How Do Data Affect Firms' Innovation Strategies: Evidence From the General Data Protection Regulation *

Lydia Wang†

November 5, 2024

[Please click here for the most recent version!]

Abstract

I examine how firms' customer data availability influences their managers' innovation resource allocation between exploitative (incremental) and exploratory (radical) innovation. I use the European Union (EU) General Data Protection Regulation (GDPR) as a plausibly exogenous shock to customer data availability in estimating a difference-in-differences specification. I employ a patent-based innovation output measure and develop and test a novel innovation input workforce-based measure. I find that, upon a negative shock to customer personal data availability, firms allocate more innovation resources to exploration relative to exploitation. The results are stronger for firms in B2C versus B2B industries and firms with more volatile customer demand. My results suggest that higher customer data availability leads firms to focus more on exploitative innovation, deterring the exploratory innovation crucial to economic growth.

Keywords: digital economy, data, innovation, exploitation, exploration, GDPR
JEL Codes: D83, K20, L86, O00, O14, O30, O31, O33

*I am deeply grateful to my dissertation committee for their unwavering support, encouragement, and valuable feedback: Chris Armstrong (co-chair), Ron Kasznik (co-chair), Brandon Gipper, Stephen Glaeser, and Jinhwan Kim. Thank you to Anat Admati, Jonathan Ashtor, Mary Barth, Robert Bartlett, Anne Beyer, Beth Blankespoor, Qi Chen, Susan Cherry, Jung Ho Choi, Hans Christensen, Ed deHaan, Stefan Droguett, Lucile Faurel, George Foster, Richard Frankel, George Gamble, Andrew Glover, João Granja, Steve Grenadier, John Hand, Junyoung Jeong, Chad Jones, Joseph Kalmenovitz, John Kepler, Nicolas Konialidis, Chris Kontz, Wayne Landsman, Mark Lang, Charles Lee, Seung Lee, Mark Lemley, Becky Lester, Iván Marinovic, Maureen McNichols, Ilan Morgenstern, Fabian Nagel, Suzie Noh, Marcel Olbert, Joe Piotroski, Huggy Rao, Devin Shanthikumar, Kevin Smith, Christopher Stewart, Xuan Su, Kristen Valentine, Rodrigo Verdi, Pei Wang, Xi Wu, Winston Xu, Li Yang, Christina Zhu, and workshop participants at Stanford, UNC Chapel Hill, the DFSI Conference, and the AAA/Deloitte Foundation/Michael J. Cook Doctoral Consortium, for very helpful advice, information, and discussions. I gratefully acknowledge the financial support of the Stanford Graduate School of Business. All errors are mine.

†Stanford Graduate School of Business. Email: lydwang@stanford.edu

1 Introduction

In the past 15 years, the digital economy has grown two and a half times faster than global GDP (*Bloomberg Tax*, 2024). In this rising digital economy, "data is the new oil" (*The Economist*, 2017; see Appendix C). Data is thought to benefit corporate performance, such as through improving investment and operating efficiency as well as revenue (e.g., Zhu, 2019; Ferracuti et al., 2024; Goldberg et al., 2024). Some commentators say that data can also contribute to innovation (e.g., Deming, 2021; Veldkamp, 2023b). Customer data, which is one of the most valuable intangible assets—and thus type of data—in the digital economy, provide insights for e-commerce and advertising, even now serving as input to machine learning (i.e., big data) models (Kepler et al., 2024). Yet, despite the many purported commercial benefits, one could argue that customer data might deter the more radical type of innovation crucial to economic growth (Schumpeter, 1942; Romer, 1990).

I examine how the availability of customer data to firms influences their managers' resource allocation between exploitative (incremental) and exploratory (radical) innovation. Customer data yield information about customer preferences that firms can use to tailor future offerings to increase sales and improve customer satisfaction and retention (Hagel and Rayport, 1997). The innovation literature distinguishes between exploitation and exploration: The former refers to the more certain, incremental improvement of existing products, services, or processes; the latter refers to the more uncertain, riskier search for and development of novel radical ideas for potentially groundbreaking products, services, or processes (March, 1991; Manso, 2011). Taking Apple Inc. as an example, its frequent upgrading of its iPhone is exploitative. In contrast, its development of the Vision Pro—which was unlike any other technology at the time—is exploratory. Firms' implemented innovation strategies can be arranged on a spectrum, from pure exploitation on one end to pure exploration on the other end (March, 1991).

I argue that higher availability of customer data leads firms to allocate more of their finite innovation resources toward exploitative innovation, relative to exploratory innovation. The outcomes of innovation resource allocation decisions can be crucial for firms' profitability and long-term survival. A firm's investment in innovation is restricted by factors, including, but not limited to, information constraints and management control (Romer, 1993; Jones, 2023; Glaeser and Lang, 2024). Firms thus cannot fund all positive net present value innovation projects and will consider which investments to fund relative to each other. However, firms also have limited foresight about the expected risk-adjusted payoffs of their innovation projects (March, 1991; Manso, 2011; Bubeck and Cesa-Bianchi, 2012; Balsmeier et al., 2017).

Insights from customer data can improve and reduce risk and uncertainty around expected payoffs from firms' offerings (Veldkamp, 2023b). However, an asymmetry exists in that customer data, a byproduct of customer interactions with a firm's existing offerings, inform more about exploitation than about exploration, because customer data on existing offerings cannot reveal as much about the prospects of radically new and nonexistent offerings. In other words, exploitation is more sensitive to changes in customer data than exploration is. However, the potential to learn more from implementing more uncertain projects, the "high risk high return" idea from the real options literature, and the ability to implement an alternative "spray and pray" approach for gathering customer insights, may temper the strength of this asymmetric influence of customer data on exploitation versus exploration (Holmström, 1979; Ewens et al., 2018).

To test my hypothesis and its causality, I use a difference-in-differences identification design that exploits the passage of the European Union (EU) General Data Protection Regulation (GDPR) in 2016. I use GDPR as a negative and arguably exogenous change to affected firms' customer data availability. GDPR regulates and restricts the collection, processing, and storage of European residents' personal data to protect their privacy. I choose this regulation as my shock because it is the "toughest privacy and security law in the world" and "global north star for privacy law" (European Data Protection Supervisor; Wolford). GDPR requires any entity (e.g., firm) in the world handling the personal data of EU residents to acquire their consent to do so. GDPR thus decreased affected firms' customer data availability by decreasing the volume and degree to which the data accurately represents firms' customer base (e.g., if certain demographics are more or less likely to grant consent for handling personal data) (Dubé et al., 2024). I determine whether a firm is affected or treated based on whether it generated significant revenues from the EU at the time of shock.

I conduct two separate main tests, the results of which support my hypothesis. First, I use a patent-based measure to study the output of firms' innovation resource allocation choices. I construct my patent-based measure based on the measure first proposed and studied by Balsmeier et al. (2017), which represents the similarity between the technological patent portfolio profiles of a firm in the current year and the prior year. The profile is the distribution of the number of patents filed in "new to the firm" technology classes versus "known to the firm" technology classes, in a given year. The technological profiles of firms that exploit more will more resemble their prior year profiles; those that explore more will less resemble their prior year profile. The innovation output test sample consists of firm-year observations in 2010–2021 for U.S. firms that operate on e-commerce platforms or have analytics technologies on their websites.

Second, I develop and test a novel workforce-based measure to study firms' resource allocation toward inputs to innovation by studying their human capital investment decisions. I do so because simply investing in innovation does not guarantee successful outcomes (Menguc and Auh, 2008; Grabner et al., 2018). However, the literature does not propose proxies for firms' resource allocation toward inputs to innovation. Therefore, I develop a novel measure, and I do so based on human capital investment because firms increasingly rely on knowledge workers to create value via innovation (Rubera and Kirca, 2012; Zumbrun, 2016; Glaeser et al., 2023). My measure is the ratio of total salaries paid to exploratory innovation role employees to total salaries paid to exploitative innovation role employees, in a given firm and year. The innovation input test sample consists of firm-year observations in the shorter period of 2013–2021, due to the greater sensitivity and exposure of firms' human capital investment to confounding noise and external factors with a longer pre-period.

In line with my hypothesis, I find that lower customer personal data availability increases firms' innovation resource allocation to exploration relative to exploitation, for both the innovation output and input tests. Specifically, the technological patent portfolio profiles of affected firms less resemble their prior year profiles for the output test, and the ratio of total salaries paid to exploratory, over exploitative, innovation role employees, increases for affected firms for the input test. Thus, my study shows that firms that experience lower customer data availability allocate a greater proportion of their innovation resources to exploration, a key driver of economic growth.

In addition, I conduct cross-sectional tests, which support the argument that customer data drives my results. I find that firms in B2C industries react more strongly to GDPR than firms in B2B industries, consistent with B2C firms' greater reliance on the personal data of individual customers. I also find that firms operating in industries with more volatile customer demand react more strongly to GDPR, consistent with these firms handling customer data insights that amortize or lose relevance more quickly.

This paper contributes to the literature in several ways. First, it contributes to the nascent literature on digital data. Despite data's significance in the digital economy, little empirical evidence is available on the topic. Accounting academics have begun to study the consequences of operating in the digital space (e.g., Armstrong et al., 2023; Chen and Srinivasan, 2024; Ferracuti et al., 2024). Finance and economics papers are working on valuing data and studying the economics of data and their nonrival nature (e.g., Jones and Tonetti, 2020; Veldkamp, 2023b). My paper contributes to this early literature on data.

Second, this paper extends the growing accounting literature on innovation. (See Glaeser and Lang, 2024, for a review.) Studies have investigated how innovation decisions can be influenced by contracting, delegation, or disclosure policy (e.g., Manso, 2011; Dutta and Fan, 2012; Laux and Stocken, 2018; Laux and Ray, 2020; Baldenius and Azinovic-Yang, 2023; Chen et al., 2024). However, my paper is the first to study the information value of customer data and its influence on firms' innovation strategies.

Third, I contribute to the literature on the implications of firms' internal information environment for managerial decision-making. Research has shown that internal information facilitates various managerial decisions, such as tax planning, hiring, firm transparency, and investment (e.g., Gallemore and Labro, 2015; Heitzman and Huang, 2019; Samuels, 2021; Ferracuti, 2022; Binz et al., 2023; Baldenius and Azinovic-Yang, 2023). My paper studies the implications of internal information on managerial decisions regarding innovation.

The paper is organized as follows. Section 2 provides background and hypotheses development. Section 3 delves into the sample and research design. Section 4 provides empirical results, robustness checks, and discussion. Section 5 concludes.

2 Background and hypothesis development

2.1 The information value of customer data

Customer data informs about expected payoffs to innovation. Data, a type of information, can broadly be described as "observations... [in] digital form... from which knowledge can be drawn" (Jones and Tonetti, 2020; Veldkamp, 2023b). This knowledge constitutes the data's value, where knowledge takes the form of strategic insights firms can use to make marginal improvements to probabilities of commercial success or payoffs (Veldkamp, 2023b; International Valuation Standards Council, 2024). These payoffs include those from firms' investments in innovation projects, where innovation "requires a deep understanding of whether customers need or desire that invention" (Keeley et al., 2013). Customer data provides this understanding that firms can apply to tailor future offerings to improve sales, customer satisfaction, and customer retention (Hagel and Rayport, 1997).

The traditionally difficult mass collection and analysis of customer data is facilitated by the advent, improvement, and now widespread use of digital technologies. These technologies, such as analytics technologies, track visitors online through code. Whenever someone visits a website, this code sends the visitor's personal information to the website's servers.

Customer data can include many types of visitor personal information, including pages viewed, time spent on pages, clicks, user flows (i.e., individuals' paths among different pages in a session), whether users are new versus returning, demographics, and even the types of devices or browsers through which users visit (Goldberg et al., 2024). Google Analytics is an example of an analytics technology provider.

The data from tracking potential customers' behavior and demographic information yield customer demand insights in aggregate (Farboodi and Veldkamp, 2022). For instance, through the number of repeat viewings or the time spent on a page, firms can identify product features that grab customers' attention. Through analyzing the number of pages visited per session, firms can gauge overall interest in their offerings. Comparing the exit pages of sessions that end in a purchase versus those that do not can help firms identify disliked products and product features. Likewise, a user's flow through product pages can help firms figure out which products customers find complementary or related to each other. These insights about customer preferences, as well as their potential variation among different demographic groups, can help firms tailor future offerings. (See Appendix D for a discussion by Fossil Group, a firm in my sample.)

Firms can use insights from customer data to improve and reduce the riskiness and uncertainty of expected payoffs from their offerings (Farboodi and Veldkamp, 2022; Veldkamp, 2023b).[1] Data-derived insights provide managers with clues for ways to better satisfy customer preferences. Even if customer data foretells negative payoffs, firms can use the derived insights to determine disliked features to avoid in future offerings, thus increasing payoffs. The riskiness of payoffs decreases because insights are based on hard data, rather than on random gambles. Customer data insights also help managers discriminate among potential avenues for exploitation or exploration, winnowing investment opportunities to only the most attractive risk-return trade-offs. In addition, customer data insights can reduce uncertainty by improving the precision of forecasted payoffs, because they are providing more informa-

[1] Expected payoffs already take into account firms' data collection and processing costs. I assume firms expect net positive value from customer data insights, or else firms would have avoided handling the data.

tion about the payoffs' characteristics (Dichev and Qian, 2022; Farboodi and Veldkamp, 2022; Veldkamp, 2023b).

Digital customer data is also unique in that it can yield valuable insights about unmet needs and preferences even if potential customers do not purchase or transact with a firm, a phenomenon less likely without digital technologies (Thomas, 2024). For instance, without digital customer data, a firm may infer the lack of revenue from a product as a lack of customer demand for the product. However, customers may actually be very interested in the product, as reflected by clicks and spending much time on the product's webpage. Thus, customer data can help firms discover these otherwise hidden insights that spur deeper digging into why customers do not make purchases despite their interest, which can then in turn potentially improve their payoffs.

2.2 The General Data Protection Regulation (GDPR)

The European Union (EU) General Data Protection Regulation (GDPR) regulates the collection, processing, and storage of EU residents' personal data to protect their privacy. It passed into law on April 27, 2016 and went into effect on May 25, 2018. The GDPR applies to any entity that handles EU residents' personal data, even if it resides outside the EU. A guiding principle of the GDPR is data minimization, which limits firms' ability to handle their EU customers' personal data due to needing their consent (Goldberg et al., 2024).[2]

The GDPR thus decreased the volume and representativeness of customer personal data available to affected firms, which lowered the depth and accuracy of the strategic insights

[2]The GDPR declares that "personal data [should] be: processed lawfully, fairly and in a transparent manner in relation to the data subject" under its principle of "lawfulness, fairness and transparency" (Regulation (EU) 2016/679). According to GDPR, personal data is "any information relating to an identified or identifiable natural person," where such a person is also called a "data subject" (Regulation (EU) 2016/679). "The data subjects are identifiable if they can be directly or indirectly identified, especially by reference to an identifier such as a name, an identification number, location data, an online identifier or one of several special characteristics, which expresses the physical, physiological, genetic, mental, commercial, cultural or social identity of these natural persons. In practice, these also include all data which are or can be assigned to a person in any kind of way" (Regulation (EU) 2016/679). I asked about and clarified with Anu Talus, current chair of the European Data Protection Board, that "processing" data, according to the GDPR text, refers to any form of handling data, such as collecting, cleaning, analyzing, and selling data.

firms could derive from this data. The value of strategic insights, where value refers to the marginal improvement in the probability of firms' commercial success, decreases with a lower volume of available data (International Valuation Standards Council, 2024). The accuracy, and thus value, of strategic insights drawn from data also decreases as the available data become less representative of a firm's potential customer base (e.g., if certain demographics are more or less likely to grant consent for personal data handling) (Dubé et al., 2024).[3]

The GDPR is described by the European Commission as the "toughest privacy and security law in the world" and has become the "global north star for privacy law" (Wolford). It replaced the 1995 EU Data Protection Directive, which was established in the Internet's infancy. The GDPR answered calls for a single EU-wide comprehensive personal data protection regulation with global application as the digital economy grew. Hence, the GDPR's effect on affected firms' customer data availability was significant.

Evidence validating the GDPR's significance in decreasing firms' data availability lies in literature, as well as anecdotally. Non-consent rates after the GDPR have ranged between 4 and 13 percent (Goldberg et al., 2024). In addition, in response to the GDPR, firms decreased data storage and processing, becoming less "data-intensive" (Demirer et al., 2024; Goldberg et al., 2024). Excerpts from annual reports also anecdotally validate the GDPR's effect on firms' handling of EU residents' personal data (see Appendix E, Exhibit 1).

Anecdotal evidence from firms' conference calls also validates the GDPR as a significant negative shock to firms' customer data availability. As an example, Meta Platforms Inc., a multinational firm reliant on its collected personal data for earning advertising revenues, predicted and observed a negative trend in monthly active users in Europe upon the implementation of the GDPR in May 2018. Mark Zuckerberg, CEO, stated in the 2018 Q2 earnings call that Meta saw monthly active users go "down by about 1 million people" after the GDPR implementation (see Appendix E, Exhibit 2).

[3] If the data available to a firm does not accurately capture the distribution of the individual characteristics of its customer base, the firm could draw erroneous insights from the data. Erroneous insights, such as inferring incorrectly that the customer base consists primarily of a certain age range, as an example, could lead firms to inefficiently target certain demographic groups and thus hurt their performance.

Firms affected by the GDPR would have begun responding to the regulation upon its passage in 2016. For example, Activision Blizzard states awareness and anticipation of the GDPR's implementation in 2018 as early as in its 2016 10-K, the year the GDPR passed. Annual report excerpts also validate that firms expected and feared penalties from enforcement if they did not abide by the regulation. (See Appendix E, Exhibits 3 and 4.)

2.3 Hypothesis development

Customer data availability could influence the innovation strategies of firms. Firms allocate their finite innovation resources between exploitative and exploratory innovation, the outcomes of which can be crucial for their profitability and long-term survival. A firm's investment toward innovation is restricted by factors, including, but not limited to, information constraints and management control (Romer, 1993; Jones, 2023; Glaeser and Lang, 2024). Firms thus cannot fund all positive net present value innovation projects and will consider which investments to fund relative to each other. However, firms also have limited certainty about the expected risk-adjusted payoffs of their innovation projects (March, 1991; Manso, 2011; Bubeck and Cesa-Bianchi, 2012; Balsmeier et al., 2017).

As described in Section 2.1, insights from customer data can improve and reduce the riskiness and uncertainty of firms' expected payoffs from their offerings (Hagel and Rayport, 1997; Jones and Tonetti, 2020; Veldkamp, 2023b).

However, an asymmetry exists in that customer data informs more about exploitation than about exploration. Customer data, a byproduct of customer interactions with a firm's existing offerings, inform more directly about payoffs from potential improvements to these existing offerings (i.e., exploitation) than about the prospects of radically new and nonexistent offerings (i.e., exploration). As March (1991) speculated: "The certainty, speed, proximity, and clarity of feedback ties exploitation to its consequences more quickly and more precisely than is the case with exploration." In other words, exploitation is more sensitive to customer data changes than exploration is. In addition, as is the standard assumption in the literature, managers are risk-averse and thus maximize their utility by investing in innovation strategies with the highest certainty equivalence values (i.e., highest expected payoff with the lowest risk) (Holmström, 1979). This leads to my hypothesis:

***H**: Firms allocate more of their finite innovation resources toward exploitation relative to exploration with higher customer data availability.*

On the other side of the same coin, firms allocate more of their finite innovation resources toward exploration relative to exploitation with lower customer data availability.

Some may argue that the magnitude of this hypothesized effect could be attenuated by several factors. First, customer data, as a type of information, could inform more about exploration than about exploitation because less is known about exploration *ex-ante*. However, what distinguishes customer data from information more broadly is that, again, the insights from customer data relate directly and specifically to existing offerings rather than to a much larger set of more orthogonal ideas for offerings that do not yet exist.

Some may also push against the argument that firms would prefer the type of innovation that becomes less risky and more certain with data availability. Such an argument may arise from the real options literature, where the option to invest in a project is analogous to a financial call option (Dixit and Pindyck, 1994). It is a well-known property of real options that they are more valuable the greater the underlying risk or uncertainty of the project's payoff. However, again, managerial decision-makers are risk-averse, as is the standard assumption in the literature. Thus, risk aversion will temper managers' utility from investing in riskier or more uncertain innovation projects.

Some may also raise the possibility that firms could seek to gather more information about innovation projects they are less certain about, by implementing a "spray and pray" approach (Ewens et al., 2018). In this approach, firms invest a little in each of a wider selection of offerings to gather clues about customer demand, assuming firms will abandon many of these offerings. One way of doing so could be through releasing prototypes, beta offerings, or "interest screeners" for concepts of future products (e.g., concept cars, some of which are never mass produced). However, customer demand insights gained from this smaller-scale approach will be noisier than insights gained from customer data derived from customer interactions with established existing offerings that are available on a larger scale.

I test my hypothesis using the GDPR as a negative shock to firms' customer data availability. The next section, Section 3, dives into the research design behind my tests.

3 Sample and research design

3.1 Sample selection

The sample consists of observations from U.S. firms that have e-commerce platforms or analytics technologies on their websites. These platforms and technologies enable the collection and processing of customer personal data. I use BuiltWith to identify firms that implement these technologies at least three years prior to the GDPR passage in 2016. I study U.S. firms and do not include EU firms, because studying firms within one country ensures more comparable samples and alleviates concerns commonly raised about cross-country studies.

The sample varies based on which of the two main tests I execute, where the tests examine firms' allocation of their finite innovation resources between exploitative and exploratory innovation before and after the GDPR. For the first main test examining the output of firms' innovation resource allocation choices, I construct a patent-based measure based on the work of Balsmeier et al. (2017) with patent data obtained from PatentsView and Kogan et al. (2017). The sample consists of firm-year observations in 2010–2021.

For the second main test that examines firms' input to innovation, I develop and test a novel measure using workforce data from Revelio Labs, which collects and standardizes hundreds of millions of online public profiles and résumés to construct aggregate measures of historical workforce composition (Revelio Labs). The sample for this innovation input test consists of firm-year observations in 2013-2021. The shorter pre-period minimizes the influence of the greater sensitivity and exposure of firms' human capital investment to confounding noise and external factors. I also exclude less specialized firms, because these firms' focus on exploitation versus exploration are less likely to manifest through clearly delineated roles, and employees are more likely to balance responsibility for both types.

I retain observations that have non-missing values for my variables of interest. This yields a final sample of 4,662 and 2,838 firm-year observations for my innovation output and input tests, respectively. For both samples, I exclude firms that generate significant revenues from advertising. I do so to alleviate the concern that the results could be driven by the GDPR's negative shock to personal data, which forced firms with significant advertising revenues as one of their primary "products" to innovate in a more exploratory way to survive. I also exclude early-stage startups by not including private firms in my sample. The only type of innovation early-stage startups can implement is exploration, due to most of them having no material existing offerings. The sample selection procedures are outlined in Appendix A.

3.2 Variable measurement

3.2.1 Key dependent variable

The key dependent variable for my analysis captures firms' allocation of their finite innovation resources between exploration and exploitation.

For the innovation output test, I use a patent-based measure as my proxy. Characteristics of patents can proxy for multiple elements of a firm's innovative activities (Hall et al., 2001; Lerner and Seru, 2022). Balsmeier et al. (2017) introduce a patent-based proxy to measure a firm's resource allocation to exploitation versus exploration. This proxy differentiates between patents filed in "new to the firm" versus "known to the firm" technology classes. Balsmeier et al. (2017) observe the distribution of the number of patents (in year of filing) per technology class and firm. Following Jaffe (1989), they then calculate the similarity between the distribution of patents across technology classes filed by a firm in the current year and the same firm's distribution of patents in the prior year. The technological profiles of firms that exploit more will more resemble their prior year profiles; those that explore more will less resemble their prior year profiles. I follow the Balsmeier et al. (2017) measure, but I use one minus their proxy and then scale it up by 100 for my measure, *Exploratory Patent Portfolio*. My proxy thus ranges from 0 to 100, where values closer to 100 represent a more exploratory, rather than exploitative, patent portfolio.

Results for the innovation output measure may manifest only after a delay, consistent with the slow-moving nature of innovation (Galasso and Schankerman, 2015; Glaeser et al., 2020). However, since my sample period extends three years after the GDPR and thus spans the two to three year lag in patent outcomes noted in the literature, this lag should not significantly affect the results.

I also test for firms' innovation input using a novel measure I develop that considers human capital investment decisions. I study firms' innovation input because simply investing in innovation does not guarantee successful outcomes (Menguc and Auh, 2008; Grabner et al., 2018). However, the literature does not propose proxies for firms' resource allocation toward inputs to innovation.[4] Therefore, I develop a novel measure, and I do so based on human capital investment because firms increasingly rely on knowledge workers to create value via innovation (Rubera and Kirca, 2012; Zumbrun, 2016; Glaeser et al., 2023).

The input proxy *Explore-to-Exploit Salary Ratio* is the ratio of total salaries paid to exploratory innovation role employees to total salaries paid to exploitative innovation role employees, in a given firm and year. More money spent on salaries of innovation job roles represents a firm's intent to invest more resources in these innovation strategies. I manually identify job roles, broken down into 150 types by Revelio Labs, as exploratory, exploitative, or neither. I mark jobs as "exploratory" if they value experimentation, give room for failure in exploring new ideas, involve potentially nonlinear processes, and are characterized by more

[4]I do not use R&D expenses as a proxy for studying firms' innovation resource allocation between exploitation and exploration. Mezzanotti and Simcoe (2023) find that R&D itself consists of both exploratory and exploitative innovation activities. They describe the 'research' or 'R' of R&D as having a more exploratory nature, while 'development' or 'D' has a more exploitative nature. However, no data with this granular breakdown of R&D exists at the firm level.

uncertainty around their deliverables or results. These jobs might focus more on developing differentiating offerings in new markets. I mark jobs as "exploitative" if they involve more predictable and normalized processes and value reliable execution to improve production efficiency and financial performance of existing products in current markets. I mark jobs as neither if they could encourage both exploration and exploitation or if neither exploratory nor exploitative innovation appears integral to them (see Appendix F for my categorization) (Keeley et al., 2013). See Appendix G for examples of job postings for exploratory and exploitative innovation job roles for a firm in my sample, Coca-Cola.

3.2.2 Key independent variable

The key independent variable for my analysis is the availability of customer data to firms. The analysis uses a difference-in-differences research design that uses the GDPR as a negative shock to customer data availability. I determine whether a firm is affected or treated (i.e., the variable *EU Rev Exposure* takes on the value 1) based on whether it generated significant revenues from the EU in 2015, the year leading up to the GDPR passage. Firms likely anticipated the GDPR in 2015, leading up to its passage in 2016, because the European Council agreed to a general approach to the GDPR in 2015, implying a high likelihood of passage soon after. I determine treatment based on segment disclosures in 10-Ks, collected from WRDS Compustat Segments. *Post-GDPR* is a dummy variable that indicates whether a firm-year observation in my sample is in or after the GDPR passage year of 2016, or instead before 2016. *Affected by GDPR* is the interaction term of *EU Rev Exposure* with *Post-GDPR*, the coefficient of which captures my main result of interest.

The GDPR satisfies the conditions necessary for the feasibility and validity of my difference-in-differences identification design. The innovation measures I study are not confounded or influenced by other shocks at the time of the GDPR passage or by factors other than changes in customer data availability. The shock is plausibly exogenous because the mission of GDPR was to protect EU individuals' privacy and not to influence innovation.

3.2.3 Discussion about using the GDPR shock for analyses

In this section, I address potential concern about whether other relevant shocks coincided with the passage of GDPR. First, the GDPR does not coincide with the passing of GDPR-like laws in other countries, as these countries followed in the footsteps of the GDPR, using it as their "north star" later on (European Data Protection Supervisor; Wolford).

U.S. state-level privacy laws are also not a concern for my analysis. Within my sample period, only a few U.S. states had a comprehensive GDPR-like data privacy law that would have lowered customer data availability for relevant firms (e.g., CCPA in California).[5] Furthermore, unless a business with an online presence restricts its customer base to only these few states, which is unlikely, firms can also collect data on customers in the states without GDPR-like laws. Therefore, even if the state-level comprehensive data privacy laws prohibited any collection of customers' personal data in the extreme case, concerns about U.S. state laws for identification would still not be significant.

The Privacy Shield is not a concern for my results either. It coincides with GDPR passage but has the same effect on the value of firms' customer data. The Privacy Shield more strictly regulated cross-border data transfers out of the EU starting in 2016. Because this regulation discouraged the aggregation of separate cross-border datasets, where analysis on bigger consolidated datasets generates more, richer, and more accurate insights, the Privacy Shield decreased the value of insights firms could generate from their customer data, due to data fragmentation across international borders. Therefore, the Privacy Shield increased the potency of the negative effect on firms' value of their customer data.

[5] CCPA, or the California Consumer Privacy Act, was the first comprehensive state data privacy law passed in the United States. It was passed in 2018 and took effect in 2020. This law and any similar law that followed in other states, should not affect my U.S.-only sample of treatment and control firms differentially.

3.2.4 Control variables

I control for factors that may relate to a firm's allocation of finite innovation resources toward exploitation versus exploration, or to a firm's value of its available customer data, to increase the precision and accuracy of my estimates. For both my innovation output and input tests, I control for financing frictions and investment opportunities through the variables *Firm Size*, *Book-to-Market*, and *Leverage*, as the literature has done (e.g., Balsmeier et al., 2017; Glaeser and Landsman, 2021; Kim and Valentine, 2021). For my patent-based innovation output test, I additionally control for *R&D*, *Sales*, *PPE*, and *Employees*, following Balsmeier et al. (2021). I construct the variables using data from CRSP/Compustat. See Appendix B for variable definitions and constructions. Firm fixed effects control for time-invariant firm characteristics that might affect outcomes, including industry characteristics. Year fixed effects account for macro factors within each annual cross-section, e.g., economic indicators or consumer sentiment and consumer propensity to spend.

3.3 Research design

The primary empirical analyses employ a difference-in-differences design around the passage of the GDPR. The model is as follows:

$$INNOVATION\ TYPE\ FOCUS_{i,t} = \beta_0 + \beta_1 \cdot Affected\ by\ GDPR_{i,t} + \gamma \cdot X_{i,t} + \delta_i + \lambda_t + u_{i,t} \quad (1)$$

The unit of observation is at the firm i and year t level. $INNOVATION\ TYPE\ FOCUS_{i,t}$ is one of the two proxies for the key dependent variable. One proxy is the Balsmeier et al. (2017) patent-based measure *Exploratory Patent Portfolio* for the innovation output test. The second is the *Explore-to-Exploit Salary Ratio* workforce-based measure for the innovation input test. The coefficient on the interaction term, *Affected by GDPR* (i.e., β_1), captures the impact of the GDPR on the affected, or treated, firms. I predict β_1 to be positive for both $INNOVATION\ TYPE\ FOCUS_{i,t}$ measures, because I expect lower data availability to induce firms to allocate more of their finite innovation resources to exploratory relative to exploitative innovation. $X_{i,t}$ represents the vector of time-varying firm controls, where the set of controls for the innovation output and input tests differ slightly, as mentioned in Section 3.2.4. δ represents firm fixed effects, and λ represents year fixed effects. I cluster standard errors at the firm level.

4 Empirical Results

4.1 Descriptive Statistics

I present descriptive statistics for the innovation output test sample in Table 1 Panel A. I present descriptive statistics for the innovation input test sample in Table 6 Panel A. All unbounded continuous variables are winsorized at the first and 99th percentile.

Descriptive statistics in Table 1 Panel A show that the innovation output dependent variable of *Exploratory Patent Portfolio* has a skew toward firms having more exploitative than exploratory patent activity in my sample. The mean of *Exploratory Patent Portfolio* is 0.316. Table 1 Panel A also shows that firms file more exploitative patents than exploratory patents, consistent with the notion that firms primarily exploit by default.

Descriptive statistics in Table 6 Panel A show that the innovation input dependent variable of *Explore-to-Exploit Salary Ratio* reflects firms' larger investment in exploitative than exploratory innovation jobs in my sample. The average ratio of total salaries paid to exploratory, over exploitative, innovation role employees is 0.418. Table 6 Panel A shows that, consistent again with the notion that firms primarily exploit by default, firms invest more in exploitative versus exploratory innovation role employees.

Table 1 Panel B and Table 6 Panel B present the 2-digit SIC code industry distribution across firm-years for the innovation output and input sample, respectively. In the innovation output test sample, the most heavily represented industry is manufacturing, with 2,903 firm-years. The second most heavily represented industry is services, with 794 firm-years. In the innovation input test sample, the most heavily represented industry is also manufacturing, with 1,275 firm-years. The second most heavily represented industry is finance, insurance, and real estate, with 643 firm-years, and services comes in third, with 442 firm-years.

Table 1 Panel C and Table 6 Panel C present the year distribution for my innovation output and input samples. The sample is balanced and well-distributed across each year, with around 200-400 firm-year observations per year. Earlier years have lower numbers of observations due to fewer firms having adopted analytics technologies in those years.

Table 1 Panel D presents the distribution of patent subclasses across patents of firms in the innovation output sample. The highest percentage of patents filed in my sample are in the subclasses of electricity (subclass H) and physics (subclass G).

4.2 Stock Market Reaction to GDPR

I validate GDPR as a plausibly exogenous and significant negative shock to firms' customer data availability by examining the presence and direction of stock market reactions for affected firms. I study the market returns upon the event of a significant announcement leading up to the GDPR's passage, as well as upon the first day the GDPR went into effect. *Market Reaction to GDPR Approach Plan* captures the market reaction to the European Council's first agreeing to a general approach for the GDPR on June 15, 2015, while *Market Reaction to GDPR Implementation* captures the market reaction to the GDPR's implementation on May 25, 2018. These two variables are defined as the stock price on the event date minus the stock price the day before, divided by the stock price the day before.

Table 2 presents results from conducting these market return tests. The table shows negative and significant coefficients on *EU Rev Exposure* in Columns (1) and (2), respectively (coef = -0.003, t-stat = -1.87, for Column (1); coef = -0.006, t-stat = -3.02, for Column (2)). The interpretation of these results is that firms in 2015 had already started to anticipate the GDPR's passage in 2016, validating the choice of 2015 as the basis for determining treatment. In addition, affected firms' strong negative reaction to the GDPR's implementation validates the GDPR as a plausibly exogenous and significant negative shock to availability of customer data, and thus the data's value, to affected firms.

4.3 Evidence from innovation output tests

My hypothesis predicts that lower customer data availability induces firms to allocate more of their finite innovation resources toward exploration relative to exploitation. The results of the innovation output test align with this hypothesis.

I present results of estimating Eq. (1) for *Exploratory Patent Portfolio* as the dependent variable in Table 3. Column (1) estimates the model without control variables. Column (4) estimates the model with controls. The positive and significant coefficients on *Affected by GDPR* indicate that the GDPR induced firms to allocate more of their innovation resources toward exploratory patenting relative to exploitative patenting. The coefficient in Column (4) indicates that a significant negative shock to a firm's customer data availability is associated with 29.9% greater exploratory relative to exploitative activity, compared to unaffected firms (coef = 0.299, t-stat = 2.23). This relative allocation shift toward exploratory patenting activity is economically significant, representing 13.5% of the sample standard deviation of *Exploratory Patent Portfolio*.[6]

The results also examine exploratory and exploitative patenting separately. Firms' shift toward exploration is driven by a greater decrease in exploitative than in exploratory patenting. Columns (2) and (3) provide results from estimating Eq. (1) without controls for *Exploratory Patenting* and *Exploitative Patenting* as dependent variables, respectively. See Appendix B for variable definitions and constructions. Columns (5) and (6) estimate the models in Columns (2) and (3), respectively, with controls. Columns (2) and (5) show that *Exploratory Patenting* decreases for firms affected by the GDPR (coef = -0.006 for Column (2); coef = -0.012 for Column (5)), while Columns (3) and (6) show that *Exploitative Patent-*

[6]The innovation outcome results are not explained by a potential alternative story driven by a negative impact to cash flow. I estimate Eq. (1) with operating cash flow scaled by assets as the dependent variable, and I find no significant change upon the GDPR shock (t-stat = -1.05).

ing also decreases, but more, for affected firms (coef = -0.168 for Column (3); coef = -0.170 for Column (6)). Consistent with the argument behind my hypothesis, exploitation is more sensitive than exploration is to customer data.[7]

A key identifying assumption for difference-in-differences specifications is pre-shock parallel trends. It requires differences between treatment and control firms to be constant in the pre-event period and to likely continue to be, absent treatment. To see whether this assumption holds, Figure 1 plots year-by-year coefficient estimates for my main difference-in-differences regression model with controls, fixed effects, and clustering by year. The omitted year is 2015. Figure 1 visually confirms that the parallel trends assumption holds.

4.4 Robustness tests

In order to alleviate potential concern about the skewness of *Exploratory Patent Portfolio*, I examine high dimensional fixed quantile regressions with controls and firm and year fixed effects at the quantiles of 10, 25, 50, 75, and 90 in Table 4 Panel A. At almost every quantile tested (i.e., 25, 50, 75, 90), coefficients on *Affected by GDPR* are positive and statistically significant, consistent with my hypothesis.

In Table 4 Panel B, I also allay potential concern about observations in GDPR transition period years influencing results. I estimate Eq. (1) for *Exploratory Patent Portfolio* as the dependent variable, after dropping observations in the transition years of 2016 to 2018. Columns (1) and (2) present results, without and with controls, respectively. Columns (1) and (2) show strong positive and significant coefficients on *Affected by GDPR* (coef = 0.382, t-stat = 2.53, for Column (2)), in line with my hypothesis.

[7] The net decrease in innovation, as captured by patenting, upon a negative shock to customer data availability, is also consistent with the notion that data benefits innovation in general, though this is not the question I study in this paper. The level of innovation in the form of patenting, as proxied by the natural log of one plus the total number of patents filed, decreases upon lower customer data availability due to the GDPR (t-stat = -2.29).

4.5 Cross-sectional tests

I conduct split subsample cross-sectional tests, which support the argument that customer data drives my results. I report these cross-sectional test results in Table 5. The analyses estimate Eq. (1) for *Exploratory Patent Portfolio* as the dependent variable. All regressions are estimated using the baseline regression with a full set of control variables, firm and year fixed effects, and clustering standard errors by firm.

4.5.1 B2C vs. B2B

I provide evidence that firms in B2C industries react more strongly to changes in the availability of customer data than do firms in B2B industries. Between Columns (1) and (2), the sample is split based on *B2C Dummy*, or whether firms are in a B2C or B2B industry, respectively. I follow Tacheva et al. (2020) in categorizing firms as B2B versus B2C. Columns (1) and (2) show that the positive coefficient on *Affected by GDPR* is on average larger among B2C firms (coef = 0.769) than B2B firms (coef = 0.055). The results are consistent with the notion that B2C firms, which sell primarily to individuals and thus have greater reliance on their personal data, react more strongly to changes in the availability of customer data than do B2B firms, which sell primarily to other businesses. This difference between subsamples is significant at the 5% level (p-value = 0.026, F-stat = 4.973).

4.5.2 Firm exposure to customer demand volatility

I provide evidence that firms operating in industries where customer demand is more volatile and thus where insights from customer data amortize more quickly, react more strongly to changes in the availability of customer data. Between Columns (3) and (4), the sample is split based on whether the value of *Customer Demand Volatility* is in the top three versus bottom three deciles of the sample, respectively, in a given year. *Customer Demand Volatility* is the standard deviation of revenues of firms within an SIC 2-digit industry in a given year. Columns (3) and (4) show that the positive coefficient on *Affected by GDPR* is on average larger among firm-years in the top three deciles of *Customer Demand Volatility* (coef = 0.656), than among firm-years in the bottom three deciles (coef = 0.014). The results are consistent with firms in industries with more volatile customer demand relying more on the timely insights customer data yield. This difference between subsamples is significant at the 5% level (p-value = 0.033, F-stat = 4.580).

4.6 Evidence from innovation input tests

My hypothesis predicts that lower customer data availability induces firms to allocate more of their finite innovation resources toward exploration relative to exploitation. The results of the innovation input test align with this hypothesis.

I present the results of estimating Eq. (1) for *Explore-to-Exploit Salary Ratio* as the dependent variable in Table 6 Panel D. Column (1) and Column (4) presents results, without and with control variables, respectively. The positive and significant coefficients on *Affected by GDPR* in Columns (1) and (4) indicate that the GDPR induced firms to allocate more resources toward exploratory relative to exploitative job roles. The coefficient in Column (4) indicates that a significant negative shock to a firm's customer data availability is associated with a 9.8% greater ratio of total salaries paid to exploratory, over exploitative, innovation role employees, compared to unaffected firms (coef = 0.098, t-stat = 2.20). This relative shift in total salaries paid to exploratory relative to exploitative innovation role employees is economically significant, as it represents 15.0% of the sample standard deviation of *Explore-to-Exploit Salary Ratio*.

The results also examine this investment in exploratory and exploitative innovation role employees separately. Columns (2) and (3) provide results from estimating Eq. (1) without controls for *Exploratory Employee Salary Ratio* and *Exploitative Employee Salary Ratio* as dependent variables, respectively. See Appendix B for variable definitions and constructions. Columns (5) and (6) estimate the models in Columns (2) and (3), respectively, with controls. Columns (2) and (5) show that investment in exploratory innovation jobs increases for firms affected by the GDPR (coef = 0.004 for Column (2); coef = 0.006 for Column (5)), while Columns (3) and (6) show that investment in exploitative innovation jobs decreases for affected firms (coef = -0.012 for Column (3); coef = -0.013 for Column (6)). Consistent with my hypothesis, upon a decrease in customer data availability, firms shift their innovation resources toward exploration and away from exploitation.[8]

4.7 Data expansion using machine learning technique SMOTE

I present results from conducting analyses on a expanded sample for the innovation output test that balances the number of treatment and control observations, in Table 7. Treatment observations are 26.6% of my original sample, where analysis with an imbalanced sample can lead to less precise estimates and lower statistical power. Therefore, I implement the machine learning synthetic minority over-sampling technique (SMOTE) to balance my sample through data expansion, where the algorithm generates synthetic observations for the minority treatment group until the number of observations in the treatment and control groups equal. Descriptive statistics for this balanced expanded sample are in Table 7 Panel A. The distribution for *EU Rev Exposure* shows that, of the 5,992 observations in the new dataset, the number of observations for both the treatment and control groups is 2,996 observations.

I present results from analysis with this new sample in Table 7 Panel B. I estimate the following difference-in-differences-like specification, where I cannot include firm or year fixed effects or cluster standard errors by firm because SMOTE does not allow for maintaining a firm-year panel structure, and where *ITF* stands for *INNOVATION TYPE FOCUS*:

$$ITF_{i,t} = \beta_0 + \beta_1 \cdot \textit{Affected by GDPR}_{i,t} + \delta \cdot \textit{EU Rev Exposure}_i + \lambda \cdot \textit{Post-GDPR}_t + \gamma \cdot X_{i,t} + u_{i,t} \tag{2}$$

[8] The slight net decrease in innovation, as captured by human capital investment, upon a negative shock to customer data availability, is also consistent with the notion that data benefits innovation in general, though this is not the question I study in this paper. The level of innovation in the form of human capital investment, as proxied by total salaries paid to all employees in innovation-related roles scaled by total salaries paid to all employees at a firm, exhibits a negative trend with lower customer data availability, though it does not change significantly upon GDPR (t-stat = -0.71).

Columns (1) and (2) present results for *Exploratory Patent Portfolio* as the dependent variable, without and with controls, respectively. The positive and significant coefficients align with my hypothesis (coef = 0.184, t-stat = 2.03, for Column (2)).

The coefficient on *Affected by GDPR* in Column (2) indicates that a significant negative shock to firms' customer data availability is associated with 18.4% greater exploratory relative to exploitative patenting, compared to unaffected firms. This magnitude is slightly lower than the magnitude in my main test. This relative allocation shift toward exploratory patenting is economically significant, representing 10.4% of the balanced sample's standard deviation of *Exploratory Patent Portfolio*.

5 Conclusion

In this study, I examine how firms' customer data availability influences their managers' resource allocation between exploitative and exploratory innovation. In estimating a difference-in-differences specification, I find that, upon a negative shock to customer personal data availability, firms allocate more of their finite innovation resources to exploration relative to exploitation. My results are stronger for firms in B2C versus B2B industries and firms with more volatile customer demand, supporting the argument that customer data drives my results. The evidence suggests higher customer data availability leads firms to focus more on exploitative innovation, deterring the exploratory innovation crucial to economic growth.

I contribute to the nascent literature on digital data, the growing accounting literature on innovation, and the literature on the implications of firms' internal information environment for managerial decision-making. The workforce-based innovation input measure that I develop and test may also be useful for future research on firms' innovation strategies.

The results of my study may suggest the counterintuitive idea that, contrary to common positive sentiment, more data is not always better. However, further research is needed to examine the implications of customer data for welfare outcomes.

References

Abrahamian, A. 2018. "Data Subjects of the World, Unite!" *The New York Times*.

Armstrong, C. S., Kepler, J., Kim, C., and Tsui, D. 2024. "Creditor Control Rights and Executive Bonus Plans." Working Paper. Available at *SSRN*: 3975781.

Armstrong, C. S., Konchitchki, Y., and Zhang, B. 2023. "Digital Traffic, Financial Performance, and Stock Valuation." Working Paper. Available at *SSRN*: 4416683.

Baldenius, T. and Azinovic-Yang, L. 2023. "Innovation in Firms: Experimentation and Strategic Communication" Working Paper. Available at *SSRN*: 4300767.

Balsmeier, B., Fleming, L., and Manso, G. 2017. "Independent boards and innovation." *Journal of Financial Economics*, 123(3): 536-557.

Balsmeier, G., Fleming, L., and Manso, G. 2021. "Heterogeneous Innovation over the Business Cycle" Working Paper.

Binz, O., Ferracuti, E., and Joos, P. 2023. "Investment, inflation, and the role of internal information systems as a transmission channel." *Journal of Accounting and Economics*, 67(2-3): 101632.

Bubeck, S. and Cesa-Bianchi, N. 2012. "Regret Analysis of Stochastic and Nonstochastic Multi-armed Bandit Problems." *Foundations and Trends in Machine Learning*, 5(1): 1-22.

Chen, H., Liang, P., and Petrov, E. 2024. "Innovation and Financial Disclosure." *Journal of Accounting Research*, 62(3): 935-979.

Chen, W. and Srinivasan, S. 2024. "Going Digital: Implications for Firm Value and Performance." *Review of Accounting Studies*, 29(2): 1619–1665.

"Cooperative Patent Classification (CPC)." European Patent Office.

Deming, David. 2021. "Balancing Privacy With Data Sharing for the Public Good." *The New York Times*.

Demirer, M., Jiménez Hernández, D. J., Li, D., Peng, S. 2024. "Data, Privacy Laws and Firm Production: Evidence from the GDPR." *NBER Working Paper* 32146.

Dichev, I. and Qian, J. 2022. "The benefits of transaction-level data: The case of NielsenIQ scanner data." *Journal of Accounting and Economics*, 74(1): 2021–2061.

Dixit, A. and Pindyck, R. 1994. *Investment under Uncertainty.* Princeton University Press.

Dubé, J., Bergemann, D., Demirer, M., Goldfarb, A., Johnson, G., Lambrecht, A., Lin, T., Tuchman, A., Tucker, C., Lynch, J. 2024. "The Intended and Unintended Consequences of Privacy Regulation for Consumer Marketing: A Marketing Science Institute Report." Available at *SSRN*: 4847653.

Dutta, S. and Fan, Q. 2012. "Incentives for innovation and centralized versus delegated capital budgeting." *Journal of Accounting and Economics*, 53(3): 592–611.

"ECLA-EUROPEAN PATENT CLASSIFICATION." *IamIP*.

Elsaify, M. and Hasan, S. 2021. "Data exchanges among firms." *Digital Business*, 1(2): 100010.

Ewens, M., Nanda, R., and Rhodes-Kropf, M. 2018. "Cost of experimentation and the evolution of venture capital." *Journal of Financial Economics*, 128(3): 422-442.

Farboodi, M. and Veldkamp, L. 2022. "A Model of the Data Economy." *NBER Technical Report Working Paper* 28427.

Ferracuti, E. 2022. "Information uncertainty and organizational design." *Journal of Accounting and Economics*, 74(1): 101493.

Ferracuti, E., Koo, M., Lee, M., and Stubben, S. 2024. "Acquisition of Customer Information and Corporate Decision Making." Working Paper. Available at *SSRN*: 4682350.

Galasso, A., and Schankerman, M. 2015. "Patents and cumulative innovation: causal evidence from the courts." *The Quarterly Journal of Economics*, 130(1): 317–370.

Gallemore, J. and Labro, E. 2015. "The importance of the internal information environment for tax avoidance." *Journal of Accounting and Economics*, 60(1): 149-167.

Glaeser, C., Glaeser, S., and Labro, E. 2023. "Proximity and the Management of Innovation." *Management Science*, 69(5): 3080-3099.

Glaeser, S. and Landsman, W. 2021. "Deterrent disclosure." *The Accounting Review*, 96(5), 291-315.

Glaeser, S. and Lang, M. 2024. "A Review of the Accounting Literature on Innovation." *Journal of Accounting and Economics*: 101720.

Glaeser, S., Michels, J., and Verrecchia, R. 2020. "Discretionary disclosure and manager horizon: Evidence from patenting." *Review of Accounting Studies*, 25(2): 597–635.

Goldberg, S., Johnson, G., Shriver, S. 2024. "Regulating Privacy Online: An Economic Evaluation of the GDPR." *American Economic Journal: Economic Policy*, 16(1): 325-358.

Grabner, I., Posh, A., Wabnegg, M. 2018. "Materializing innovation capability: A management control perspective." *Journal of Management Accounting Research*, 30(2):163–185.

Hagel, J., III and Rayport, J. 1997. "The coming battle for customer information." *Harvard Business Review*, 75(1): 53-65.

Hall, B., Jaffe, A., Trajtenberg, M. 2001. "The NBER patent citation data file: Lessons, insights and methodological tools." *NBER Working Paper* No. 8498.

Heitzman, S. and Huang, M. 2019. "Internal Information Quality and the Sensitivity of Investment to Market Prices and Accounting Profits." *Contemporary Accounting Research*, 36(3):1699-1723.

"The History of the General Data Protection Regulation." European Data Protection Supervisor.

Holmström, B. 1979. "Moral Hazard and Observability." *The Bell Journal of Economics*, 10(1): pp. 74-91.

Jaffe, A. 1989. "Characterizing the 'technological position' of firms, with application to quantifying technological opportunity and research spillovers." *Research Policy*, 18(2), 87–97.

Jones, C. 2023. "Recipes and economic growth: a combinatorial march down an exponential tail." *Journal of Political Economy*, 131(8).

Jones, C. and Tonetti, C. 2020. "Nonrivalry and the Economics of Data." *American Economic Review*, 110(9): 2819-58.

Keeley, B., Walters, H., Pikkel, R., and Quinn, B. 2013. *Ten Types of Innovation: The Discipline of Building Breakthroughs*. Wiley; 1st edition.

Kepler, J., McClure, C., and Stewart, C. 2024. "Competition Enforcement and Accounting for Intangible Capital." Working Paper. Available at *SSRN*: 4861968.

Kim, J., and Valentine, K. 2021. "The innovation consequences of mandatory patent disclosures." *Journal of Accounting and Economics*, 71(2-3): 101381.

Kogan, L., Papanikolaou, D., Seru, A., and Stoffman, N. 2017. "Technological innovation, resource allocation, and growth." *The Quarterly Journal of Economics*, 132(2):665–712.

Koh, P. and Reeb, D. 2015. "Missing R&D." *Journal of Accounting and Economics*, 60(1): 73-94.

Laux, V., and Ray, K. 2020. "Effects of accounting conservatism on investment efficiency and innovation." *Journal of Accounting and Economics*, 70(1): 101319.

Laux, V. and P. Stocken. 2018. "Accounting standards, regulatory enforcement, and innovation." *Journal of Accounting and Economics*, 65(2-3): 221–236.

Lerner, J. and Seru, A. 2022. "The Use and Misuse of Patent Data: Issues for Finance and Beyond." *The Review of Financial Studies*, 35(6): 2667–2704.

Manso, G. 2011. "Motivating Innovation." *Journal of Finance*, 66(5): 1823-60.

March, J. 1991. "Exploration and Exploitation in Organizational Learning." *Organization Science*, 2(1): 71-87.

Menguc B. and Auh S. 2008. "The asymmetric moderating role of market orientation on the ambidexterity-firm performance relationship for prospectors and defenders." *Industrial Marketing Management*, 37(4): 455–470.

Mezzanotti, F. and Simcoe, T. 2023. “Research and/or Development? Financial Frictions and Innovation Investment.” *NBER Working Paper* No. 31521.

”The OECD and Digital Services Taxes.” 2024. *Bloomberg Tax*.

Ovide, Shira. 2022. “The Messy Progress on Data Privacy.” *The New York Times*.

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ L 119, 4.5.2016, p.1-88 (2016, May 4).

Romer, P. 1990. “Endogenous Technological Change.” *Journal of Political Economy*, 98(5): S71–S102.

Romer, P. 1993. “Two strategies for economic development: using ideas and producing ideas.” *Proceedings of the World Bank Annual Conference on Development Economics*, pp. 63-115.

Rubera, G., and Kirca, A. 2012. “Firm innovativeness and its performance outcomes: A meta-analytic review and theoretical integration.” *Journal of Marketing*, 76(3): 130–147.

Samuels, D. 2021. “Government Procurement and Changes in Firm Transparency.” *The Accounting Review*, 96(1): 401-430.

Schumpeter, J. 1942. *Capitalism, Socialism and Democracy.* Harper & Row, New York, Vol. 36, 132-145.

Stanford Graduate School of Business Library. 2024. Revelio Labs Workforce Data (Version 4.0) [Data set]. Redivis (RRID:SCR_023111).

Tacheva, Z., Simpson, N., and Ivanov, A. 2020. “Examining the Role of Top Management in Corporate Sustainability: Does Supply Chain Position Matter?” *Sustainability*, 12(18): 7518.

Thomas, J. 2024. “Using Data Analytics to Improve Customer Experience.” *Mouseflow*.

“Valuing Data.” 2024. *International Valuation Standards Council Perspectives Paper*.

Veldkamp, Laura. 2023a. “6 approaches to valuing data.” *IESE Business School Insight*, no. 163.

Veldkamp, Laura. 2023b. “Valuing Data as an Asset.” *Review of Finance*, 27(5): 1545–1562.

Wolford, Ben. “What is GDPR, the EU’s new data protection law?” *GDPR.EU*.

“The world’s most valuable resource is no longer oil, but data.” 2017. *The Economist*.

Zhu, C. 2019. “Big Data as a Governance Mechanism.” *The Review of Financial Studies*, 32(5): 2021–2061.

Zumbrun, J. 2016. “The rise of knowledge workers is accelerating despite the threat of automation.” *Wall Street Journal*.

Table 1

Panel A: Descriptive Statistics for Innovation Output Sample

	N	*Mean*	*Std. Dev.*	*25%*	*Median*	*75%*
Main Test						
Exploratory Patent Portfolio	4,662	0.316	2.210	0	0.0004	0.035
Exploratory Patenting	4,662	0.126	0.323	0	0	0
Exploitative Patenting	4,662	1.399	1.757	0	0.693	2.398
Affected by GDPR	4,662	0.143	0.350	0	0	0
EU Rev Exposure	4,662	0.266	0.442	0	0	1
Post-GDPR	4,662	0.522	0.500	0	1	1
Assets	4,662	7.618	2.281	5.973	7.590	9.273
Leverage	4,636	0.239	0.201	0.054	0.222	0.360
Book-to-Market	4,367	0.459	0.386	0.200	0.368	0.621
Employees	4,585	1.278	2.183	-0.212	1.387	2.890
R&D	4,662	2.751	2.516	0	2.958	4.533
Sales	4,549	7.135	2.313	5.784	7.291	8.729
PPE	4,484	6.228	2.578	4.445	6.270	8.013
Cross-Sectional and Heterogeneity Test						
B2C Dummy	4,582	0.294	0.456	0	0	1
Customer Demand Volatility	4,662	1,781	5,811	184.2	389.9	1,058
Validation Test						
Market Reaction to GDPR Approach Plan	402	-0.006	0.018	-0.014	-0.007	0.0007
Market Reaction to GDPR Implementation	358	0.002	0.018	-0.007	0.001	0.009

Notes: This table presents descriptive statistics for the variables examined in estimated regressions for the innovation output sample. Detailed definitions of all variables are provided in Appendix B.

Table 1

Panel B: Industry Distribution for Innovation Output Sample

2-Digit SIC Code	*Industry Title*	*Frequency*	*Percent*
10-14	Mining	148	3.17
15-17	Construction	43	0.92
20-39	Manufacturing	2,903	62.27
40-49	Transportation, Communications, Electric, Gas, And Sanitary Services	309	6.63
50-51	Wholesale Trade	71	1.52
52-59	Retail Trade	122	2.62
60-67	Finance, Insurance, And Real Estate	252	5.41
70-89	Services	794	17.03
99	Non-classifiable Establishments	20	0.43
Total		**4,662**	**100**

Notes: This table presents the 2-digit SIC code industry distribution of firm-year observations in the innovation output sample.

Table 1

Panel C: Year Distribution for Innovation Output Sample

Year	*Frequency*	*Percent*
2010	184	3.95
2011	321	6.89
2012	404	8.67
2013	432	9.27
2014	443	9.5
2015	446	9.57
2016	436	9.35
2017	425	9.12
2018	414	8.88
2019	392	8.41
2020	388	8.32
2021	377	8.09
Total	**4,662**	**100**

Notes: This table presents the year distribution of firm-year observations in the innovation output sample.

Table 1

Panel D: CPC Patent Subclass Distribution

CPC 1-digit Subclass	*Frequency*	*Percent*
A: Human Necessities	68,842	7.7
B: Performing Operations; Transporting	76,233	8.53
C: Chemistry, Metallourgy	40,065	4.48
D: Textiles; Paper	3,206	0.36
E: Fixed Constructions	13,463	1.51
F: Mechanical Engineering	45,074	5.04
G: Physics	317,392	35.5
H: Electricity	329,773	36.89
Total	**894,048**	**100**

Notes: This table presents descriptive statistics for the CPC 1-digit patent subclasses of patents for firms in the innovation output sample.

Table 2

Market Reaction to GDPR

	(1) *Market Reaction to GDPR Approach Plan*	(2) *Market Reaction to GDPR Implementation*
EU Rev Exposure	-0.003*	-0.006***
	(-1.87)	(-3.02)
Constant	-0.005***	0.003***
	(-4.39)	(2.96)
Observations	402	358
R-squared	0.005	0.020

Notes: This table presents results from studying market reactions to GDPR events. Column (1) features the results of estimating the regression of *Market Reaction to GDPR Approach Plan* on *EU Rev Exposure*. Column (1) executes a test around the June 15, 2015 event that the European Council reached a general approach on the GDPR. Column (2) features the results of estimating the regression of *Market Reaction to GDPR Implementation* on *EU Rev Exposure*. Column (2) executes a test around the first day GDPR is implemented or enforced on May 25, 2018. The unit of observation is at the firm level. The values in the parentheses represent t-statistics. ***, **, * indicates statistical significance at the 1%, 5%, and 10% levels, respectively (two-tailed). Detailed definitions of all variables are provided in Appendix B.

Table 3

The Effect of GDPR on Innovation Strategy Focus: Results for Output Tests

	(1) *Exploratory Patent Portfolio*	(2) *Exploratory Patenting*	(3) *Exploitative Patenting*	(4) *Exploratory Patent Portfolio*	(5) *Exploratory Patenting*	(6) *Exploitative Patenting*
Affected by GDPR	0.322**	-0.006	-0.168**	0.299**	-0.012	-0.170**
	(2.57)	(-0.23)	(-2.41)	(2.23)	(-0.44)	(-2.40)
Assets				0.041	0.023	0.141***
				(0.37)	(1.47)	(3.05)
Leverage				0.141	0.059	-0.109
				(0.86)	(1.26)	(-0.81)
Book-to-Market				-0.238	-0.004	-0.075
				(-1.49)	(-0.19)	(-1.41)
Employees				0.161	0.022	0.148**
				(1.12)	(1.06)	(2.16)
R&D				0.100	0.016	0.047
				(1.24)	(1.33)	(1.17)
Sales				0.094*	0.008	-0.093***
				(1.80)	(0.83)	(-2.64)
PPE				-0.378***	-0.041**	-0.043
				(-2.77)	(-2.32)	(-0.73)
Observations	4,662	4,662	4,662	4,100	4,100	4,100
R-squared	0.207	0.320	0.907	0.202	0.329	0.913
Firm FEs	YES	YES	YES	YES	YES	YES
Year FEs	YES	YES	YES	YES	YES	YES

Notes: This table presents difference-in-differences regression results from estimating Eq. (1) with either *Exploratory Patent Portfolio*, *Exploratory Patenting*, or *Exploitative Patenting* as the dependent variables. Column (1) estimates the baseline regression with *Exploratory Patent Portfolio* and firm and year fixed effects. Column (2) is the baseline regression with firm and year fixed effects except with *Exploratory Patenting*, and Column (3) is the baseline regression with firm and year fixed effects except with *Exploitative Patenting*. Column (4-6) adds the vector of appropriate control variables to the regressions run in Columns (1-3). The unit of observation is at the firm-year level. The standard errors are clustered by firm. The values in the parentheses represent t-statistics. ***, **, * indicates statistical significance at the 1%, 5%, and 10% levels, respectively (two-tailed). Detailed definitions of all variables are provided in Appendix B.

Table 4

Panel A: The Effect of GDPR on Innovation Strategy Focus: Quantile Tests

Quantile:	(1) 10th	(2) 25th	(3) 50th	(4) 75th	(5) 90th
Affected by GDPR	-0.265**	0.115**	0.196***	0.509**	1.144**
	(-2.06)	(2.15)	(2.69)	(2.29)	(1.99)
Assets	0.081	0.054	0.048	0.026	-0.019
	(0.53)	(1.30)	(0.70)	(0.13)	(-0.04)
Leverage	-0.134	0.051	0.091	0.243	0.553
	(-0.56)	(0.53)	(0.77)	(0.90)	(0.86)
Book-To-Market	0.165	-0.107**	-0.165**	-0.388	-0.842
	(0.54)	(-2.18)	(-2.11)	(-1.23)	(-1.00)
Employees	-0.125	0.068	0.109	0.268	0.591
	(-0.82)	(1.06)	(1.22)	(1.11)	(1.03)
R&D	-0.155	0.017	0.054	0.195	0.483
	(-1.48)	(0.49)	(1.11)	(1.40)	(1.40)
Sales	-0.054	0.046*	0.068**	0.150*	0.316
	(-0.83)	(1.86)	(2.07)	(1.74)	(1.50)
PPE	0.133	-0.212***	-0.285***	-0.569**	-1.144*
	(0.73)	(-3.83)	(-3.81)	(-2.36)	(-1.91)
Observations	4,113	4,113	4,113	4,113	4,113
Firm FEs	YES	YES	YES	YES	YES
Year FEs	YES	YES	YES	YES	YES

Notes: This table presents quantile regression results from estimating Eq. (1) with *Exploratory Patent Portfolio* as the dependent variable. Columns (1)-(5) present results from estimating the baseline regressions with the vector of appropriate control variables along with firm and year fixed effects for the quantiles of 10, 25, 50, 75, and 90, respectively. The unit of observation is at the firm-year level. The standard errors are clustered by firm. The values in the parentheses represent t-statistics. ***, **, * indicates statistical significance at the 1%, 5%, and 10% levels, respectively (two-tailed). Detailed definitions of all variables are provided in Appendix B.

Table 4

Panel B: Results with Dropped Transition Period

	(1) *Exploratory Patent Portfolio*	(2) *Exploratory Patent Portfolio*
Affected by GDPR	0.378***	0.382**
	(2.70)	(2.53)
Assets		0.046
		(0.28)
Leverage		-0.009
		(-0.04)
Book-to-Market		-0.397
		(-1.43)
Employees		0.067
		(0.40)
R&D		0.087
		(0.97)
Sales		0.141**
		(2.35)
PPE		-0.268*
		(-1.81)
Observations	3,379	2,959
R-squared	0.248	0.228
Firm FEs	YES	YES
Year FEs	YES	YES

Notes: This table presents difference-in-differences regression results from estimating Eq. (1) with *Exploratory Patent Portfolio* as the dependent variable, but with observations for the years 2016 to 2018 dropped. Column (1) includes the baseline regression with firm and year fixed effects. Column (2) adds the vector of appropriate control variables. The unit of observation is at the firm-year level. The standard errors are clustered by firm. The values in the parentheses represent t-statistics. ***, **, * indicates statistical significance at the 1%, 5%, and 10% levels, respectively (two-tailed). Detailed definitions of all variables are provided in Appendix B.

Table 5

Cross-sectional analyses

	B2C vs. B2B		Customer Demand Volatility	
	(1) B2C	(2) B2B	(3) High	(4) Low
Affected by GDPR	0.769** (2.56)	0.055 (0.48)	0.656** (2.25)	0.014 (0.18)
p-value	*0.026*		*0.033*	
F-stat	*4.973*		*4.580*	
Observations	1,188	2,912	1,204	1,528
R-squared	0.201	0.216	0.190	0.294
Controls	YES	YES	YES	YES
Firm FEs	YES	YES	YES	YES
Year FEs	YES	YES	YES	YES

Notes: This table reports the results of cross-sectional analyses of innovation output strategy focus around the GDPR. The analyses present difference-in-differences regression results from estimating Eq. (1) with *Exploratory Patent Portfolio* as the dependent variable. Between Columns (1) and (2), the sample is split based on *B2C Dummy*, or whether firms are in a B2C or B2B industry. Columns (3) and (4) mirror the presentation in Columns (1) and (2), but the sample is split based on whether the customer demand volatility level *Customer Demand Volatility* is in the top three deciles versus the bottom three deciles in the sample in a given year. All columns estimate the baseline regression with the full set of appropriate control variables, as well as firm and year fixed effects. The unit of observation is at the firm-year level. The standard errors are clustered by firm. ***, **, * indicates statistical significance at the 1%, 5%, and 10% levels, respectively (two-tailed). The values in the parentheses represent t-statistics. Detailed definitions of all variables are provided in Appendix B.

Table 6

Panel A: Descriptive Statistics for Innovation Input Sample

	N	*Mean*	*Std. Dev.*	*25%*	*Median*	*75%*
Explore-to-Exploit Salary Ratio	2,838	0.418	0.655	0.075	0.197	0.449
Exploratory Employee Salary Ratio	2,838	0.114	0.137	0.016	0.062	0.157
Exploitative Employee Salary Ratio	2,838	0.392	0.270	0.133	0.368	0.637
Affected by GDPR	2,838	0.125	0.331	0	0	0
EU Rev Exposure	2,838	0.186	0.389	0	0	0
Post-GDPR	2,838	0.704	0.457	0	1	1
Assets	2,838	6.676	1.726	5.637	6.725	7.848
Leverage	2,819	0.241	0.233	0.034	0.187	0.389
Book-to-Market	2,652	0.517	0.450	0.227	0.431	0.704

Notes: This table presents descriptive statistics for the variables examined in estimated regressions for the innovation input sample. Detailed definitions of all variables are provided in Appendix B.

Table 6

Panel B: Industry Distribution for Innovation Input Sample

2-Digit SIC Code	*Industry Title*	*Frequency*	*Percent*
10-14	Mining	117	4.12
15-17	Construction	43	1.52
20-39	Manufacturing	1,275	44.93
40-49	Transportation, Communications, Electric, Gas, And Sanitary Services	213	7.51
50-51	Wholesale Trade	70	2.47
52-59	Retail Trade	35	1.23
60-67	Finance, Insurance, And Real Estate	643	22.66
70-89	Services	442	15.57
Total		**2,838**	**100**

Notes: This table presents the 2-digit SIC code industry distribution of firm-year observations in the innovation input sample.

Table 6

Panel C: Year Distribution for Innovation Input Sample

Year	*Frequency*	*Percent*
2013	258	9.09
2014	283	9.97
2015	299	10.54
2016	316	11.10
2017	329	11.56
2018	343	12.12
2019	337	11.91
2020	345	12.16
2021	328	11.56
Total	**2,838**	**100**

Notes: This table presents the year distribution of firm-year observations in the innovation input sample.

Table 6

Panel D: The Effect of GDPR on Innovation Strategy Focus: Results for Input Tests

	(1)	(2)	(3)	(4)	(5)	(6)
	Explore-to-Exploit Salary Ratio	*Exploratory Employee Salary Ratio*	*Exploitative Employee Salary Ratio*	*Explore-to-Exploit Salary Ratio*	*Exploratory Employee Salary Ratio*	*Exploitative Employee Salary Ratio*
Affected by GDPR	0.086*	0.004	-0.012	0.098**	0.006	-0.013
	(1.75)	(0.65)	(-1.52)	(2.20)	(1.26)	(-1.59)
Assets				0.020	0.001	-0.005
				(1.26)	(0.31)	(-0.91)
Leverage				-0.089*	-0.010	0.003
				(-1.95)	(-1.55)	(0.22)
Book-to-Market				0.048**	0.006	-0.003
				(2.33)	(1.63)	(-0.65)
Observations	2,838	2,838	2,838	2,629	2,629	2,629
R-Squared	0.928	0.964	0.976	0.932	0.965	0.975
Firm FEs	YES	YES	YES	YES	YES	YES
Year FEs	YES	YES	YES	YES	YES	YES

Notes: This table presents difference-in-differences regression results from estimating Eq. (1) with either the innovation input measure *Explore-to-Exploit Salary Ratio*, *Exploratory Employee Salary Ratio*, or *Exploitative Employee Salary Ratio* as the dependent variables. Column (1) estimates the baseline regression with *Explore-to-Exploit Salary Ratio* and firm and year fixed effects. Column (2) is the baseline regression with firm and year fixed effects except with *Exploratory Employee Salary Ratio*, and Column (3) is the baseline regression with firm and year fixed effects except with *Exploitative Employee Salary Ratio.* Column (4-6) adds the vector of appropriate control variables to the regressions run in Columns (1-3). The unit of observation is at the firm-year level. The standard errors are clustered by firm. The values in the parentheses represent t-statistics. ***, **, * indicates statistical significance at the 1%, 5%, and 10% levels, respectively (two-tailed). Detailed definitions of all variables are provided in Appendix B.

Table 7

Panel A: Descriptive Statistics for SMOTE Sample

	N	*Mean*	*Std. Dev.*	*25%*	*Median*	*75%*
Exploratory Patent Portfolio	5,992	0.250	1.765	0	0.002	0.042
Affected by GDPR	5,992	0.271	0.445	0	0	1
EU Rev Exposure	5,992	0.500	0.500	0	0.500	1
Post-GDPR	5,992	0.532	0.499	0	1	1
Assets	5,992	7.729	2.206	6.154	7.743	9.301
Leverage	5,992	0.243	0.186	0.085	0.236	0.356
Book-to-Market	5,992	0.451	0.356	0.217	0.373	0.594
Employees	5,992	1.436	2.099	0.021	1.591	2.991
R&D	5,992	3.010	2.551	0	3.235	4.708
Sales	5,992	7.257	2.257	5.919	7.477	8.837
PPE	5,992	6.394	2.526	4.755	6.484	8.134

Notes: This table presents descriptive statistics for the variables examined in estimated regressions for the innovation output test dataset balanced through data expansion using the machine learning technique of synthetic minority over-sampling technique (SMOTE).

Table 7

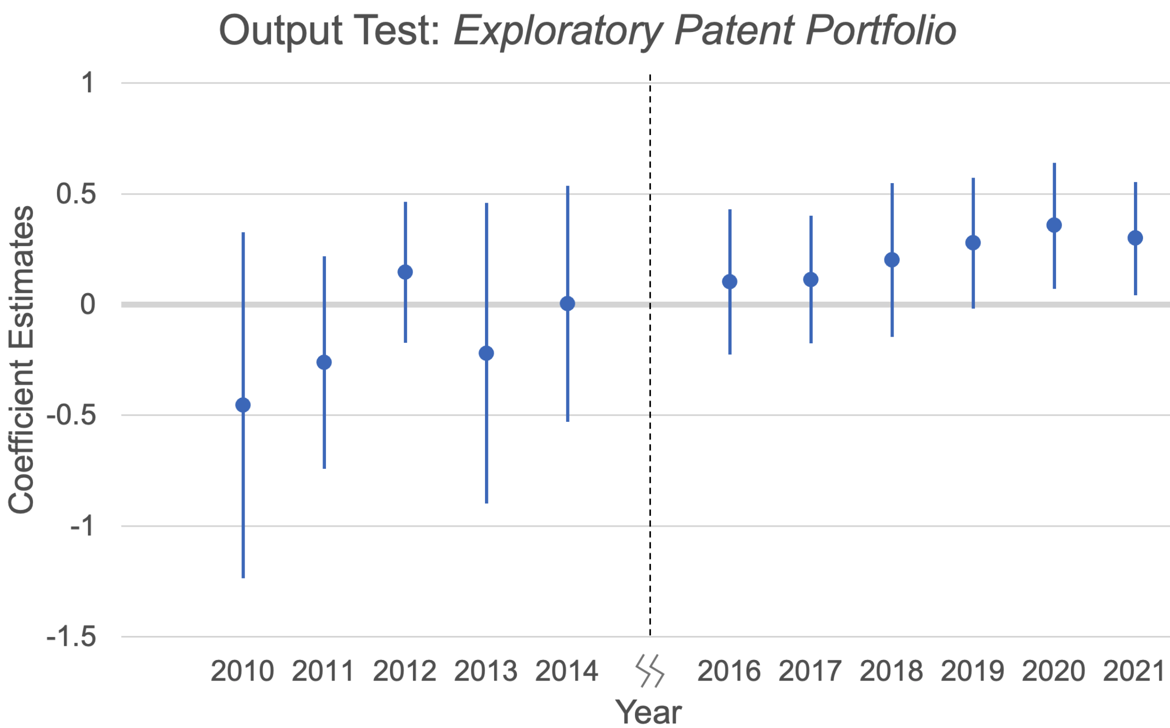
Panel B: Results with SMOTE-generated dataset

	(1) *Exploratory Patent Portfolio*	(2) *Exploratory Patent Portfolio*
Affected by GDPR	0.180**	0.184**
	(1.98)	(2.03)
EU Rev Exposure	-0.317***	-0.274***
	(-4.78)	(-4.08)
Post-GDPR	-0.317***	-0.292***
	(-4.93)	(-4.52)
Assets		0.058*
		(1.74)
Leverage		-0.306**
		(-2.22)
Book-to-Market		-0.100
		(-1.47)
Employees		-0.072**
		(-2.31)
R&D		-0.078***
		(-7.63)
Sales		0.080**
		(2.12)
PPE		-0.021
		(-0.76)
Observations	5,992	5,992
R-squared	0.009	0.020

Notes: This table presents difference-in-differences regression results from estimating Eq. (2) with *Exploratory Patent Portfolio* as the dependent variable, run on the innovation output dataset balanced through data expansion using the machine learning technique of synthetic minority over-sampling technique (SMOTE). Columns (1) and (2) study the innovation output variable *Exploratory Patent Portfolio* without and with appropriate control variables, respectively. The unit of observation is at the firm-year level. The values in the parentheses represent t-statistics. ***, **, * indicates statistical significance at the 1%, 5%, and 10% levels, respectively (two-tailed). Detailed definitions of all variables are provided in Appendix B.

Figure 1

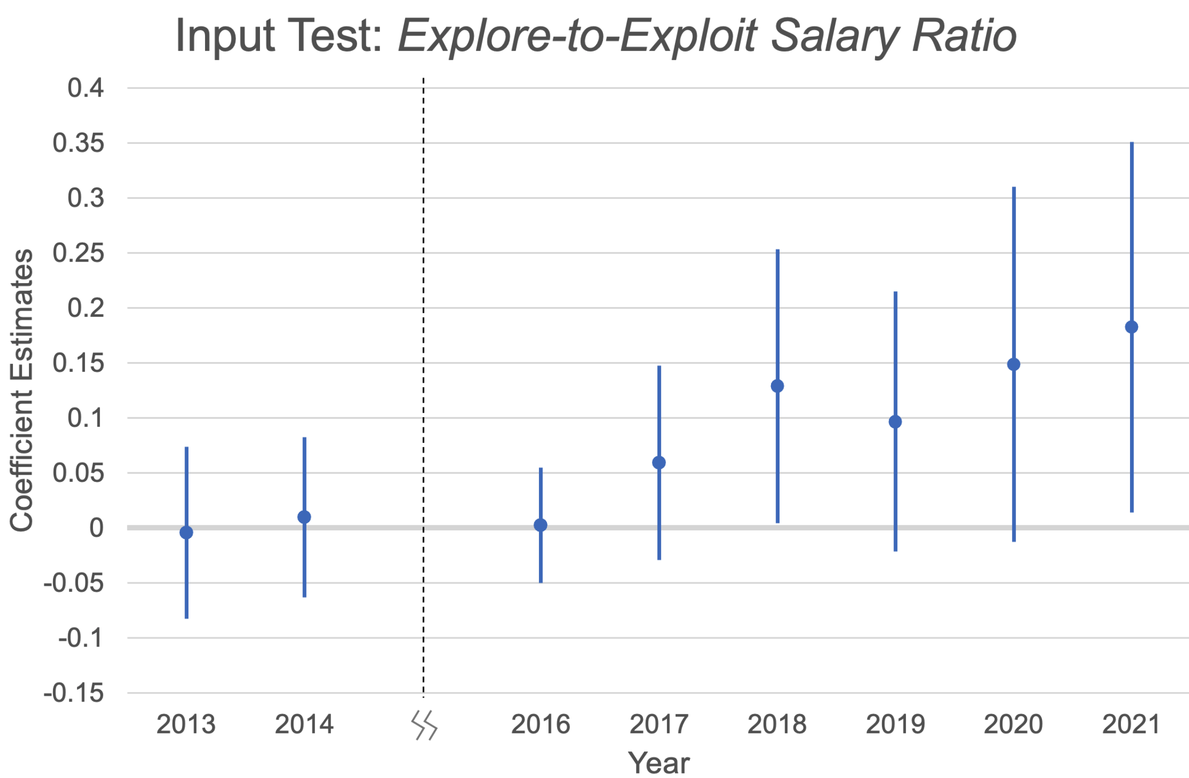
Output Test: *Exploratory Patent Portfolio*



Notes: This figure displays the year-by-year coefficients from estimating the Eq. (1) difference-in-differences regression model with *Exploratory Patent Portfolio* as the dependent variable, with all appropriate controls, firm and year fixed effects, and clustering standard errors by firm. The omitted year is 2015. The definition of *Exploratory Patent Portfolio* is in Appendix B.

Figure 2

Input Test: *Explore-to-Exploit Salary Ratio*



Notes: This figure displays the year-by-year coefficients from estimating the Eq. (1) difference-in-differences regression model with *Explore-to-Exploit Salary Ratio* as the dependent variable, with all appropriate controls, firm and year fixed effects, and clustering standard errors by firm. The omitted year is 2015. The definition of *Explore-to-Exploit Salary Ratio* is in Appendix B.

Appendix A. Sample Selection Criteria

This table presents the sample selection criteria for my main analyses.

Criteria	**Firm-year observations**	**Number of firms**
Firm-year observations for public U.S. firms on CRSP/Compustat from 2010-2021	**65,925**	**9,647**
Firm-year observations for public U.S firms with e-commerce or analytics technologies on websites at least three years before GDPR passage, from 2010-2021	**10,235**	**1,814**
Innovation output test (patent-based):		
After merging with Kogan et al. (2017) extended patent data and keeping observations that adhere to the innovation output regression test design:	4,662	533
Innovation input test (workforce-based):		
After merging with Revelio Labs workforce sample and keeping observations that adhere to the innovation input regression test design:	2,838	383

Appendix B. Variable Definitions

Variable	Definition
Exploratory Patent Portfolio	One minus the Balsmeier et al. (2017) variable that calculates the similarity between the distribution of patents across 3-digit CPC technology classes filed by a given firm in a year and the same firm's prior year distribution of patents across technology classes. The distribution is of the number of patents (in year of filing) per technology class and firm. The variable ranges from 0 and 100, where the 0 end indicates exploitation and the 100 end indicates exploration. Patent data from PatentsView and Kogan et al. (2017).
Exploratory Patenting	The natural log of one plus the number of patents filed in a new technology class by a firm, in a given year. Patent data from PatentsView and Kogan et al. (2017).
Exploitative Patenting	The natural log of one plus the number of patents filed in a technology class previously filed in by a firm, in a given year. Patent data from PatentsView and Kogan et al. (2017).
Explore-to-Exploit Salary Ratio	The ratio of total salaries paid to exploratory innovation role employees to total salaries paid to exploitative innovation role employees in a given firm and year. Workforce data from Revelio Labs.
Exploratory Employee Salary Ratio	The total salaries paid to exploratory innovation role employees divided by the total salaries paid to all employees in a given firm and year. Workforce data from Revelio Labs.
Exploitative Employee Salary Ratio	The total salaries paid to exploitative innovation role employees divided by the total salaries paid to all employees in a given firm and year. Workforce data from Revelio Labs.
EU Rev Exposure	A dummy that equals 1 if a firm earns significant percentage of revenues from the EU in 2015, 0 otherwise. Significance is determined by whether revenues are disclosed in the 10-K footnotes as from the EU. Data for revenue significance by geographical region is collected from Compustat Segments.
Post-GDPR	A dummy that equals 1 if a firm-year observation is in or after the GDPR passage year of 2016, 0 otherwise.
Affected by GDPR	The interaction term of *EU Rev Exposure* with *Post-GDPR.*
Assets	The natural logarithm of total assets (AT) from Compustat.
Leverage	Total liabilities divided by total assets (DLTT+DLC)/AT from Compustat.
Book-to-Market	Book value of equity divided by market capitalization (CEQ+TXDB)/(PRCC_F*CSHO) from CRSP/Compustat.
Employees	The natural logarithm of employees (EMP) in the prior year from Compustat.
R&D	The natural logarithm of research and development expense (XRD) in the prior year from Compustat, with missing values replaced with 0, in accordance with Koh and Reeb (2015).
Sales	The natural logarithm of sales (REVT) in the prior year from Compustat.

PPE	The natural logarithm of total gross property, plant and equipment (PPEGT) in the prior year from Compustat.
B2C Dummy	A dummy that equals 1 for a firm in a B2C 4-digit SIC industry, 0 for a firm in a B2B 4-digit SIC industry, based on classification from Tacheva et al. (2020).
Customer Demand Volatility	The standard deviation of revenues of firms within an SIC 2-digit industry in a given year. Data on revenues (REVT) from Compustat.
Market Reaction to GDPR Approach Plan	The stock price return to GDPR-affected treatment firms upon the event of the European Council first reaching a general approach on GDPR on June 15, 2015. Defined as the price on the event date minus the price the day before the event, divided by the price the day before the event. Data from CRSP.
Market Reaction to GDPR Implementation	The stock price return to GDPR-affected treatment firms upon the implementation or enforcement of GDPR on May 25, 2018. Defined as the price on the event date minus the price the day before the event, divided by the price the day before the event. Data from CRSP.

**All continuous unbounded variables are winsorized at the 1^{st} and 99^{th} percentile.*

Appendix C. Discussions of Data in the Media

Below are quotes that support the common positive sentiment about data:

- *The Economist* article from 2017/05/06 (The Economist, 2017): "The world's most valuable resource is no longer oil, but data."
- *New York Times* article from 2018/05/28 (Abrahamian, 2018): "Data is currency; creating and holding it is power."
- *New York Times* article from 2021/02/19 (Deming, 2021): "[Data has] increasing value as an economic resource..."; "'Data is the new oil.'"
- *New York Times* article from 2022/05/12 (Ovide, 2022): "Tech companies... mostly unknown data middlemen, and even the local supermarket harvest any morsel of data on us that might help their businesses"; data is good for all types of businesses and the economy.
- Laura Veldkamp (Veldkamp, 2023a): "[Data has] central importance to our modern economy"; the most valuable firms today are valuable because of their data.

Appendix D. Anecdotal Validation of Customer Data's Value

I present anecdotal evidence from Fossil Group's 2018 10-K that discusses its use of customer data to generate insights, innovate, and succeed.

"Owned Brands" (page 1):

- "Our **consumer-first mindset** drives every decision we make. By capitalizing on major fashion trends and **leveraging proprietary data and insights**, we are able to deliver **relevant, high-value product and experiences** to consumers across a diverse range of price points, **style preferences** and **geographies**."

"Building Strong Brands" (page 3):

- "Our **product designs are fueled** by a combination of **creativity**, fashion trends and **consumer insights**. Over the past 30 plus years, we have built an incredible in-house design team that works in partnership with our **consumer insights** and trend teams to **ideate, design, test and deliver new product concepts to market**. We also employ more than 200 research and development ("R&D") team members who focus on **innovation and product development** across our watch and smartwatch categories...

- In order to respond to and capitalize on fast-paced changes in the global marketplace, we have created a process that allows us to **design and develop consumer insight driven product** in as little as 30 days. We have found speed coupled with **insight-driven product** to be a true differentiator and **key revenue driver** across our business...

- We are also able to deliver a **high level of personalization** through the **consumer insight and predictive analytics capabilities** we have built over the past few years and through our partnerships with leading **online** third-party retailers."

"Operating Strategy" (page 4):

- "Our goal is to **drive shareholder value** by **increasing earnings** and making a positive impact on our people, planet and communities. While we currently operate in a challenging business environment we are leveraging our business strengths while continuing to lead a significant internal transformation to strengthen our business model. We plan to achieve our **business strategy** by focusing on the following strategic initiatives...

- We are focused on improving our overall profitability through revenue-management strategies to price and position our products optimally and most effectively...

- We are driving innovation across every aspect of our business. We continue to form new partnerships with leading brands, which helps us leverage our vertical structure, size and scale. We are bringing new and innovative functions to smartwatches across both our hybrid and display platforms—as well as expanding our distribution and increasing our addressable market. We are also **driving innovations** in traditional watches and our accessory categories **through our investment** in R&D and **data analytics**...

- We are investing in our **digital infrastructure** and commerce capabilities across our owned and third-party **e-commerce sites**. We plan to continue to expand our **digital capabilities** for **consumer insight, analytics and the use of data** throughout our organization, as well as shifting to a new marketing and commerce platform to drive greater personalization and a better consumer experience."

"Any material disruption of our information systems could disrupt our business and reduce our sales" (page 17):

- "We are **increasingly dependent on information systems to operate our websites, process transactions, manage inventory, monitor sales and purchase, sell and ship goods on a timely basis**...

- In addition, we have **e-commerce** and **other websites in the U.S. and internationally**. In addition to changing consumer preferences and buying trends relating to Internet usage, **we are vulnerable** to certain additional risks and uncertainties associated with the Internet, including changes in required technology interfaces, website downtime and other technical failures, security breaches, and **consumer privacy concerns**. Our **failure to successfully respond to these risks and uncertainties could reduce e-commerce sales, increase costs and damage the reputation of our brands**."

Appendix E. Anecdotal GDPR Validation

I present anecdotal evidence that validates GDPR as a plausibly exogenous and significant negative shock to affected firms' data value.

Exhibit 1. Annual report excerpts about GDPR's effect on the collection and analysis of consumer information

Activision Blizzard 2018 10-K:

- "We **collect and store information about our consumers**, including consumers who play these games. In addition, we collect and store information about our employees. We are **subject to laws from a variety of jurisdictions regarding privacy and the protection of this information, including the E.U.'s General Data Protection Regulation (the "GDPR")**... Failure to comply with any of these laws or regulations may increase our costs, subject us to expensive and distracting government investigations, and result in substantial fines."

Groupon 2018 10-K:

- "**We are subject to a variety of federal, state and international laws and regulations governing consumer data**. The General Data Protection Regulation ("GDPR"), which was recently adopted by the European Union became effective in May 2018, **requires companies to satisfy new requirements regarding the handling of personal and sensitive data, including its collection, use, protection and the ability of persons whose data is stored to correct or delete such data about themselves. Complying with the GDPR caused us to update certain business practices and systems**. Non-compliance with GDPR could result in proceedings against us by governmental entities or others and fines up to the greater of €20 million or 4% of annual global revenue... As a result of GDPR, in particular, we may also experience **difficulty retaining or obtaining new European or multi-national customers**..."

Twitter 2018 10-K:

- "[W]e strive to **comply with applicable data protection laws and regulations**, as well as our own posted privacy policies and other obligations we may have with respect to privacy and data protection... A number of proposals have recently been adopted or are currently pending before federal, state and foreign legislative and regulatory bodies that could **significantly affect our business**... For example, the GDPR has been adopted and went into effect in May 2018. The **GDPR includes more stringent operational requirements for entities processing personal information** and significant penalties for non-compliance, including fines of up to €20 million or 4% of total worldwide revenue, whichever is higher."

Exhibit 2. Meta Conference Call Transcript Excerpts

Before GDPR:

- "[W]ith regards to GDPR and other initiatives around data usage... we believe that European MAU and DAU may be flat to slightly down sequentially in Q2 as a result of the GDPR roll out... that's just based on what we're expecting, given that you're having to bring people through these consent flows" - *David Wehner, CFO, (2018 Q1 earnings call transcript on April 25, 2018)*

After GDPR:

- "GDPR was an important moment for our industry. We did see a decline in monthly actives in Europe – down by about 1 million people as a result." - *Mark Zuckerberg, CEO, 2018 Q2 earnings call transcript on July 25, 2018)*

Exhibit 3. Annual report excerpt about awareness and anticipation of GDPR as early as its 2016 passage

Activision Blizzard 2016 10-K (with fiscal year end 12/31/2016):

- "For example, the **E.U.'s General Data Protection Regulation (the "GDPR")**, which will come into effect in May 2018, imposes a range of new compliance obligations for us and other companies with European users, and increases financial penalties for noncompliance significantly."

Exhibit 4. Annual report excerpts about enforcement of GDPR

<u>Activision Blizzard 2018 10-K:</u>

- "The laws and regulations concerning data privacy are continually evolving. **Failure to comply with these laws and regulations could harm our business**...
- We are subject to laws from a variety of jurisdictions regarding privacy and the protection of this information, including the E.U.'s General Data Protection Regulation (the "GDPR")...
- If we fail to comply with our posted privacy policies, EULAs, or terms of service, or if we fail to comply with existing privacy-related or data protection laws and regulations, it **could result in proceedings or litigation against us by governmental authorities or others, which could result in fines or judgments against us, damage our reputation, impact our financial condition, and harm our business**. If regulators, the media, consumers, or employees raise any concerns about our privacy and data protection or consumer protection practices, even if unfounded, this **could also result in fines or judgments against us, damage our reputation, negatively impact our financial condition, or damage our business**."

<u>Groupon 2018 10-K:</u>

- "**Failure to comply with** federal, state and **international privacy laws and regulations**, or the expansion of current or the enactment of new privacy laws or regulations, **could adversely affect our business**...
- **Noncompliance could result in proceedings against us by governmental entities or others and fines. For example, fines under GDPR could be up to the greater of €20 million or 4% of annual global revenue and damage our reputation and brand**...
- **Any failure, or perceived failure, by us to comply with our posted privacy policies or with any data-related consent orders, Federal Trade Commission requirements or orders or other federal, state or international privacy or consumer protection-related laws, regulations or industry self-regulatory principles could result in claims, proceedings or actions against us by governmental entities or other third-parties or other liabilities, which could adversely affect our business. In addition, a failure or perceived failure to**

comply with industry standards or with our own privacy policies and practices could result in a loss of subscribers or merchants and adversely affect our business."

Twitter 2018 10-K:

- "[T]he European Union, or EU, and its member states traditionally have taken broader views as to types of data that are subject to privacy and data protection, and have imposed greater legal obligations on companies in this regard. For example, the GDPR has been adopted and went into effect in May 2018. The GDPR includes **more stringent operational requirements** for entities processing personal information and **significant penalties for non-compliance, including fines of up to €20 million or 4% of total worldwide revenue, whichever is higher**."

Appendix F. Manual Classification of Innovation Job Roles

I present my classification, through manual textual analysis, of different types of job roles into innovation strategy types.

Job Roles (50 Role Granularity)	Job Roles (150 Role Granularity)	Job Category	Dummy for Exploratory Job
Legal	Compliance	Admin	.
Legal	Fraud Analyst	Admin	.
Legal	Legal	Admin	.
Recruiter	Recruiter	Admin	.
Coordinator	Coach	Admin	.
Coordinator	Coordinator	Admin	.
Coordinator	Student Intern	Admin	.
Coordinator	Corporate Trainer	Admin	.
Development Manager	Training Facilitator	Admin	.
Development Manager	Development Manager	Admin	1
Operations Administrator	Operations Administrator	Admin	0
Operations Administrator	Support Staff	Admin	.
Operations Administrator	Information Specialist	Admin	.
Operations Administrator	Facilities Manager	Admin	.
Human Resources Specialist	HR Business Partner	Admin	.
Human Resources Specialist	Payroll Specialist	Admin	.
Human Resources Specialist	Human Resources Specialist	Admin	.
Human Resources Specialist	Benefits Specialist	Admin	.
Cashier	Cleaner	Sales	.
Cashier	Conseiller commercial	Sales	.
Cashier	Pharmacy Technician	Sales	.
Cashier	Cashier	Sales	.
Crew Member	Crew Member	Sales	.
Crew Member	Restaurant Manager	Sales	.
Merchandiser	Stylist	Sales	.
Merchandiser	Merchandiser	Sales	.
Receptionist	Security Specialist	Sales	.
Receptionist	Pilot	Sales	.
Receptionist	Receptionist	Sales	.
Retail Sales	Genius	Sales	.
Retail Sales	Branch Manager	Sales	.
Retail Sales	Sales Service Representative	Sales	0
Retail Sales	Loss Prevention	Sales	.
Retail Sales	Retail Sales	Sales	.
Product Manager	Product Manager	Sales	1

Product Manager	Corporate Strategy	Sales	1
Product Manager	Market Research	Sales	1
Sales Associate	Customer Success Specialist	Sales	0
Sales Associate	Sales Engineer	Sales	0
Sales Associate	Account Manager	Sales	1
Sales Associate	Inside Sales	Sales	.
Sales Associate	Commercial Manager	Sales	1
Sales Associate	Sales Associate	Sales	0
Customer Service	Sales Support	Sales	0
Customer Service	Customer Support	Sales	0
Customer Service	Customer Service	Sales	0
Customer Service	Vendor Management	Sales	0
Customer Service	Customer Experience Specialist	Sales	0
Customer Service	Subject Matter Expert	Sales	0
Sales Representative	Sales Representative	Sales	.
Sales Representative	Sales Training Specialist	Sales	.
Solutions Specialist	Solutions Specialist	Sales	.
Accountant	Accountant	Finance	.
Accountant	Auditor	Finance	.
Accountant	Financial Controller	Finance	.
Accountant	Financial Analyst	Finance	.
Client Services	Client Services	Finance	.
Client Services	Business Support	Finance	.
Claims Specialist	Collections Specialist	Finance	.
Claims Specialist	Claims Specialist	Finance	.
Financial Advisor	Financial Advisor	Finance	.
Financial Advisor	Wealth Manager	Finance	.
Financial Advisor	Agent	Finance	.
Financial Advisor	Property Manager	Finance	.
Financial Advisor	Realtor	Finance	.
Financial Advisor	Banker	Finance	.
Billing Specialist	Billing Specialist	Finance	.
Investment Specialist	Investment Specialist	Finance	.
Investment Specialist	Economist	Finance	.
Investment Specialist	Risk Analyst	Finance	.
QA Tester	QA Tester	Engineer	0
Technician	Technical Support Engineer	Engineer	0
Technician	Technician	Engineer	0
Data Analyst	Data Scientist	Engineer	0
Data Analyst	Data Analyst	Engineer	0
Data Analyst	Business Analyst	Engineer	1

IT Specialist	Technical Support	Engineer	0
IT Specialist	Network Specialist	Engineer	0
IT Specialist	IT Specialist	Engineer	0
IT Specialist	Database Administrator	Engineer	0
IT Specialist	IT Analyst	Engineer	0
IT Specialist	Application Support	Engineer	0
Delivery Manager	Business Process Specialist	Engineer	0
Delivery Manager	Service Delivery Manager	Engineer	0
Delivery Manager	Delivery Manager	Engineer	0
Delivery Manager	Transformation Specialist	Engineer	0
Machine Operator	Machine Operator	Engineer	0
Machine Operator	Foreman	Engineer	0
Machine Operator	Mechanic	Engineer	0
Quality Assurance	Laboratory Technician	Engineer	0
Quality Assurance	Quality Assurance	Engineer	0
Software Engineer	Technology Analyst	Engineer	0
Software Engineer	Software Developer	Engineer	1
Software Engineer	Systems Engineer	Engineer	1
Software Engineer	Software Engineer	Engineer	0
Software Engineer	Technology Lead	Engineer	1
Software Engineer	DevOps Engineer	Engineer	0
Software Engineer	Web Developer	Engineer	.
Software Engineer	Data Engineer	Engineer	0
IT Project Manager	SAP Consultant	Engineer	0
IT Project Manager	IT Project Manager	Engineer	1
Mechanical Engineer	Automation Engineer	Engineer	0
Mechanical Engineer	Mechanical Engineer	Engineer	1
Mechanical Engineer	Structural Engineer	Engineer	0
Mechanical Engineer	Process Engineer	Engineer	1
Mechanical Engineer	Electrical Engineer	Engineer	.
Production Operator	Production Operator	Engineer	0
Production Operator	Manufacturing Associate	Engineer	0
Production Operator	Quality Engineer	Engineer	0
Technical Architect	Documentation Specialist	Engineer	0
Technical Architect	Construction Manager	Engineer	0
Technical Architect	Contracts Specialist	Engineer	0
Technical Architect	Technical Architect	Engineer	1
Technical Architect	Planning Manager	Engineer	0
Application Engineer	Test Engineer	Engineer	0
Application Engineer	Application Engineer	Engineer	1
Infrastructure Engineer	Infrastructure Engineer	Engineer	0

Infrastructure Engineer	Information Security	Engineer	.
Sustainability Specialist	Sustainability Specialist	Engineer	.
Sustainability Specialist	Safety Officer	Engineer	.
Designer	Graphic Designer	Marketing	1
Designer	Technical Writer	Marketing	.
Designer	UX Designer	Marketing	.
Designer	Designer	Marketing	1
Producer	Writer	Marketing	.
Producer	Content Specialist	Marketing	.
Producer	Producer	Marketing	.
Marketing	Digital Marketing Specialist	Marketing	.
Marketing	Brand Manager	Marketing	.
Marketing	Marketing	Marketing	.
Communications Specialist	Communications Specialist	Marketing	.
Communications Specialist	Public Relations	Marketing	.
Geologist	Geologist	Scientist	.
Scientist	Scientist	Scientist	.
Medical Rep	Case Manager	Scientist	.
Medical Rep	Medical Rep	Scientist	.
Clinical Research Associate	Regulatory Affairs Associate	Scientist	.
Clinical Research Associate	Clinical Research Associate	Scientist	.
Logistics	Procurement Specialist	Operations	0
Logistics	Logistics	Operations	0
Project Manager	Project Manager	Operations	1
Project Manager	Project Administrator	Operations	0
Operations Manager	Operations Manager	Operations	1
Operations Manager	AM	Operations	1
Operations Manager	Officer	Operations	1
Operations Manager	MD	Operations	1
Business Operations	Business Operations	Operations	0
Distribution Specialist	Operations Coordinator	Operations	0
Distribution Specialist	Distribution Specialist	Operations	0
Distribution Specialist	Driver	Operations	0

Appendix G. Examples of Exploratory and Exploitative Jobs

I present examples of job postings for a firm in my sample, Coca-Cola, that represent jobs that contribute to exploratory and exploitative innovation.

Exploratory Job Role: Coca-Cola's "Scientist IV"

Scientist IV

The Coca-Cola Company · Atlanta, GA · 4 days ago · **Be an early applicant**

On-site · Full-time

10,001+ employees · Food and Beverage Services

Skills: Cell Biology, High-Performance Liquid Chromatography (HPLC), +8 more

About the job

Duties

The Coca Cola Company has an opening for a Scientist IV in Atlanta, GA. Duties are to be responsible for leading all aspects of new product development from concept to commercialization. Work closely with several cross-functional groups (Marketing, Knowledge & Insights, Technical Operations, Quality, etc.) to lead and influence new product innovation in Retail & Food Service Channels. The position will perform the following:

- Responsible for planning, organizing, and executing new product development activities to support business growth and continuity. Designs, tests and execute activities/processes to
- ensure that all cycles of product development process are executed successfully and on time.
- Provide technical leadership and expertise to projects from inception through commercialization and launch. Supports other development projects, operations and related activities as necessary.
- Responsible for creating and managing technical timeline, ensuring product brief is complete and aligned upon, ensures risks analysis is complete and communicated. Executes projects within given timelines to meet desired business objectives (consumer, financial, operational, etc.). Exercises deep cross-functional involvement and alignment with project teams and management.
- Apply scientific, engineering and consumer insights principles in research, formulation, processing and commercialization of beverages.
- Identify, evaluate, and select new ingredients, developments, trends & technologies in food and beverage industry to guide product development initiatives. Partners with internal and external parties to stay abreast of our consumer's needs and identifies actions plans to address those needs. Drives and leverages internal and external suppliers to deliver ingredient and product innovation.

Education

Requires a minimum of a Master's Degree, or foreign equivalent, in Food Science, Chemical Engineering, or related field plus three years of experience in job offered, related position, or research assistant experience while pursuing a post-bachelor's degree.

Experience

Experience must include 3 years of experience developing and launching beverage or food products for a food or beverage company; 3 years of experience evaluating the chemistry of flavors, food ingredients, processing, packaging, nutrition, and stability testing. 3 years of experience working within a food science laboratory. 3 years of experience working with cross-functional teams ; 3 years of experience leading a food science laboratory and junior scientists.

Skills

Product Development; Leadership; Chemistry; Influencing; Continual Improvement Process; Environmental Science; Sensory Processing; Researching; organization; Microbiology Laboratory; Project Management; Food Sciences; Food Technology; Communication; Laboratory Testing; Data Compilation; Quality Control (QC); Food Safety Management; Green Solutions

Exploitative Job Role: Coca-Cola's "Advisor for Customer Logistics and Supply Chain"

Advisor, Customer Logistics and Supply Chain

The Coca-Cola Company · Atlanta, GA · 19 hours ago

Full-time

10,001+ employees · Food and Beverage Services

Skills: Attention to Detail, Social Influence, +8 more

About the job

This role is responsible for providing dedicated support for our Retail, Bottler, Distributor, McDonald's customers and Business partners by processing orders and inquiries within the Order To Cash (OTC) Supply Chain organization. The Advisor, Customer Logistics and Supply Chain Specialist will research, influence and resolve issues for our Customers and Business partners using Coca-Cola North America (CCNA) order management system & tools based on our established service level agreements.

Key Activities

- Single point of contact for our customers across Business units
- Partner with sales teams and brokers to address customer satisfaction needs related to pricing, delivery method changes, account management and strategic projects (product launches, packing changes etc.)
- Partner with Product supply, Supply planning, logistics and transportation teams to influence and resolve order flow concerns (late trucks, dwell time, refused deliveries)
- Responsible for managing phone calls and emails from customers, internal stakeholders and partners and accurately tracking and resolving the business need
- Customer order management – Acts as liaison between external departments and customers to process order changes, communicate all order changes, track and reschedule customer orders to meet customer requested delivery dates
- Apply best practices and business knowledge to make timely decisions and resolve issues with product orders to meet requested customer delivery dates
- Proactively manage order flow, communication of all In Full issues and partner with transportation to ensure delivery appointments and pick-up appointments are set timely to mitigate order failures
- Recognize and track trends with customer issues, evaluate and suggest process improvements for Product order management and our partners
- Understand Company goals and performance metrics and improve quality and speed to meet and exceed customer expectations
- Work collaboratively but often independently on daily tasks and resolve escalations that are diverse in scope using strategic thinking, people and company resources
- Build and leverage collaborative relationships with our customers, sales, logistics and other internal partners based on customer specific operations
- Analyze, coordinate and participation in implementation of design projects to support department and or key customer objectives.

Required Skills/Experience

- Strong customer service and relationship background, strategic thinking, effective communication and business writing skills
- Must have strong research or analytical skills, attention to detail, effective problem solver and influencing skills
- Must have a strong sense of urgency and be empowered to make timely and informed decisions
- Building collaborative relationships within Product Order Management (POM), Supply chain/OTC organization and our external partners
- Knowledge of the CCNA Product Supply system, Order Management and Salesforce & SAP a plus. Other systems used include Customer Relationship Management (CRM) Database, MS Office; Outlook, Word, Excel, SharePoint, Genesys
- Ability to excel and contribute in a fast paced and changing work environment with accuracy while meeting Service Level Agreement (SLA's) and deadlines
- Ability to understand overall Company objectives and manage competing project and tasks
- Roles within this organization include teleworking and remote working. Candidate must be able to work independently with minimal daily supervision

Types of additional activities managed

- Expedited Orders
- Customer Order Change Request
- Customer Pick Up
- Transportation Issues
- Dock Cuts
- Ship Withs
- General Product Information
- Plant Redirects
- Over, short and damaged process
- Transportation Method Change
- Duplicate, missing, or incorrect orders