on EU ventures founded on or after 2015. An advantage of focusing on 0-3 year old firms is that our dataset can provide their total amounts raised since their founding. The average age for ventures in this group is 1.21. For each firm in this group, its total dollar amount raised is divided by the bounds on its employee range, which provides a lower and upper bound on its total dollar amount raised per current employee. We then obtain the average lower and upper bounds across this subsample and divide it by the average firm age to obtain a crude back-of-the-envelope calculation of the dollar amount raised per employee per year. This calculation gives an average lower bound of \$0.123 million and an average upper bound of \$1.02 million for 0-3 year old firms.

We can use these average annualized bounds of dollar amount raised per employee and the estimated annual investment losses following the rollout of GDPR to obtain a back-of-the-envelope estimate of the number of technology venture jobs lost in the EU. We obtain a lower bound of 4,705 jobs lost and an upper bound of 38,931 jobs lost from our first approach (using the number of deals times the dollar amount per deal), and a lower bound of 3,604 jobs and an upper bound of 29,819 jobs by using the total dollar amount approach. We tally up the ranges of employees for all firms founded since 2015 in our sample.¹⁸ Using the approach based on the aggregate dollar amount invested, as a percentage of the total

¹⁸Firms that are missing employee ranges in the subsample are counted by assuming they have the subsample average employee range. This may mean that our results are understated since some of those firms may have ceased operating. Constructing the subsample with firms founded on or after January 2015 may also understate our results since some of those firms may be entering their fourth year of operation in 2018.

range on employment by EU ventures in the subsample comprising firms in this age group, the estimated job loss bounds translate to a 4.09%-11.20% loss in the number of individuals employed by those firms.

7 Conclusion

We presented analyses of the short-term effects of the rollout of GDPR on investment in technology ventures. We found evidence suggesting negative and pronounced effects following the rollout of GDPR on the number of venture deals, the size of those deals, and the overall amount of dollars invested. We broke down those effects according to two venture categories and four venture age groups, and presented a rough estimate of the effect on the number of jobs for 0-3 year old technology ventures.

It is important to emphasize that our dataset is not a complete universe of venture funding, but rather a partial snapshot of primarily venture capital and angel investments in technology ventures. As such, our results must be taken with a bit of caution, given that the effects we observe may be incomplete. At the same time, our findings indicate that it is exactly those nascent ventures that are in the process of transitioning from angel to venture capital that may be most impacted by GDPR.

Another caveat is that the impact estimated on EU ventures is relative to their US counterparts. To the extent that capital flows freely across continents, it is unclear whether the reduced investment in the EU may have in tandem translated to additional support for US ventures or that it reflects the reluctance by investors to invest anywhere. If it is the former, our estimates may have overestimated the effects of GDPR; if it is the latter, our estimates may be conservative as our sample does not include ventures that could serve EU residents but are based in other countries.

While our analyses concern the amount of dollars invested in technology, they are not necessarily translatable into welfare implications. For instance, a reduction in investment

dollars in technology ventures could benefit welfare if firms that are potentially harmful from a societal perspective do not come to fruition. Similarly, it could be that data regulation encourages new types of innovation further down the road.

It is also important to emphasize that given our data, our measure of the effect on jobs can only provide rough back-of-the-envelope ranges on job loss estimates. This is because, on the one hand, we have no insight into whether investors are taking a wait-and-see approach nor do we know the outside options of affected firms or of those individuals who would have been employed by the firms in our dataset had it not been for GDPR. There may also be jobs created as a result of GDPR (for instance, data privacy compliance officers, data security and management ventures, etc). On the other hand, the potential for job losses may well extend and intensify past our four months post-GDPR dataset period, in which case the effect on jobs is understated. Moreover, our estimates do not incorporate potential foregone related job creation (for instance, downstream jobs to service the additional employees that may have been employed had it not been for the rollout of GDPR). The long-term impact of GDPR on the EU technology venture scene will become clearer in the years ahead.

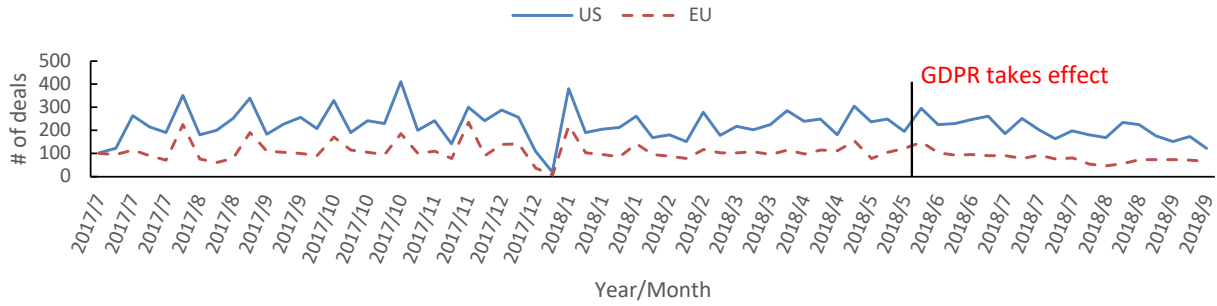
References

- [1] ACQUISTI, A., C. TAYLOR, AND L. WAGMAN (2016): “The economics of privacy,” *Journal of Economic Literature*, 54(2), 442–492.
- [2] BERTRAND, M., E. DUFLO, AND S. MULLAINATHAN (2004). “How much should we trust differences-in-differences estimates?” *Quarterly Journal of Economics*, 119(1), 249–275.
- [3] CAMPBELL, J., A. GOLDFARB, AND C. TUCKER (2015): “Privacy regulation and market structure,” *Journal of Economics & Management Strategy*, 24(1), 47–73.

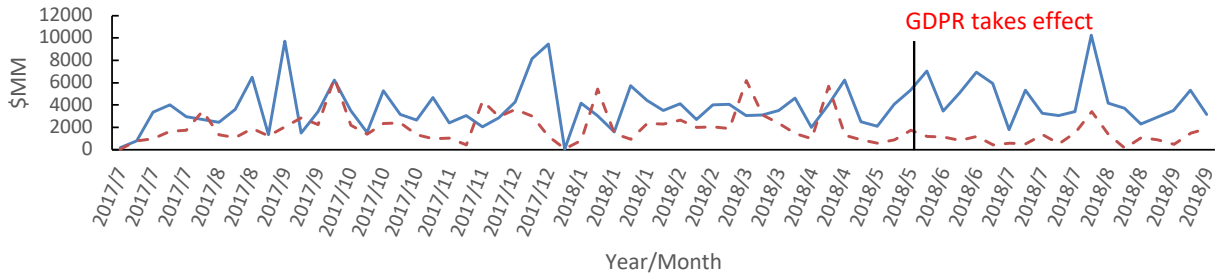
- [4] CHATTERJI, A., S. DELECOUR, S. HASAN, AND R.M. KONING (2018). “When does advice impact startup performance?” *NBER Working Paper*.
- [5] DJANKOV, S., C. MCLIESH, AND A. SHLEIFER (2007): “Private credit in 129 countries,” *Journal of Financial Economics*, 84(2): 299–329.
- [6] DOBLAS-MADRID, A., AND R. MINETTI (2013): “Sharing information in the credit market: Contract-level evidence from U.S. firms,” *Journal of Financial Economics*, 109(1), 198–223.
- [7] FEDERAL TRADE COMMISSION (2014): “Data brokers: A call for transparency and accountability.” Accessible at <https://www.ftc.gov/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014>.
- [8] FEDERAL TRADE COMMISSION (2016): “Big data: A tool for inclusion or exclusion?” Accessible at <https://www.ftc.gov/system/files/documents/reports/big-data-tool-inclusion-or-exclusion-understanding-issues/160106big-data-rpt.pdf>.
- [9] GOLDBERG, S., G. JOHNSON, AND S. SHRIVER (2018): “Regulating privacy online: An early look at Europe’s GDPR,” *Working paper*.
- [10] GOLDFARB, A., AND C.E. TUCKER (2011): “Privacy regulation and online advertising,” *Management Science*, 57(1), 57–71.
- [11] GOLDFARB, A., AND C.E. TUCKER (2012): “Privacy and innovation,” *Innovation Policy and the Economy*, 12(1), 65–90.
- [12] HALTIWANGER, J., R. S. JARMIN, AND J. MIRANDA (2013): “Who creates jobs? Small versus large versus young,” *Review of Economics and Statistics*, 95(2), 347–361.
- [13] HERTZBERG, A., J. LIBERTI, AND D. PARAVISINI (2011): “Public information and coordination: Evidence from a credit registry expansion,” *Journal of Finance*, 66(2), 379–412.

- [14] HOCHBERG, Y.V. (2016). “Accelerating entrepreneurs and ecosystems: The seed accelerator model,” *Chapter in NBER book Innovation Policy and the Economy*, 16(2016), 25–51.
- [15] HOEL, M. AND T. IVERSON (2002): “Genetic testing when there is a mix of compulsory and voluntary health insurance,” *Journal of Health Economics*, 21, 253–270.
- [16] JAFFE, A.B. AND K. PALMER (1997): “Environmental regulation and innovation: A panel data study?” *Review of Economics and Statistics*, 1997: 610–619.
- [17] KAPLAN, S.N. AND J. LERNER (2017): “Venture capital data: Opportunities and challenges,” *Measuring Entrepreneurial Businesses: Current Knowledge and Challenge*.
- [18] KERR, W.R., J. LERNER, AND A. SCHOAR (2014). “The consequences of entrepreneurial finance: Evidence from angel financings,” *Review of Financial Studies*, 27(1), 20–55.
- [19] KIM, J.-H. AND L. WAGMAN (2015): “Screening incentives and privacy protection in financial markets: A theoretical and empirical analysis,” *RAND Journal of Economics*, 46(1), 1–22.
- [20] KORTUM, S. AND J. LERNER (2000): “Assessing the contribution of venture capital to innovation,” *Rand Journal of Economics*, 31(4), 674–692.
- [21] KRASTEVA, S., P. SHARMA, AND L. WAGMAN (2015): “The 80/20 rule: Corporate support for innovation by employees.” *International Journal of Industrial Organization*, 38: 32–43.
- [22] LERNER, J., A. SCHOAR, AND S. SOKOLINSKI (2018). “The globalization of angel investments: Evidence across countries,” *Journal of Financial Economics*, 127(1), 1–20.

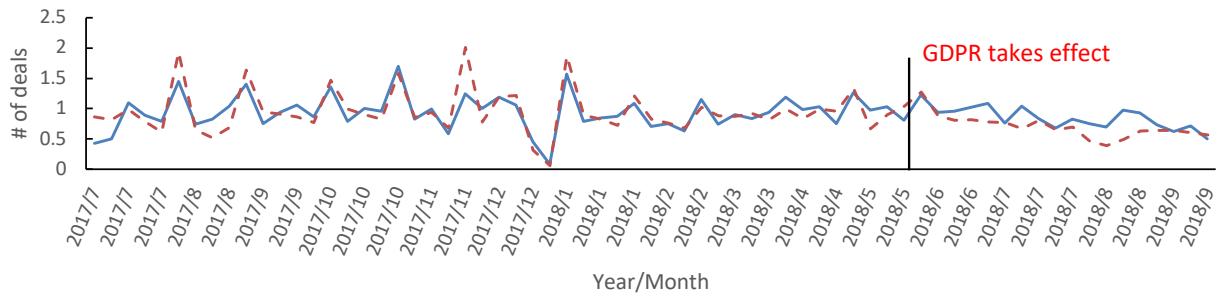
- [23] MILLER, A.R. AND C.E. TUCKER (2009): “Privacy protection and technology diffusion: The case of electronic medical records,” *Management Science*, 55(7), 1077–1093.
- [24] MILLER, A.R. AND C.E. TUCKER (2011): “Can health care information technology save babies?” *Journal of Political Economy*, 119(2), 289–324.
- [25] MILLER, A.R. AND C.E. TUCKER (2017): “Privacy protection, personalized medicine, and genetic testing,” *Forthcoming Management Science*.
- [26] PAGANO, M. AND T. JAPPELLI (1993): “Information sharing in credit markets,” *Journal of Finance*, 48(5), 1693–1718.
- [27] PAGANO, M. AND T. JAPPELLI (2002): “Information sharing, lending and defaults: Cross-country evidence,” *Journal of Banking and Finance*, 26(10), 2017–2045.
- [28] PORTER, M.E. (1991): “America’s green strategy,” *America’s Green Strategy*, 264(4), 168.
- [29] QIAN, J., AND P. STRAHAN (2007): “How laws and institutions shape financial contracts: The case of bank loans,” *Journal of Finance*, 62(6), 2803–2834.
- [30] TUCKER, C.E., (2014): “Social networks, personalized advertising, and privacy controls,” *Journal of Marketing Research*, 51(5), 546–562.



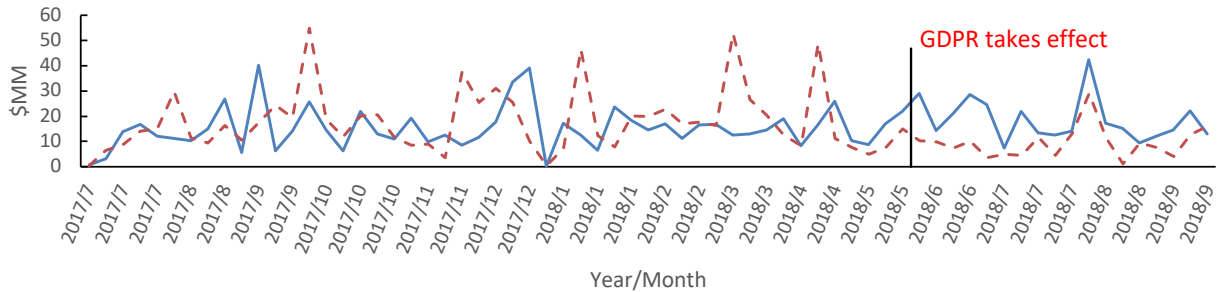
(a) Total # of deals per week between EU and US



(b) Total \$ amount raised per week between EU and US

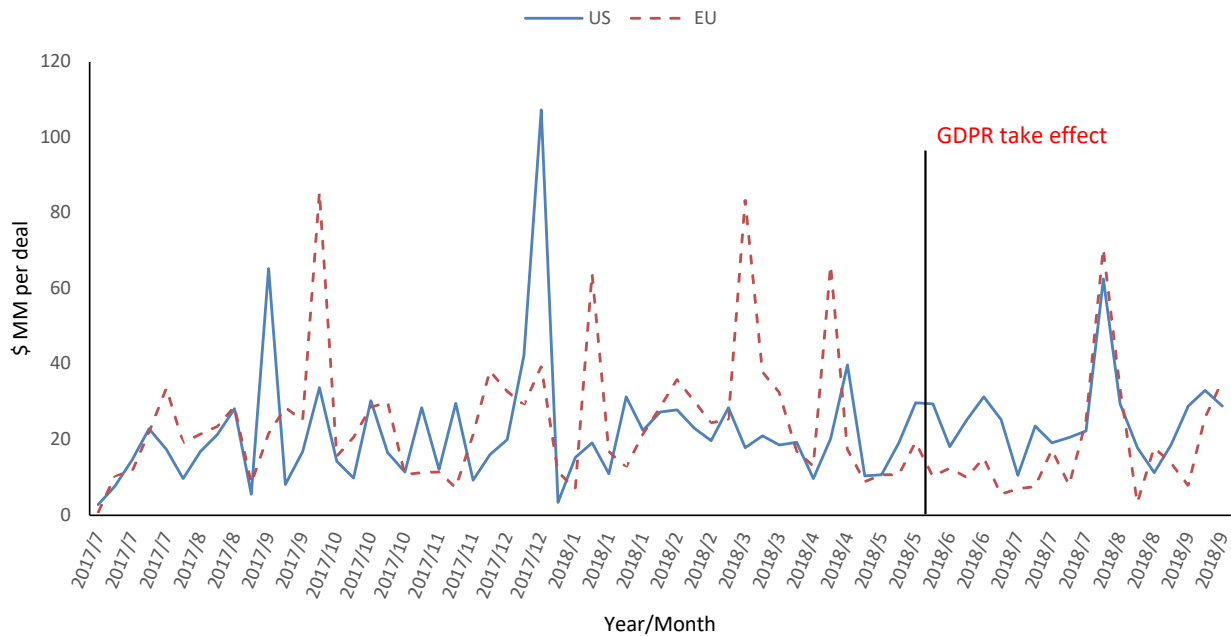


(c) Average weekly # of deals per category per state between EU and US

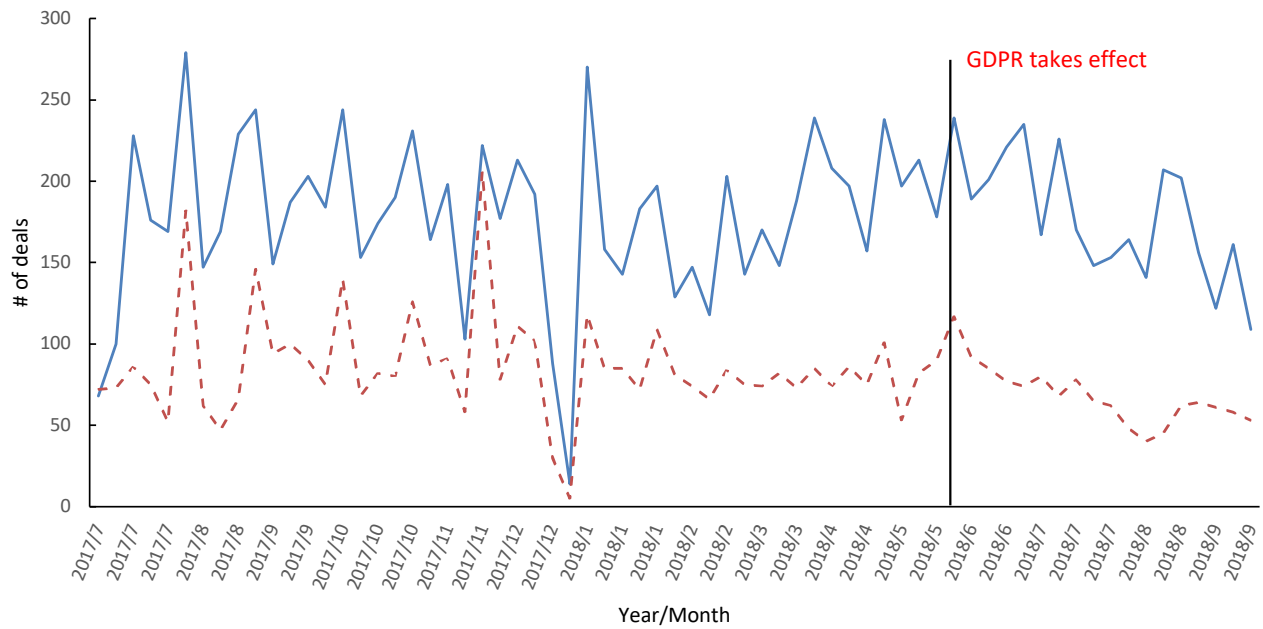


(d) Average weekly \$ raised amount per category per state between EU and US

Figure1. Aggregate-level data plots

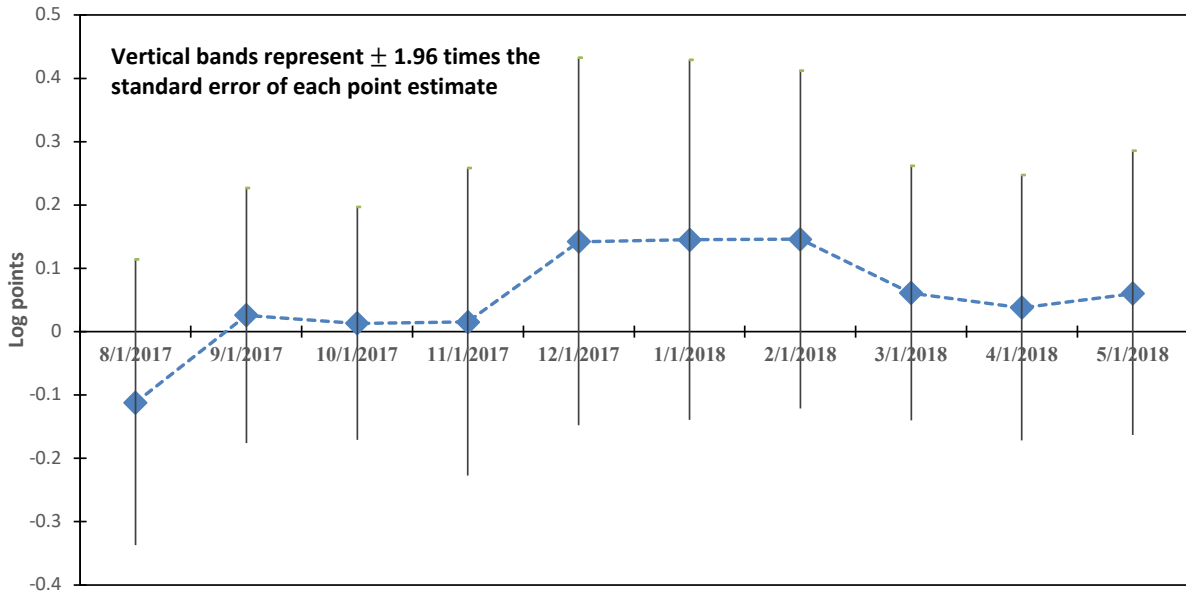


(a). Average \$ amount raised per deal between EU and US

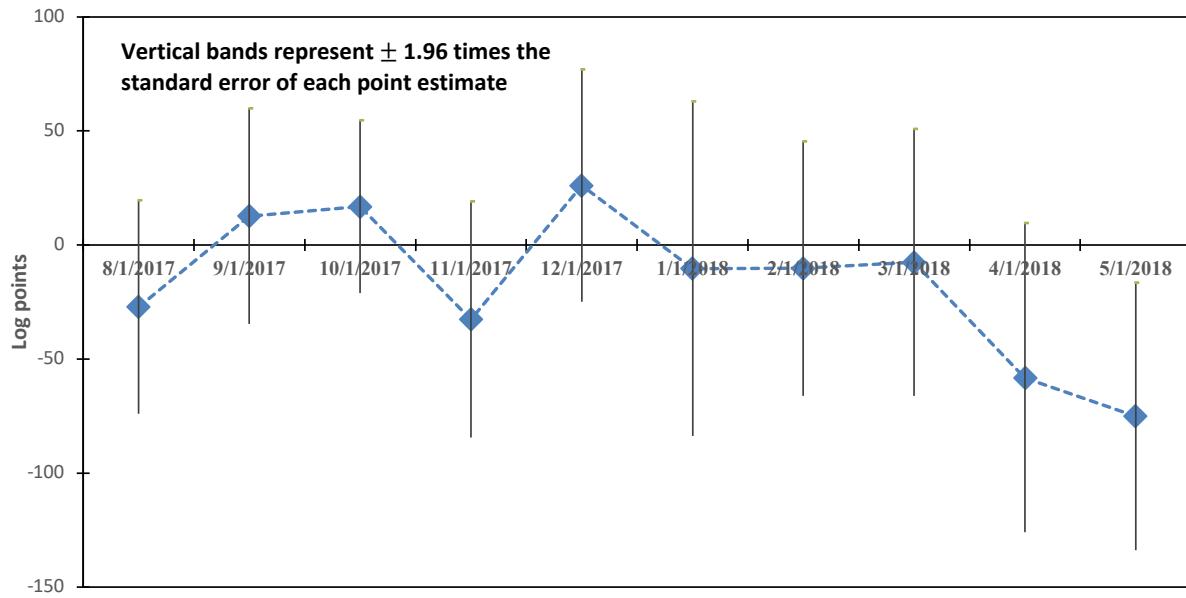


(b). Average weekly # of deals between EU and US

Figure 2. Deal-level data plots

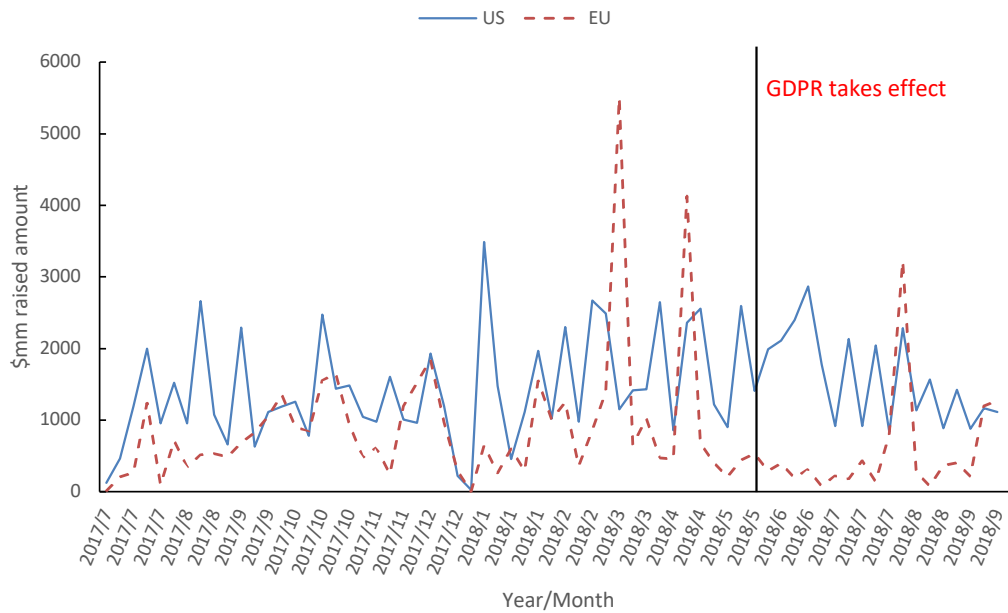


(a) Pre-treatment test for aggregate level # of deals – Poisson regression

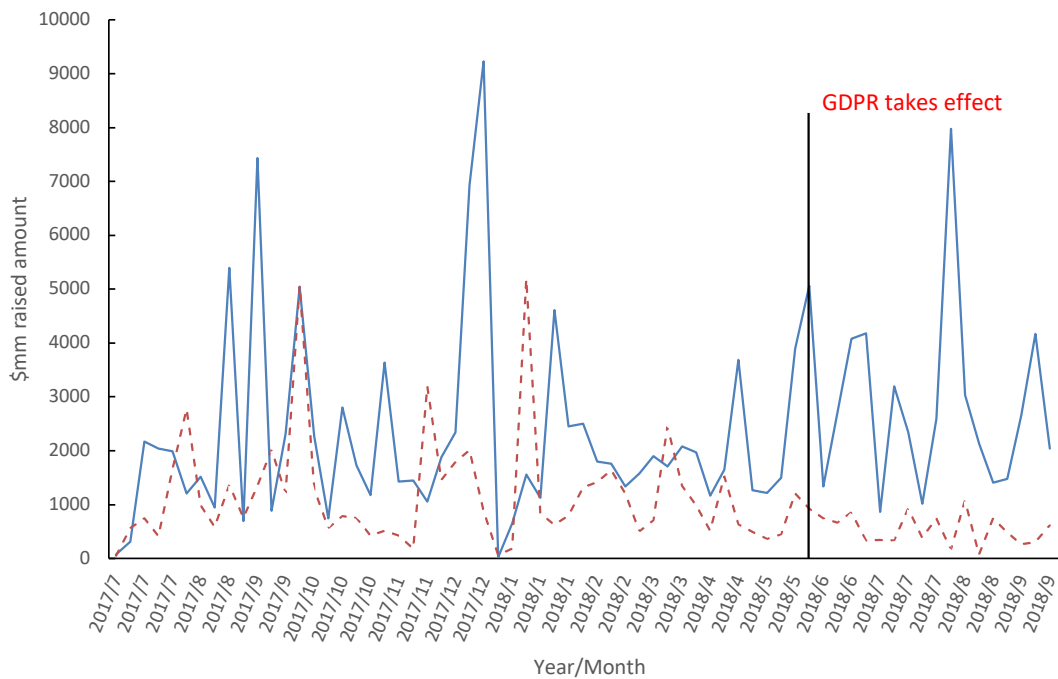


(b) Pre-treatment test for aggregate level \$ raised amount – Tobit regression

Figure 3. Pre-treatment tests for aggregate level # of deals and \$ raised amount

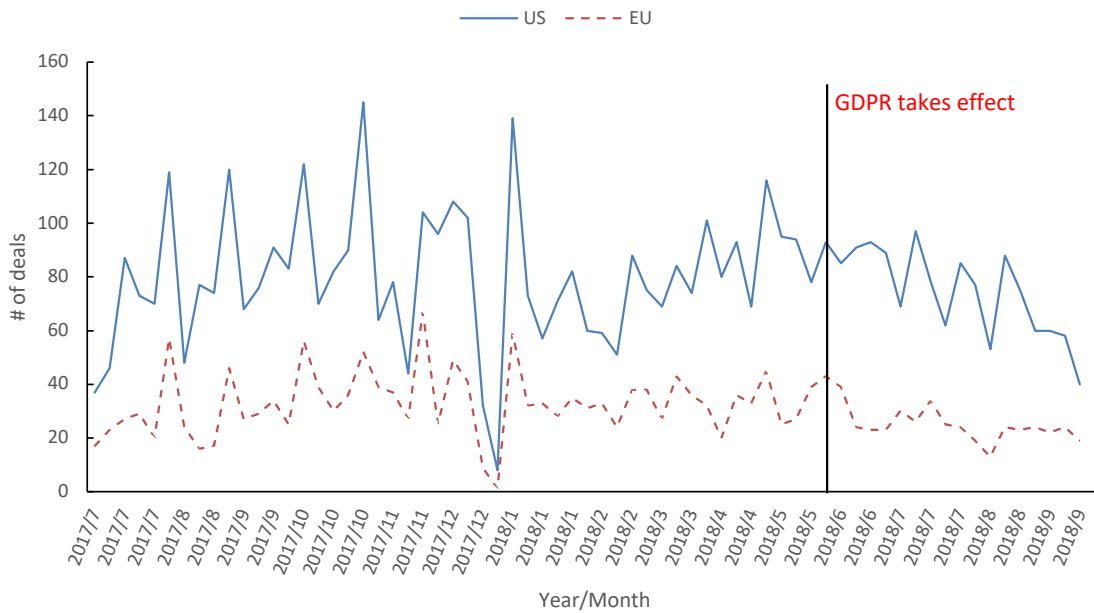


(a) Total weekly \$mm raised amount – healthcare-financial category group

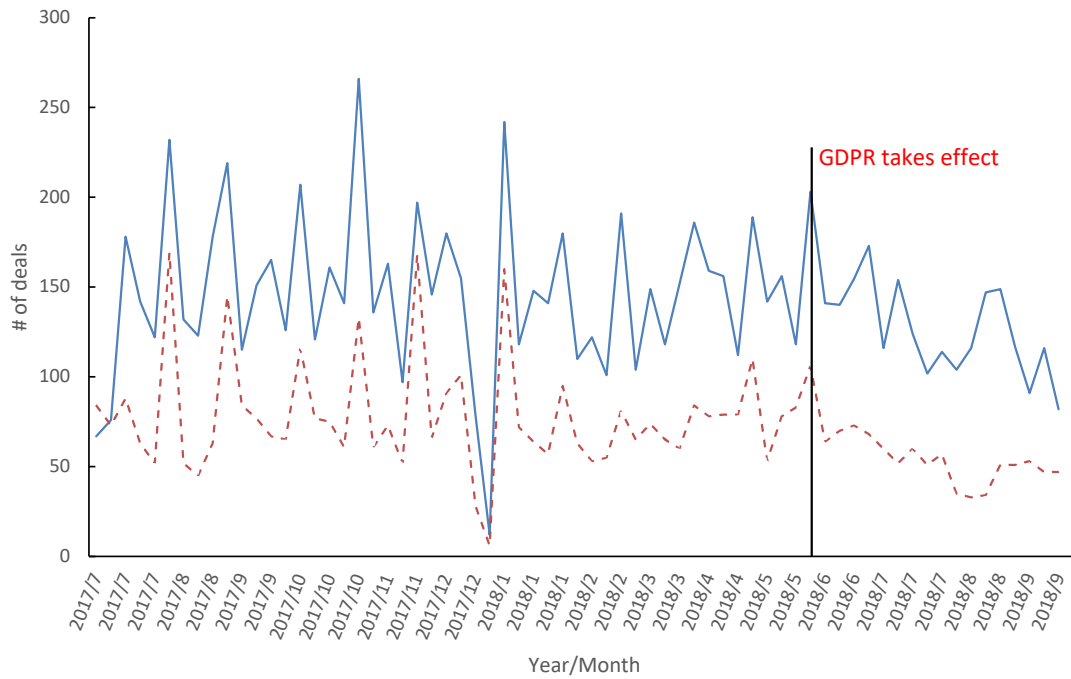


(b) Total weekly \$mm raised amount – other-categories group

Figure 4. Total weekly \$mm raised amount per category group

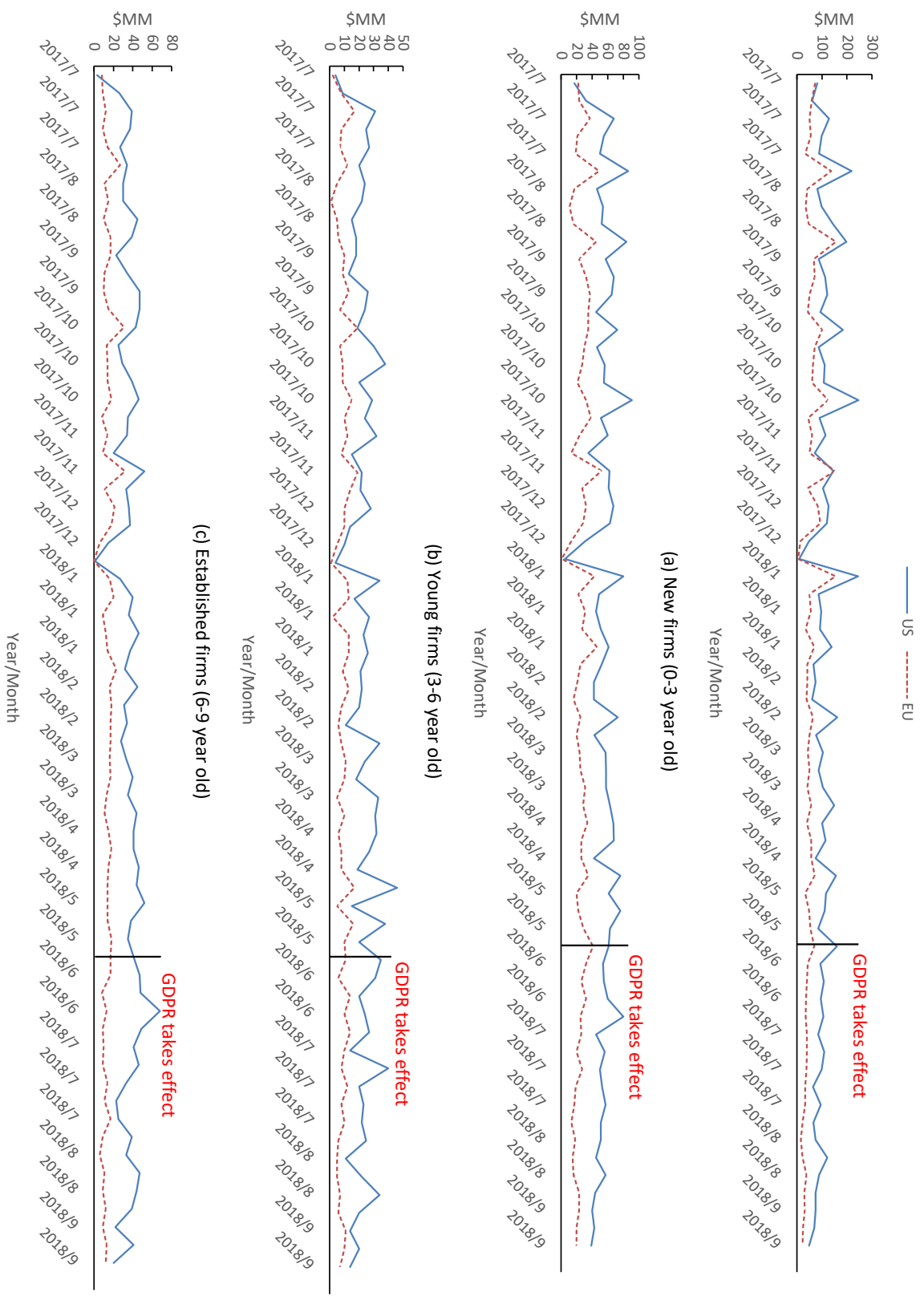


(a) Total weekly # of deals – healthcare-financial category group



(b) Total weekly # of deals – other-categories group

Figure 5. Total weekly # of deals per category group



(d). Mature firms (9+ year old)
Figure 6. Weekly # of deals by firm age

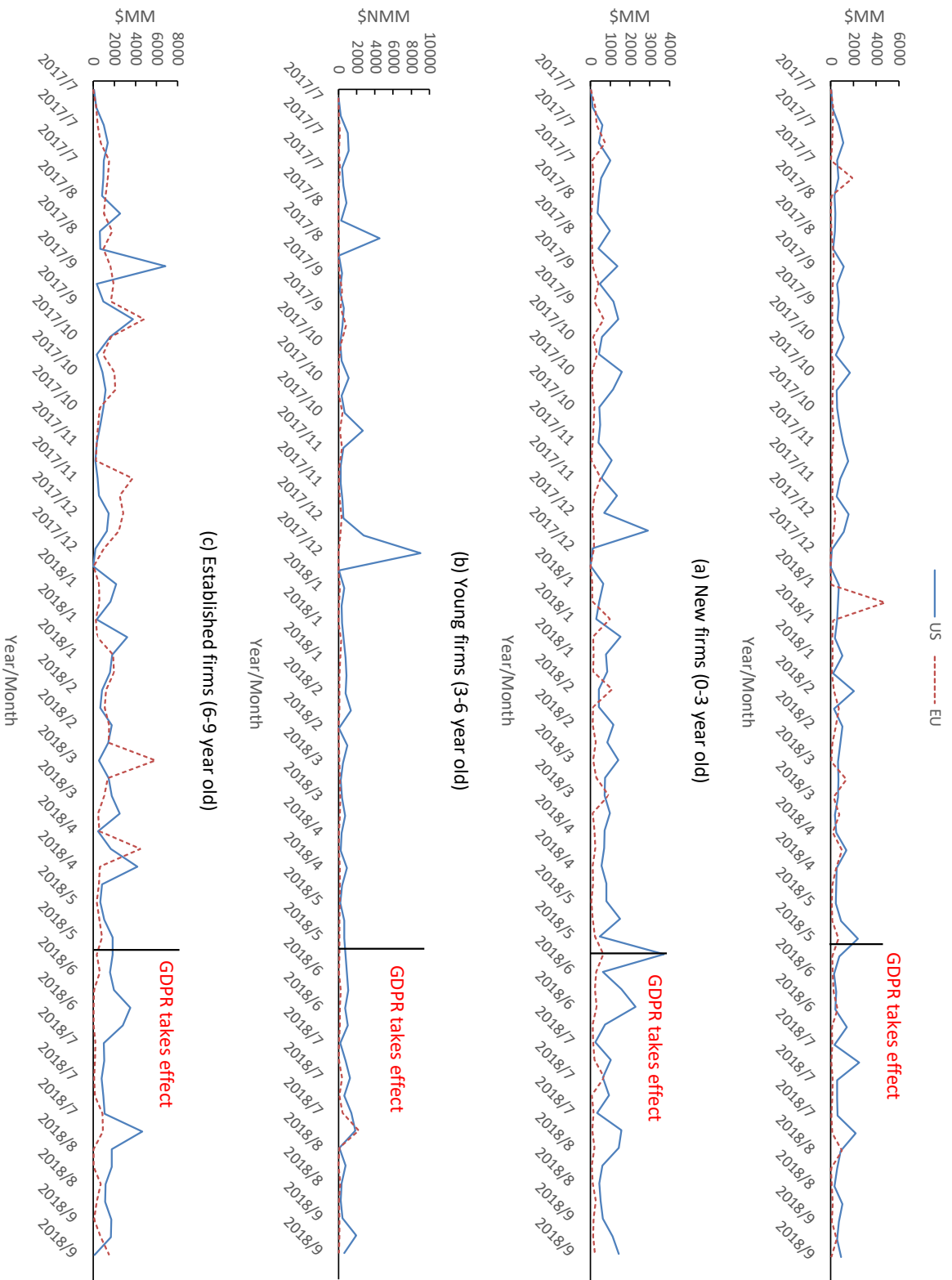
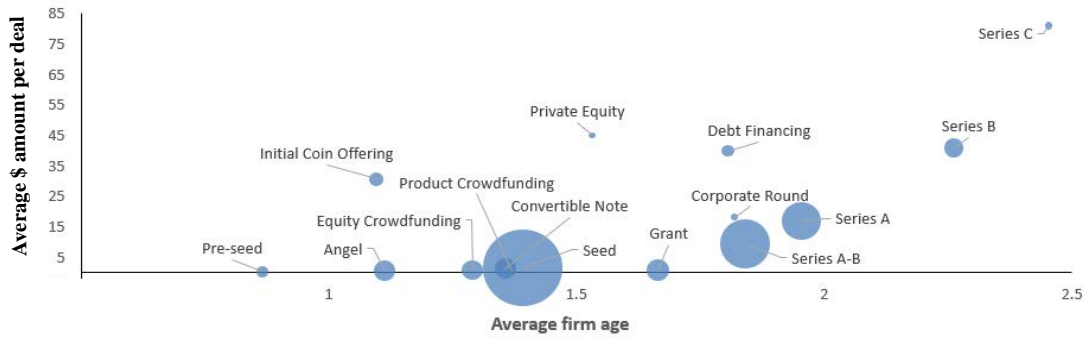
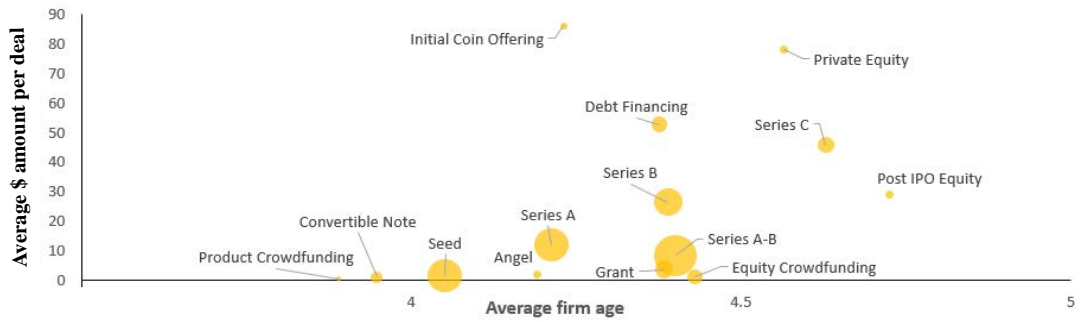


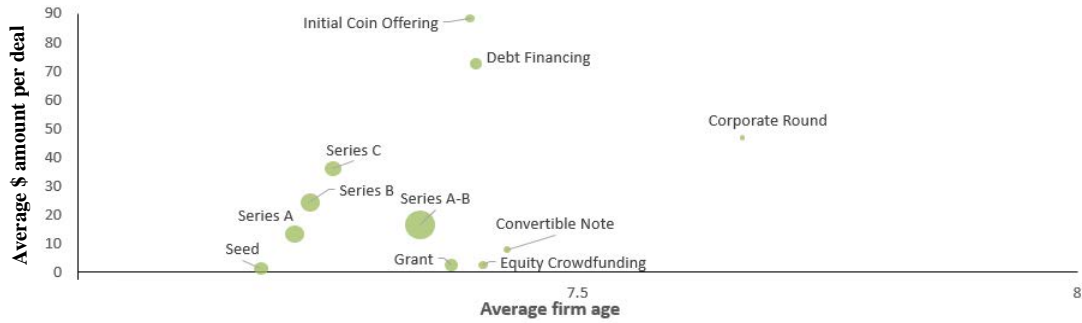
Figure 7. Weekly \$ amount raised by firm age



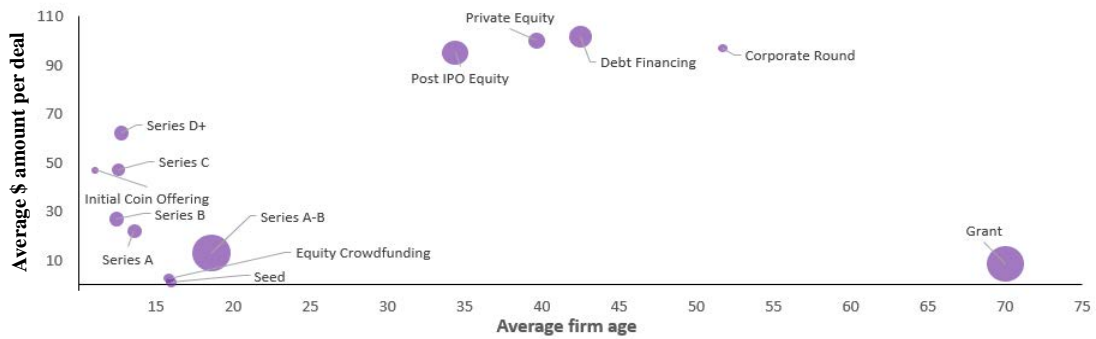
(a) New firms (0-3 year old)



(b) Young firms (3-6 year old)



(c) Established firms (6-9 year old)



(d) Mature firms (9+ year old)

Figure 8. Funding types and size as a function of firm age (observations fewer than 10 are not depicted).

Table 1: Summary Statistics

	EU					US				
	Mean	Median	75- percentile	95- percentile	N	Mean	Median	75- percentile	95- percentile	N
Aggregate Level:										
<i>Panel A: Whole Sample</i>										
# of countries/states	-	-	-	-	24	-	-	-	-	51
# of weeks	-	-	-	-	67	-	-	-	-	67
# of categories	-	-	-	-	2	-	-	-	-	2
\$ MM amount raised	38.04	0	3.67	174.81	3,216	38.12	0	5.14	159.82	6,834
# of deals	2.15	0	2	11	3,216	2.17	0	2	8	6,834
Unemployment	7.10%	-	-	-	-	4.00%	-	-	-	-
GDP (in billion)	298.77	-	-	-	-	398.28	-	-	-	-
CPI	109.95	-	-	-	-	114.09	-	-	-	-
Interest	-0.33%	-	-	-	-	1.56%	-	-	-	-
<i>Panel B: Sub-group by category</i>										
<i>Healthcare – financial:</i>										
\$ MM amount raised	33.67	0	1.85	135.14	1,608	28.72	0	4.60	150.00	3,417
# of deals	1.29	0	1	6	1,608	1.54	0	1	7	3,417
<i>Others:</i>										
\$ MM amount raised	42.41	0	6.52	198.22	1,608	47.52	0	5.51	38.61	3,417
# of deals	3.02	0	3	13	1,608	2.79	0	2	10	3,417
Deal Level:										
<i>Panel C: Whole Sample</i>										
\$ MM amount raised / deal	22.27	1.42	11.69	80	5,369	21.79	3	11.22	70	11,823
Firm age (exclude mature firms)	2.94	2.56	-	-	4,544	3.05	2.67	-	-	9,620
Firm age (whole sample)	8.66	3.13	-	-	5,369	10.46	3.51	-	-	11,823
<i>Panel D: Subgroup by category</i>										
<i>Healthcare-financial:</i>										
\$ MM amount raised / deal	31.76	2.36	21.99	102.39	1,692	22.09	4.2	15.51	82.00	4,379
Employee range midpoint average	705	-	-	-	1,692	464	-	-	-	4,379
<i>Others:</i>										
\$ MM amount raised / deal	17.90	1.19	4.43	70.54	3,677	21.61	2.37	9.6	60.00	7,444
Employee range midpoint average	1049	-	-	-	3,677	909	-	-	-	7,444

Table 1 Continued

	EU					US				
	Mean	Median	75- percentile	95- percentile	N	Mean	Median	75- percentile	95- percentile	N
<i>Panel E: Subgroup by funding stage</i>										
<i>Pre Stage:</i>										
\$ MM amount raised / deal	1.63	0.51	1.22	3.58	1,684	1.44	0.60	1.80	4.77	3,489
Employee range midpoint average	872	-	-	-	1,684	646	-	-	-	3,489
<i>Main Stage:</i>										
\$ MM amount raised / deal	18.59	3	21.71	77.44	2,630	14.43	5.85	15	50	6,066
Employee range midpoint average	741	-	-	-	2,630	685	-	-	-	6,066
<i>Late Stage:</i>										
\$ MM amount raised / deal	65.81	1.73	34.60	283.56	1,055	72.62	7.5	45	300	2,268
Employee range midpoint average	1394	-	-	-	1,055	1605	-	-	-	2,268
<i>Panel F: Subgroup by firm age</i>										
<i>New firms (0-3 year):</i>										
\$ MM amount raised / deal	8.31	0.65	2.10	43.16	2,607	7.76	1.40	4.50	27	5,293
Employee range midpoint average	32	-	-	-	2,607	26	-	-	-	5,293
<i>Young firms (3-6 year):</i>										
\$ MM amount raised / deal	17.07	2.32	17	70.86	1,415	18.82	4.20	13.32	60	3,057
Employee range midpoint average	50	-	-	-	1,415	64	-	-	-	3,057
<i>Pre-mature firms (6-9 year):</i>										
\$ MM amount raised / deal	31.45	4	35.06	100.59	522	44.48	7.5	25	125	1,270
Employee range midpoint average	123	-	-	-	522	153	-	-	-	1,270
<i>Mature firms (9+ year):</i>										
\$ MM amount raised / deal	69.50	7.65	50	277.23	825	46.54	6.3	25	200	2,203
Employee range midpoint average	3464	-	-	-	825	2986	-	-	-	2,203

Table 2. GDPPR impact on aggregate level weekly \$ raised amount

	(1)	(2)	(3)	(4)	(5)
	Dependent variable: Aggregate \$MM raised amount				In (1+\$ amount)
	Baseline	Baseline with top-coded	Adding linear trend	Excluding May	OLS
postGDP	180.955** (76.124)	42.993*** (16.157)	44.571*** (16.106)	44.880*** (16.947)	0.977*** (0.268)
EU_postGDP	-54.819* (31.597)	-11.574*** (4.436)	-11.668*** (4.366)	-16.659*** (5.689)	-0.265*** (0.098)
Unemployment	15.936* (9.242)	1.258 (1.733)	1.233 (1.683)	0.161 (1.752)	-0.004 (0.031)
GDP	0.044 (0.027)	0.006 (0.006)	0.007 (0.006)	0.007 (0.006)	0.000 (0.000)
CPI	-32.596* (17.709)	-6.924* (3.962)	-6.833* (4.004)	-5.866 (3.908)	-0.108 (0.075)
Interest rate	64.848** (28.565)	26.208*** (6.801)	26.534*** (6.910)	20.557*** (7.066)	0.231* (0.123)
Effect at post GDP mean	-13.896** (6.695)	-3.380*** (1.074)	-3.403*** (1.054)	-4.488*** (1.155)	
State FE	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes
Linear Trend	No	No	Yes	No	No
Top Coded	No	Yes	Yes	Yes	No
Observations	10,050	10,050	10,050	9,300	10,050
R-squared	-	-	-	-	0.578
F-test on pre-treatment (p-value)	0.475	0.101	0.108	0.142	0.116

Note: We group our sample into two different sub-categories (i.e., healthcare-financial, and others). The standard errors are clustered by state level (i.e., country level in EU and state level in US) in all specifications. We use 95 percentile value of \$ raised amount (i.e., 175 million) as the value to top coded in Tobit regression. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.

Table 3. GDPR impact on aggregate level # of deals

	(1)	(2)	(3)	(4)
	Dependent variable: # of deals			ln (1+ # of deals)
	Baseline	Adding linear trend	Excluding May	OLS
postGDPR	-0.183 (0.190)	-0.248 (0.205)	-0.164 (0.194)	-0.006 (0.074)
EU_postGDPR	-0.194*** (0.070)	-0.194*** (0.068)	-0.259*** (0.085)	-0.062* (0.033)
Unemployment	0.043 (0.035)	0.036 (0.037)	0.025 (0.041)	0.016 (0.012)
GDP	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
CPI	0.065 (0.064)	0.068 (0.065)	0.074 (0.062)	-0.006 (0.017)
Interest rate	-0.004 (0.109)	-0.018 (0.112)	-0.043 (0.111)	0.089*** (0.032)
State FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
Linear Trend	No	Yes	No	No
Observations	10,050	10,050	9,300	10,050
R-squared	-	-	-	0.733
F-test on pre-treatment (p-value)	0.585	0.582	0.670	0.117

*Note: We group our sample into two different sub-categories (i.e., healthcare – financial, and others). The standard errors are clustered by state level (i.e., country level in EU and state level in US) in all specifications. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.*

Table 4 GDPR impact on \$ raised amount per deal

	(1)	(2)	(3)	(4)
	Dependent variable: ln (\$ amount per deal)			
	Baseline	Baseline with top-coded	Adding linear trend	Excluding May
postGDPR	-0.127 (0.285)	-0.103 (0.272)	-0.125 (0.284)	2.446*** (0.322)
EU_postGDPR	-0.396*** (0.074)	-0.380*** (0.071)	-0.397*** (0.074)	-0.355*** (0.087)
Firm age	0.004** (0.001)	0.004*** (0.001)	0.004** (0.001)	0.004*** (0.001)
Unemployment	-0.002 (0.038)	-0.001 (0.038)	-0.007 (0.038)	-0.008 (0.039)
GDP	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)
CPI	-0.359*** (0.095)	-0.331*** (0.093)	-0.358*** (0.095)	-0.320*** (0.093)
Interest rate	0.079 (0.119)	0.054 (0.107)	0.072 (0.120)	0.117 (0.112)
State FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
Linear Trend	No	No	Yes	No
Top Coded	No	Yes	No	No
Observations	17,192	17,192	17,192	15,914
R-squared	0.382	0.391	0.382	0.386

*Note: We group our sample into two different sub-categories (i.e., healthcare - financial, and others). The standard errors are clustered by state level (i.e., country level in EU and state level in US) in all specifications except for specifications. We use 95 percentile value of \$ raised amount (i.e., 75 million) as the value to top coded in Tobit regression. ***, **, and * indicate significance at the 1%, 5%, and 10% levels.*

Table 5. GDPPR impact on healthcare-financial and others

	(1)	(2)	(3)	(4)	(5)	(6)
	Poisson regression on # of deals		Tobit regression on total \$ amount		OLS on ln (\$ amount per deal)	
postGDPR	0.092 (0.335)	-0.270 (0.198)	59.561*** (19.638)	25.822 (16.256)	-0.322 (0.441)	-0.111 (0.301)
EU_postGDPR	-0.209*** (0.060)	-0.186*** (0.090)	-22.009*** (5.824)	-2.934 (5.597)	-0.566*** (0.171)	-0.284*** (0.080)
Unemployment	0.078 (0.061)	0.027 (0.041)	1.054 (2.798)	1.306 (2.652)	0.005*** (0.002)	0.005*** (0.002)
GDP	-0.000 (0.000)	-0.000 (0.000)	0.011 (0.007)	0.003 (0.008)	-0.053 (0.084)	0.042 (0.068)
CPI	0.051 (0.080)	0.063 (0.068)	-7.149 (5.051)	-5.743 (3.701)	0.000** (0.000)	0.000*** (0.000)
Interest rate	-0.069 (0.158)	0.025 (0.100)	8.977 (9.218)	38.208 (8.589)	-0.590*** (0.144)	-0.245** (0.094)
Effect at post GDPPR mean			-5.216*** (0.891)			
Category	Healthcare- financial	Others	Healthcare- financial	Others	Healthcare- financial	Others
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Top coded	-	-	Yes	Yes	No	No
Observations	5,025	5,025	5,025	5,025	6,071	11,121
R-squared					0.410	0.372
F-test on pre- treatment (p-value)	0.100	0.985	0.102	0.302		

Note: The standard errors are clustered by state level (i.e., country level in EU and state level in US). We also top code for each funding stage by their 95-percentile value on \$ raised amount per deal, respectively. ***, **, * and * indicate significance at the 1%, 5%, and 10% levels.

Table 6. GDPR impact on total \$ amount per week per state by firm age

	(1)	(2)	(3)	(4)
Tobit regression on aggregate \$ amount				
postGDPR	8.500** (3.471)	13.228** (6.495)	42.128*** (15.108)	74.527** (31.400)
EU_postGDPR	-3.204* (1.863)	0.637 (2.718)	4.100 (5.050)	-58.154*** (11.703)
Firm age	0.783 (0.482)	0.372 (1.117)	1.170 (1.074)	-1.397 (3.596)
Unemployment	0.002 (0.001)	0.002 (0.002)	0.017*** (0.003)	-0.002 (0.014)
GDP	-1.796* (0.942)	0.126 (2.022)	-6.431** (2.691)	-21.523** (8.555)
CPI	2.948 (1.863)	2.817 (4.060)	12.041** (5.234)	50.506*** (16.939)
Interest rate	8.500** (3.471)	13.228** (6.495)	42.128*** (15.108)	74.527** (31.400)
Effect at post GDPR mean	-0.902** (0.422)			-7.099*** (0.248)
Firm age group	New Firm (0-3 y.o.)	Young Firm (3-6 y.o.)	Established Firm (6-9 y.o.)	Mature Firm (9+ y.o.)
State FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
Top Coded	Yes	Yes	Yes	Yes
Observations	5,025	5,025	5,025	5,025
F-test on pre- treatment (p-value)	0.319	0.102	0.509	0.130

*Note: The standard errors are clustered by state level (i.e., country level in EU and state level in US). We also top code for each funding stage by their 95-percentile value on \$ raised amount per deal, respectively. ***, **, and * indicate significance at the 1%, 5%, and 10% level.*

Table 7. GDPR impact on # of deals per week per state by firm age

	(1)	(2)	(3)	(4)
	Poisson regression on # of deals			
postGDPR	-0.681** (0.266)	0.256 (0.317)	0.984* (0.580)	0.732 (0.540)
EU_postGDPR	-0.211** (0.087)	0.022 (0.085)	0.047 (0.188)	-0.350*** (0.105)
Firm age	0.004 (0.038)	0.115* (0.061)	0.062 (0.066)	0.029 (0.056)
Unemployment	-0.000** (0.000)	-0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
GDP	-0.081 (0.064)	0.044 (0.079)	0.064 (0.092)	0.162 (0.152)
CPI	0.415*** (0.086)	0.120 (0.125)	0.082 (0.182)	-0.100 (0.229)
Interest rate	-0.681** (0.266)	0.256 (0.317)	0.984* (0.580)	0.732 (0.540)
Firm age group	New Firm (0-3 y.o.)	Young Firm (3-6 y.o.)	Established Firm (6-9 y.o.)	Mature Firm (9+ y.o.)
State FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
Observations	5,025	5,025	5,025	5,025
F-test on pre-treatment (p-value)	0.116	0.586	0.160	0.193

*Note: The standard errors are clustered by state level (i.e., country level in EU and state level in US). We also top code for each funding stage by their 95-percentile value on \$ raised amount per deal, respectively. ***, **, and * indicate significance at the 1%, 5%, and 10% level.*

Table 8. GDPR impact on \$ raised per deal by firm age

	(1)	(2)	(3)	(4)
	OLS on ln (\$ amount per deal)			
postGDPR	0.379 (0.397)	-0.420 (0.342)	-0.034 (1.057)	-1.763** (0.823)
EU_postGDPR	-0.271* (0.159)	-0.314* (0.176)	-0.773** (0.343)	-0.436 (0.284)
Firm age	0.212*** (0.027)	0.161*** (0.047)	0.012 (0.051)	-0.003*** (0.001)
Unemployment	0.051 (0.071)	0.032 (0.135)	-0.119 (0.141)	-0.272* (0.159)
GDP	0.000*** (0.000)	-0.000 (0.000)	0.000*** (0.000)	0.000*** (0.000)
CPI	-0.185* (0.095)	-0.313** (0.147)	-0.733** (0.344)	-0.666* (0.344)
Interest rate	0.111 (0.138)	0.190 (0.198)	0.288 (0.428)	0.577 (0.433)
Firm age group	New Firm (0-3 y.o.)	Young Firm (3-6 y.o.)	Established Firm (6-9 y.o.)	Mature Firm (9+ y.o.)
State FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
Top Coded	No	No	No	No
Observations	7,900	4,472	1,792	3,028
R-squared	0.414	0.374	0.408	0.275

*Note: The standard errors are clustered by state level (i.e., country level in EU and state level in US). We also top code for each funding stage by their 95-percentile value on \$ raised amount per deal, respectively. ***, **, and * indicate significance at the 1%, 5%, and 10% level.*

Table 9. Back-of-the-envelope estimates of a range on the job losses incurred by 0-3 year-old firms

Variables	(1)	(2)	(3)	(4)	(5)
	Poisson Regression # of deals	OLS \$ amount per deal	Tobit Regression Aggregate \$ amount	Back-of-the-envelope calculations # of deals * \$ amount/deal Aggregate \$ amount	
EU_postGDPR	-0.211** (0.087)	-0.271* (0.159)	-3.204* (1.863)		
% reduction in # of deals	19.02%				
Average \$mm % change per deal		27.1%			
Aggregate \$mm amount change			0.902		
\$mm raised/deal (post GDPR)				3.321	
# of deals (post GDPR)				690	
Aggregate \$mm per week per state					14.41
Estimated total \$mm reduction				1589.77	1217.70
Average firm age				1.21	1.21
Annualized \$mm/employee lower bound				0.123	0.123
Annualized \$mm/employee upper bound				1.020	1.020
Job loss lower bound (post GDPR)				1,559	1,194
Job loss upper bound (post GDPR)				12,899	9,880
Ratio of Jun–Sep 2017 deal activity to that in the entirety of 2017				0.331	0.331
Annualized job loss lower bound				4,705	3,604
Annualized job loss upper bound				38,931	29,819
Subsample employee # lower bound				88,092	88,092
Subsample employee # upper bound				266,352	266,352
% job loss calculation lower bound				5.34%	4.09%
% job loss calculation upper bound				14.62%	11.20%

Note: The estimates of the effects on the \$ mm raised per deal, # of deals, and aggregate \$mm in the post-GDPR period are used for back-of-the-envelope calculations of the effect on jobs. First, an estimate decrease in total \$mm invested due to GDPR is calculated. Next, ventures founded on or after 2015 are used to calculate the average \$ amount raised per employee (total \$ amount raised by each firm divided by the firm's employee range). Those bounds are averaged and annualized. A job-loss range in 2018 Q3 is obtained by dividing the total \$mm reduction by the annualized \$mm/employee range. We extrapolate 2018 Q3 to an entire calendar year by using the ratio of deals in 2017 Q3 to 2017 in order to annualize the job loss bounds. In the final step, these bounds are given as % of the total range on the # of employees for ventures founded on or after 2015.