# Three Essays on the Consequences of Disclosure


Dissertation

zur Erlangung des akademischen Grades
doctor rerum politicarum
(Doktor der Wirtschaftswissenschaft)

eingereicht an der
Wirtschaftswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von
Sarah Kröchert, M.Sc.

Präsidentin der Humboldt-Universität zu Berlin:
Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Wirtschaftswissenschaftlichen Fakultät:
Prof. Dr. Christian D. Schade

Gutachter: 1. Prof. Dr. Joachim Gassen
2. Dan Amiram, Philip H. Geier Associate Professor of Business

Tag des Kolloquiums: 21. Dezember 2017

## To the Reader

Financial disclosure regulation is paramount. In particular after the recent financial crises, jurisdictions around the globe have revamped the financial disclosure environment with the objective to re-establish trust in financial markets. Yet, we know surprisingly little about how these disclosure requirements actually work. Do they help institutional investors to structure their portfolios? Do ownership disclosure requirements trigger rebalancing decisions by investors? How does the international diversity of regulations interact with the labor mobility of the professionals that have to deal with them?

These are the questions that Sarah Kröchert is exploring in her thesis. Although all questions address consequences of disclosure, they capture widely different aspects of that theme. While this can be understood as an indication of how far-reaching the consequences of financial disclosure regulation are, it is also an impressive indication of Sarah's widely spread research interests.

Her findings are fascinating. She documents that institutional investors cater to accounting-related notions of comparability, indicating that one of the key features of financial reporting is valued by the investment community. This is relevant as there is an ongoing debate in the financial accounting community about whether to strengthen the comparability or the individual informativeness of financial accounting information. In her second study, she leverages her institutional expertise to address a key question related to ownership disclosures. Using proprietary data from the Deutsche Bundesbank, she studies whether institutional investors manage their ownership stakes to stay "below the radar" of mandatory ownership disclosures (they do). Her setting is neat for two reasons: Since the data collected by Deutsche Bundesbank is independent from public ownership disclosures, she is capable to look "behind the curtain". In addition, as she can observe the complete ownership structure, she can also provide some insights on who is picking up the discarded ownership stakes. In her final study, which is co-authored with Ulf Brüggemann and myself, we explore the determinants of short-term international labor mobility in the professional services sector. The direct labor market consequences of financial market regulation have been largely overlooked by prior work. Yet again, the international regulatory harmonization can be expected to affect the mobility of the professionals in that field. While our findings are exploratory in nature, they are consistent with regulatory harmonization positively affecting the short-term labor mobility of affected employees.

Sarah Kröchert has developed an impressive research program that provides a holistic view on what disclosure regulation can and cannot achieve. Understanding the depth and interconnectedness of regulations is highly relevant to regulators, market participants and academics alike. The work of Sarah Kröchert contributes to this debate by using fresh data and designs to answer new questions. I hope that her studies will be widely read and used.

Berlin, May 19, 2018

Joachim Gassen

## Acknowledgements

This dissertation is the result of my Ph.D. studies in Accounting at Humboldt University of Berlin. I am grateful to the German Research Foundation (DFG), which supported my initial years in a structured Ph.D. program, and to the Caroline von Humboldt Grant Programme, which provided the financial means for the stay at Columbia University in New York in the final year of my Ph.D. Yet, most importantly, I am grateful to the large number of people who I met in all those years and who helped me understand that academia is, professionally, the right place for me.

I am indebted to the members of my dissertation committee. I thank my supervisor, Joachim Gassen, for sharing his tremendous amount of knowledge, not only when it comes to research, but also when it comes to the workings of the profession. I thank Dan Amiram for his immediate willingness to support my stay at Columbia, his constant availability during my job search and his trust in my abilities. I thank Ulf Brüggemann for his limitless time and his remarkable readiness to listen to anything from minor econometric problems to the darker moments in the final months of the Ph.D. I thank Roland Strausz for providing me with insights into the analytical mind. All four of them demonstrate exceptional dedication to their research and, by doing this in very different ways, each one of them gave me the chance to learn something.

I equally want to thank my fellow Ph.D. students. There are likely only a few people as personable and understanding as Nader Hemaidan is. And there are likely only a few people as enthusiastic and brilliant as Maximilian Muhn is. Those two were around during most of my time in Berlin and I am very happy about the experiences we made together. The group of Ph.D. students that introduced me to academic life, primarily Jens Günther, Marcus Witzky and Jochen Pierk, keeps reminding me of the right balance between work and private life, which at times involves your colleagues. The group of Ph.D. students that currently struggles with their dissertation, Rico Chaskel, Janja Brendel, Tom Fischer and Tobias Witter, as well as our post-doc, Martin Bierey, keep showing me what the Ph.D. is about: diligence, hard work and, at the same time, fun.

All this time in Berlin has been extremely enjoyable and a large portion of the good moments is to be attributed to our amazing secretaries, Andrea Häußler and Heidlinde Völker, who, relentlessly, made and make sure that the slightly chaotic academics do not forget about their duties.

There are many other people from inside and outside of academia who had an impact on my decisions, my work and my development; our student assistants, the accounting group at Columbia, researchers I spoke to on conferences, my friends and my family. I truly value having spent moments with all of them and I have to apologize for not mentioning every single name.

The persons I do mention in the final paragraph stand out because of their patience with me and their confidence in my work. They helped me in many of the more demanding situations. Among my friends there is one person who not only knows exactly when I need a break (as many do), but has gone through endless versions of my papers. Anna has pointed me towards inconsistencies in my write-ups without having ever attended an accounting class. Thanks for always being there, no matter what it is about. Finally, I would not be where I am without my

family. My sister does not need words to understand me. My mother is one of the strongest women I have met so far in my life. And my father is the one I would have loved to have around in so many moments in the last years. To him, I dedicate this dissertation.

Berlin, March 2018

Sarah Kröchert

# Contents

## Introductory Summary

The requirement to disclose information can affect behavior, both of the disclosing agent and others consuming the disclosure. The disclosing agent often reveals information about his own actions and thus makes himself susceptible to critical evaluations. Apart from reputational consequences, he regularly faces monetary incentives to present himself in a particular manner. Agents consuming the disclosure obtain, directly or indirectly, more information, which they can incorporate into their decision-making. Jin and Leslie [2003], for instance, provide evidence for consequences of the public disclosure of restaurant hygiene scores, for the restaurants subject to disclosure and their guests. Notably, the authors study the introduction of the dislosure regime, not the underlying hygiene inspections, from which the scores are derived and which are in place before. Their findings suggest that the mere requirement to disclose incentivizes restaurants to improve hygiene, serving, in turn, to increase revenues. At the same time, potential guests can better choose the restaurant that matches their preferences [e.g., Akerlof, 1970; Spence, 1973]. This thesis comprises three essays on the consequences of mandatory, financial disclosures. In each essay, the agent of interest assumes a different role; he is primarily disclosing, consuming disclosures or supporting the preparation of the same.

The agents of interest are sophisticated, knowledgeable and well informed. They have the necessary expertise to understand financial disclosures and the impact of disclosing certain items. Despite their expertise, even they are limited in their capabilities of processing information. The benefits of fully grasping content are frequently offset by the costs of acquiring relevant pieces of information. In the first and second essay, I focus on one group of sophisticated capital market participants, institutional investors. Although subsumed by a common term, they are heterogeneous entities [e.g., Cronqvist and Fahlenbrach, 2009; Edmans, 2014]. Some are forthcoming about their own activities, others are not; some deal intensively with the activities of third parties, others do not. In the final essay, I concentrate on highly skilled professionals who assist in the preparation of financial disclosures, employees in the accounting industry. Accounting is characterized by a comparatively high density of rules and a high degree of standardization [e.g., Madsen, 2011]. Across countries, these rules can differ and, in this vein, change incentive patterns in other markets, such as the international labor market.

The essays are all empirical in nature and take methodological approaches aimed at gaining insight into the questions at hand. They do so by exploiting settings that offer variation in the constructs of interest along with data sources with a sufficient level of detail. As a result, settings range from single countries, Germany or the United States (US), to a broad set of country pairs. They encompass disclosure regulation that varies in the cross section or over

time, in part allowing for causal inference strategies. Finally, they entail data of a granularity rarely found in academic studies.

The first essay, "Ownership Disclosure and Ownership Structure: Investors' Response to Lower Reporting Thresholds", assesses the reaction of disclosing agents to a change in disclosure requirements. Disclosure instruments are ownership disclosures, timely, non-periodic announcements, through which investors are to reveal their holdings. Rules refer to all investors on regulated capital markets, but only to certain holdings in that they condition disclosure on crossing reporting thresholds. Hence, disclosing agents have some discretion as to whether to disclose at all. Remaining below reporting thresholds does not trigger any disclosure requirement. Disclosing agents also have potential incentives to avoid disclosure. Transparency can be costly, for example, if it involves the revelation of privately generated information [e.g., Fishman and Hagerty, 1995; Wermers, 2001]. These costs should be especially pronounced for sophisticated investors with holdings close to reporting thresholds. I therefore differentiate between different types of (institutional) investors and investigate whether they adjust their holdings around disclosure changes.

The study exploits the German setting for three reasons. First, the German regulator lowered reporting thresholds, enabling the analysis of rule changes that are plausibly exogenous from the perspective of the investor. Second, I can work with data on privately reported and publicly disclosed holdings of all different investor types in the market. The comprehensive data coverage makes it possible to analyze which investors decrease their holdings, but also which investors are on the other side of the trades. Third, the German stock market has two segments, the regulated and unregulated market, of which only one is subject to the ownership disclosure regime. Consequently, I conduct a difference-in-differences analysis, with the rule changes as treatment and unaffected holdings as control group.

Findings show that mutual funds and banks, sophisticated investor types with small stakes, reduce their holdings while non-financial corporations, typical blockholders whose stakes are public knowledge, increase their holdings. They underline that, in the extreme, changes in ownership disclosure can induce changes in ownership structure. By this means, they can influence other firm outcomes, e.g., stock liquidity as discussed below or corporate governance structures as discussed in extant work [e.g., Appel, Gormley and Keim, 2016; Edmans, 2014]. Further analyses of mutual funds, the investor type with the most robust results, show that mutual funds tend to concentrate their holdings just below the initial reporting threshold, consistent with disclosure avoidance incentives driving the observed reduction in their holdings. Lastly, I address liquidity consequences of mutual funds' response. The reduction in their holdings ap-

pears to weaken the positive relation between disclosure and stock liquidity that is documented in prior literature [e.g., Agarwal, Mullally, Tang and Yang, 2015; Christensen, Hail and Leuz, 2016].

The second essay, "Accounting Comparability in Mutual Funds' Portfolios", centers on mutual funds in their role as agents who consume disclosures. Precisely, as agents who process and evaluate periodic financial statements issued by potential and existing portfolio firms. Mutual funds, or rather their fund managers, need accounting information to make investment decisions. In addition, they need accounting information that is comparable and standardized to benchmark firms against each other. Although they certainly have the expertise to make financial statements comparable on their own, they increasingly rely on automated data processing for large and diversified portfolios, implying that they might demand ex ante similar accounting information. In this study, I examine whether accounting comparability of potential and existing portfolio firms affects mutual funds' investment decisions.

I resort to the US market because of the availability of frequent portfolio disclosures of mutual funds. I measure accounting comparability of firms in line with De Franco, Kothari and Verdi [2011], but extend the concept in two respects. Next to the return-based model proposed by De Franco et al., I employ a cash flow-based model. Moreover, I emphasize the investor perspective. The assessment of comparability requires the definition of a peer group. Since mutual funds decide about the composition of their portfolios, the most relevant peers, from their perspective, should be portfolio firms. The comparability proxies thus capture accounting similarities with portfolio peers. They are defined at the holding level and vary with firms' own accounting practices as well as the set of portfolio peers they are compared to.

The analysis at first establishes that accounting comparability is high in mutual fund portfolios. As benchmark, I use analyst portfolios, constructed from analysts' coverage decisions. Analyst portfolios are ambitious benchmarks in that analysts themselves should benefit from covering firms that are, in accounting terms, more comparable [e.g., De Franco et al., 2011; Neel, 2017]. Yet, findings show that comparability is higher in mutual fund portfolios. Among mutual funds, comparability is also higher in nonindexer than in indexer portfolios, strengthening the economic rationale. If similar accounting information facilitates portfolio management, comparability effects should be concentrated in portfolios of nonindexer funds that actively manage their portfolios. The analysis then turns to the source of the comparability effects and mutual funds' investment decisions. I can show that the probability of a firm being included into a portfolio is increasing in accounting comparability. Besides, I study comparability around the point in time when the firm is included. Until inclusion into the portfolio, accounting compa-

rability increases, but does not change much afterwards. Taken together, the findings suggest that mutual funds value accounting comparability in selecting investments. The findings do not suggest that mutual funds induce changes in comparability of their portfolio firms once they are invested.

The third essay, "Temporary Migration within Multinational Corporations: Evidence from the Accounting Industry" (co-authored with Ulf Brüggemann and Joachim Gassen), focuses on agents that support the preparation of financial disclosures. Regulations governing accounting and tax disclosures are complex in themselves and differ across countries. Auditors and (tax) consultants assist in setting up the necessary infrastructure in firms. Their expertise originates from both specialization in local rules and international experience with different institutional environments. Mobility in an international context may thereby depend on the similarity in the respective regulatory frameworks. In this study, we provide initial evidence on the relative importance of these occupation-level determinants for temporary, within-firm migration.

We explore a proprietary dataset of a large accounting firm, informing about all international assignments of its employees. The data are granular in the sense that they allow to observe mobility in country pairs. We initially verify that typical, country-level migration determinants also play a role for the type of mobility in our setting; temporary, within-firm migration of the highly skilled. We can show that mobility is higher in country pairs with cultural links and economic ties and country pairs that are geographically closer. In the second step, we test whether the similarity in accounting and tax regimes in country pairs, the factor specific to the occupations, is associated with the decision to migrate. We can show that mobility is increasing in the similarity of rules, in particular for auditors. We hence add to the still evolving literature on temporary migration of (accounting) professionals [e.g., Beaverstock, 2017; Bloomfield, Brüggemann, Christensen and Leuz, 2017] by documenting that occupation-level factors are related to mobility patterns.

## References


Agarwal, V., Mullally, K. A., Tang, Y. and Yang, B. [2015], 'Mandatory portfolio disclosure, stock liquidity, and mutual fund performance', *The Journal of Finance* **70**(6), 2733–2776.

Akerlof, G. A. [1970], 'The market for "lemons": Quality uncertainty and the market mechanism', *The Quarterly Journal of Economics* **84**(3), 488–500.

Appel, I. R., Gormley, T. A. and Keim, D. B. [2016], 'Passive investors, not passive owners', *Journal of Financial Economics* **121**(1), 111–141.

Beaverstock, J. V. [2017], 'Migration: Skilled international labor', *The International Encyclopedia of Geography*.

Bloomfield, M. J., Brüggemann, U., Christensen, H. B. and Leuz, C. [2017], 'The effect of regulatory harmonization on cross-border labor migration: Evidence from the accounting profession', *Journal of Accounting Research* **55**(1), 35–78.

Christensen, H. B., Hail, L. and Leuz, C. [2016], 'Capital-market effects of securities regulation: Prior conditions, implementation, and enforcement', *The Review of Financial Studies* **29**(11), 2885–2924.

Cronqvist, H. and Fahlenbrach, R. [2009], 'Large shareholders and corporate policies', *The Review of Financial Studies* **22**(10), 3941–3976.

De Franco, G., Kothari, S. P. and Verdi, R. S. [2011], 'The benefits of financial statement comparability', *Journal of Accounting Research* **49**(4), 895–931.

Edmans, A. [2014], 'Blockholders and corporate governance', *Annual Review of Financial Economics* **6**(1), 23–50.

Fishman, M. J. and Hagerty, K. M. [1995], 'The mandatory disclosure of trades and market liquidity', *The Review of Financial Studies* **8**(3), 637–676.

Jin, G. Z. and Leslie, P. [2003], 'The effect of information on product quality: Evidence from restaurant hygiene grade cards', *The Quarterly Journal of Economics* **118**(2), 409–451.

Madsen, P. E. [2011], 'How standardized is accounting?', *The Accounting Review* **86**(5), 1679–1708.

Neel, M. [2017], 'Accounting comparability and economic outcomes of mandatory IFRS adoption', *Contemporary Accounting Research* **34**(1), 658–690.

Spence, M. [1973], 'Job market signaling', *The Quarterly Journal of Economics* **87**(3), 355–374.

Wermers, R. [2001], 'The potential effects of more frequent portfolio disclosure on mutual fund performance', *Investment Company Institute Perspective* **7**(3), 1–11.

# Ownership Disclosure and Ownership Structure: Investors' Response to Lower Reporting Thresholds*

**Sarah Kröchert**
Humboldt University of Berlin

***Abstract***

The study investigates which investors adjust their holdings around the introduction of lower reporting thresholds for ownership disclosures. I work with a dataset that includes both publicly disclosed and undisclosed holdings. I exploit segment-specific rules in the stock market that provide variation in disclosure regimes. Mutual funds and, to a lesser extent, banks decrease while non-financial corporations increase their holdings. I then examine in greater detail whether the reduction in mutual funds' holdings is consistent with the avoidance of reporting thresholds. In the public disclosures, mutual funds concentrate their holdings just below the initial threshold. They also seem to react gradually over time, supporting a response to stricter enforcement. Finally, I study liquidity consequences. The reduction in mutual funds' holdings appears to weaken the previously documented positive relation between disclosure and stock liquidity. The findings suggest that changes in ownership disclosure can induce changes in ownership structure.

*I am grateful to Deutsche Bundesbank for the provision of the Securities Holdings Statistics; to EQS Group for the notifications; to Martin Schmidt and Richard Stehle from Humboldt University of Berlin for the list of ISIN changes; and to Erik Theissen, Esad Smajlbegovic and Thomas Johann from the University of Mannheim for the translation file for historical and current ISIN. I thank Dan Amiram, Martin Bierey, Mary Billings, Ulf Brüggemann, Joachim Gassen, Alon Kalay, Maximilian Muhn and Peter Pope for their thoughtful feedback. I further thank workshop participants at Humboldt University, Columbia University, the 2017 EAA Doctoral Colloquium, the 2017 EAA Annual Congress and the 2017 AAA Annual Meeting for helpful comments.

# 1 Introduction

There is evidence that investors prefer to avoid the disclosure of their holdings. Under a transparency regime that conditions on reporting thresholds, such a preference can translate into the accumulation of holdings just below the disclosure threshold. Academic research suggests, for instance, that transparency regimes for short sales result in investors keeping positions low and hence avoiding their public disclosure [e.g., Jank, Roling and Smajlbegovic, 2016; Jones, Reed and Waller, 2016]. Investors themselves regularly emphasize that disclosure is costly and regulators care about potential reactions to these costs. In discussions about the adequate level of ownership disclosure, the European Commission acknowledges that "[l]arge institutional investors [...] have a tendency to remain below the regulatory disclosure threshold in listed companies" [European Commission, 2008, p. 9]. This study examines a setting in which the regulator lowers reporting thresholds for ownership disclosures. While the above rationale implies that, in response, investors reduce their holdings, it is incomplete in that it cannot describe the behavior of *all* market participants. Widespread selling, even if only by institutional investors, induces price pressure. In the end, some investors need to take the other side of the trades and thus add to their holdings. Ex ante, it is not obvious who ends up on which side of the market. I investigate which investors decrease and which investors increase their positions. In addition, I study the liquidity consequences of the reaction to the regulatory changes.

Empirical research is relatively silent on the relation between ownership disclosure and ownership structure, which is mostly due to the frequent unobservability of investors' actions. If they successfully avoid disclosure, there is no public record of their holdings. Even if data are available, they usually relate to subsets of investors and impede inferences on the overall ownership structure [e.g., Aragon, Hertzel and Shi, 2013; Jank et al., 2016]. Yet, understanding the relation is important because it impacts firm-level outcomes. Ownership structure shapes stock liquidity or the corporate governance of firms. Studies show that different investors can affect firm decisions and capital market assessments, but, again, often focus on a particular investor type [e.g., Appel, Gormley and Keim, 2016; Boone and White, 2015; Edmans, 2014].

I explore the research question in the German setting, since it offers a rare combination of datasets. The data cover the holdings of (almost) all investors, including investors who successfully avoid disclosure. Precisely, they encompass *privately* reported and *publicly* disclosed holdings. I use privately reported holdings from the Securities Holdings Statistics (SHS), a database of Deutsche Bundesbank, the German central bank. Every quarter, German financial institutions inform the central bank about all holdings in shares deposited with them, separately for single securities and different investor types. In this dataset, I have holdings only at the level

of the investor type and not the individual investor. However, I can observe all actions of investors, disclosed and undisclosed, without reporting bias. It is this feature that distinguishes my study from prior work relying on assumptions about undisclosed choices [e.g., Agarwal, Gay and Ling, 2014; Jones et al., 2016] and that makes the data attractive for research in general [e.g., Baltzer, Jank and Smajlbegovic, 2015; Jank, 2011; Kick, Onali, Ruprecht and Schaeck, 2014]. In the SHS, I can differentiate between foreign investors and the following domestic investor types: mutual funds, banks, insurance companies and pension funds, non-financial corporations, governmental institutions, other financial investors and households. In what follows, I refer to these data as holdings. I use publicly available holdings from notifications disseminated by the DGAP, the main distributor in Germany. Notifications are the instrument through which investors disclose their ownership when they cross a threshold and which is directly affected by the regulatory changes.[1] Notably, they contain information at the level of the individual investor. In what follows, I refer to these data as notifications.

I expect that investors adjust their holdings in response to the introduction of lower reporting thresholds because disclosure entails costs. Disclosure can be costly, since it invites free riding and front running [e.g., Brunnermeier and Pedersen, 2005; Fishman and Hagerty, 1995]; raises agency costs [e.g., Edmans, Heinle and Huang, 2016; Prat, 2005]; and imposes direct administrative costs [e.g., European Commission, 2008, 2012]. As an example, Verbeek and Wang [2013] show that mimicking mutual funds' holdings from periodic reports can be profitable, consistent with free riders extracting rents by exploiting mutual funds' privately generated information. In contrast to periodic reports that follow fixed reporting intervals, notifications become public a few days after the triggering event, the trade that results in crossing a threshold. Their timeliness might make disclosure particularly costly. Disclosure can also be beneficial in that it accelerates the realization of returns [e.g., Frank, Poterba, Shackelford and Shoven, 2004; Ljungqvist and Qian, 2016]. Anecdotally, incidences of misreporting suggest that, at least for some investors, costs outweigh benefits. In my setting, the asset manager BlackRock repeatedly discloses notifications with errors or delay, effectively hiding information on its holdings.[2]

I argue that investors are heterogeneous and therefore face these costs and benefits to a different extent. However, a priori, I refrain from ranking them along their (net) costs as such a ranking is not obvious. For instance, mutual funds' agency relationship might involve consid-

[1] As described in Section 2.1, it is the issuer and not the investor who discloses the notification; the investor only notifies the issuer. Although slightly negligent, I ignore the role of the issuer.

[2] BlackRock turns itself in to the BaFin, the supervisory authority. For the reaction of the BaFin, see the official announcement at: www.bafin.de/SharedDocs/Veroeffentlichungen/EN/Meldung/2015/meldung_150320_bussgeld_blackrock_en.html.

erable disclosure costs. Fund investors closely monitor fund managers and punish undesirable actions by redeeming shares in a timely manner. Fund managers, in turn, ensure that disclosed holdings satisfy the current expectations of their investors and not necessarily the optimum [Prat, 2005]. At the same time, mutual funds' activities are already quite transparent through other channels, e.g., periodic portfolio disclosures or daily return reports. Albeit imperfectly, other market participants can guess which securities the fund manager trades. Alternatively, non-financial corporations, usually long-term strategic investors, likely prefer not to disclose their private information. Anecdotally, cases of hidden stake building imply that they regard the revelation of such information as costly.[3] Yet, existing rules typically require them to disclose their holdings before thresholds are lowered, again questioning incremental costs. Thus, I merely argue that costs differ across investors and that this heterogeneity is precisely what makes holding adjustments possible. Investors with the highest disclosure costs reduce their holdings below reporting thresholds and can do so *only* because there are other investors with lower disclosure costs who accept their shares.

In the holdings analysis, I examine whether investors adjust their holdings upon the introduction of stricter reporting thresholds in 2007, 2009 and 2012. I use investors' percentage stakes since thresholds are formulated as percentage of overall voting rights. The identification strategy makes use of the segmentation of the German stock market. Ownership disclosure rules apply only to firms listed on the regulated, not the unregulated market. Hence, I conduct a difference-in-differences analysis, comparing changes in investors' holdings in firms in the regulated market (treatment group) with changes in holdings in firms in the unregulated market (control group). Despite being relatively large, the unregulated market is the explicit segment choice of smaller and younger firms [Vismara, Paleari and Ritter, 2012]. I undertake two steps to improve comparability between treatment and control group. First, I drop from the treatment group the largest firms. Treated firms in the resulting sample are on average still larger than control firms, but substantially closer in size. This step should further shift the focus to firms for which holdings around reporting thresholds translate into typical investment amounts of institutional investors. The German Investment Funds Association (BVI) stresses that mutual funds might easily exceed thresholds with their holdings in midcaps.[4] Second, I work with different estimation approaches. I mainly rely on firm (and quarter-year) fixed effects. In additional specifications, I include firm-specific linear time trends and time-varying firm characteristics or

[3] In Germany, the most prominent example for hidden stake building is the Porsche–Volkswagen case, which becomes public in 2008. See, e.g.: http://www.economist.com/node/12523898.

[4] See their comments at: http://webarchiv.bundestag.de/archive/2009/0626/ausschuesse/a07/anhoerungen/033/Stellungnahmen/04-BVI.pdf.

restrict the sample to a balanced panel and matched subsets of firms.

Across the different specifications, mutual funds and, to a lesser extent, banks decrease while non-financial corporations increase their stakes. The result for mutual funds is pronounced for the disclosure increase in 2007, the strongest event for identification. In 2007, the regulator reduces the initial reporting threshold from 5% to 3%, in all likelihood prompting investors with on average small stakes, to which mutual funds belong, to consider adjustments. Theoretically, mutual funds' reaction is consistent with agency costs accounting for a sizeable component of overall disclosure costs. The result for banks is somewhat weaker in that it is less robust across models and not as clearly related to the disclosure increase in 2007. It supports the existence of other components of disclosure costs. Like mutual funds, banks are frequently trading investors with small stakes and hence exposed to free riding and front running. The result for non-financial corporations reveals who is on the other side of the market with presumably lower disclosure costs. Non-financial corporations typically trade less regularly and their holdings are often public knowledge under the previous disclosure regime. I then provide descriptive evidence that hints at the specific actions investors take around the disclosure increase in 2007. Findings so far show that some investors decrease their holdings. If the decrease is motivated by the threat of disclosure, these investors will only reduce and not completely sell off their holdings. Aggregate statistics are consistent with this rationale. Mutual funds and banks slightly increase the number of firms they are invested in and the market value of their holdings drops.

The above results suggest that an increase in ownership disclosure can provoke a shift in ownership structure in the market. The reduction in holdings that I document is large from the standpoint of the investor type. For mutual funds, the type with the most robust results, stakes in firms in the regulated market decline relative to stakes in firms in the unregulated market by 38bp or 17% of their pre-treatment level. In monetary terms, the decline corresponds to asset sales of around €2,074m or 1% of total market capitalization. Admittedly, a change in ownership of 0.38% might not seem large from the standpoint of assessing ownership structure as a whole. However, I cannot observe holding changes of all mutual funds. Among the investor types I resort to, foreign investors include foreign mutual funds. Moreover, I eliminate from the sample the largest and most visible firms, the preferred investment target of mutual funds [e.g., Dahlquist and Robertsson, 2001; Ferreira and Matos, 2008]. Thus, my results can well represent a lower bound.

In the next step, I focus on mutual funds and examine in greater detail whether the avoidance of reporting thresholds is the most likely explanation for the observed reduction in their holdings. I can show that mutual funds reduce their holdings gradually over time, indicative of adjustments

in response to stricter enforcement of ownership disclosure rules. I then turn to the notifications. The dataset starts with the disclosure increase in 2007, contains stakes of individual investors and distinguishes between mutual funds and other investors. In it, mutual funds report stakes more frequently just below 3% and less frequently at zero, relative to other investors. Both findings strengthen the disclosure avoidance argument: mutual funds appear to reduce their holdings to be below the initial reporting threshold and do not completely sell them off, which could also be motivated by, for instance, a change in investment strategy. Besides, mutual funds disclose a decreasing number of notifications for falling below a threshold over the sample period, relative to all notifications they submit. This finding is again in line with a delayed response because of initially weak enforcement. In the early years, there can still be (more) notifications by funds that intend to avoid disclosure, but adjust holdings not before the regulator displays sufficient scrutiny.

Finally, I study the liquidity consequences of the reduction in mutual funds' holdings. In principle, there are two contradictory effects. An increase in ownership disclosure improves the public information set. Information asymmetry should decline and liquidity should increase [e.g., Agarwal, Mullally, Tang and Yang, 2015; Huddart, Hughes and Levine, 2001; Kyle, 1985]. At the same time, I find that an increase in ownership disclosure results in mutual funds reducing their holdings, in part by selling to non-financial corporations. A decline in the number of frequently trading investors and a rise in the number of blockholders should impair liquidity [e.g., Demsetz, 1968; Heflin and Shaw, 2000; Rubin, 2007]. I first replicate results from prior literature and provide evidence for both channels separately. I then address their interaction and show that the positive impact on liquidity is mitigated or even completely offset for firms likely most affected by the reduction in mutual funds' holdings.

The paper contributes to the literature on the consequences of ownership disclosure. Theoretically, extant research analyzes how ownership disclosure rules impact strategies of (potentially) informed investors [e.g., Fishman and Hagerty, 1995; Huddart et al., 2001; John and Narayanan, 1997]. From a legal perspective, authors discuss how the design of existing regulation shapes the investor base in practice [e.g., Hu, 2015; Hu and Black, 2006; Schouten, 2010]. Empirically, studies suggest that investors adjust positions around periodic reporting dates [e.g., Agarwal et al., 2014; He, Ng and Wang, 2004; Lakonishok, Shleifer, Thaler and Vishny, 1991] and in response to increased, not necessarily periodic disclosure requirements [e.g., Agarwal, Vashishtha and Venkatachalam, 2017; Duong, Huszár and Yamada, 2015; Jank et al., 2016]. They further reveal that investors do not always revert to holding adjustments; in some settings, non-disclosure is an option [e.g., Agarwal, Jiang, Tang and Yang, 2013; Aragon et al., 2013]. Common to the

empirical work is the isolated study of specific subsets of investors. Investors need to find others that are willing to trade with them, which makes incentive structures in the market interdependent. I attempt to, at least in part, account for these interdependencies by simultaneously investigating all different investor types.

The remainder of the paper proceeds as follows. Section 2 characterizes the setting and formulates expectations. Section 3 describes the identification strategy, sample and data. Section 4 presents the main results for all investor types and additional analyses for mutual funds. Section 5 reports the results for liquidity consequences and Section 6 concludes.

# 2 Setting and Expectations

## 2.1 Institutional Setting

In Germany, firms listed on the regulated market are subject to ownership disclosure rules. I work with regulatory changes in 2007, 2009 and 2012 that mandate stricter reporting thresholds for the disclosure of ownership in these firms. All of them are well anticipated, evidenced, for instance, by early adoptions of the corresponding laws. In addition, all of them comprise various measures aimed at increasing transparency about ownership structures. In what follows, I focus on the main elements resulting in lower thresholds.

The instrument to inform about ownership is the notification of major holdings or, in short, notification. Notifications reveal the identity of the individual or entity that can exert influence on the issuer via the exercise of voting rights. To capture actual voting power, rules on attribution obligate not only the holder, but all investors with legal or de facto influence to report holdings [secs. 22 ff. of the Securities Trading Act (WpHG)].[5] Attribution rules are responsible for a dominance of large entities among reporting investors, such as fund families with many individual funds, as they usually require the aggregation of individual stakes at the entity level. The disclosure process involves both the investor, who first notifies the issuer, and the issuer, who then disseminates the notification. Disclosure occurs within a few days. Disclosed information encompasses, among other things, identifying information on the investor and issuer, the stake as percentage of overall voting rights and the date of the trade triggering the disclosure requirement [secs. 17, 19 of the Securities Trading Notification and Insider List Ordinance (WpAIV) in the version applicable until 11/26/2015]. During the sample period, disclosure formats are

[5] To be precise, there are very few exceptions for which rules on attribution do not obligate the holder, but only other investors to report, e.g., shares provided as security [BaFin, 2013*b*]. In general, rules on attribution follow the principle of mutual attribution, i.e., all involved parties have to notify.

not standardized and vary in practice; Appendix A.2 presents one example for a typical, short notification.[6] Prior research shows that the disclosure of notifications is associated with abnormal returns, suggesting that the information therein constitutes news for market participants [e.g., Mietzner and Schweizer, 2014; Veil, Ruckes, Limbach and Doumet, 2015].

Investors are required to notify whenever their stake, i.e., the number of voting rights they can exercise over the total number of voting rights, reaches or crosses a reporting threshold.[7] In January 2007, the regulator adds to the existing reporting thresholds of 5%, 10%, 25%, 50% and 75% the new thresholds of 3%, 15%, 20% and 30% [sec. 21 WpHG]. Besides, it prescribes analogous rules for financial instruments that grant the right to acquire shares with attached voting rights (e.g., call options with physical settlement) with the exception of the 3% threshold [sec. 25 WpHG in the version applicable until 3/1/2009]. In conjunction with shorter disclosure deadlines and broader dissemination principles [secs. 21, 26 WpHG, sec. 3a WpAIV], the regulatory change hence ensures faster and more granular information about ownership.

Regulatory changes in 2009 and 2012 differ in two respects. First, they are partly motivated by prominently discussed cases in which investors secretly accumulate large positions with instruments outside the scope of existing rules [e.g., European Commission, 2012]. Second, they are more subtle. Since March 2009, voting rights from shares and other financial instruments are to be aggregated, practically resulting in stricter reporting thresholds [sec. 25 WpHG in the version applicable until 11/26/2015]. Since February 2012, financial instruments that merely facilitate the acquisition of shares (e.g., call options with cash settlement) are likewise subject to disclosure. Corresponding voting rights are to be aggregated with the other positions, again tightening thresholds [secs. 25, 25a WpHG in the version applicable until 11/26/2015]. These later changes represent disclosure increases only for investors who are active in equity and derivative markets. In Germany, derivative markets are in principal open to all investors and even the less sophisticated ones, households, access them [DDV, 2010]. For some of the more sophisticated investors, the regulator limits investments in derivatives; constraints are most binding for insurance companies and pension funds and less so for mutual funds and banks. Eventual usage of derivatives appears to depend on both investment strategy and regulatory constraints.

[6] I redact identifying information on investor and issuer according to the publication guidelines of Deutsche Bundesbank.

[7] Normally, it is the investor's trade, i.e., the change in the number of voting rights the investor can exercise, that triggers the disclosure requirement. Occasionally, it is the change in the issuer's capital structure, i.e., the change in the total number of voting rights. I interpret the disclosure as deliberate choice of the investor as raw data suggest that capital structure changes are not material. For the period 2007–2014, issuers report changes in the total number of voting rights on average only every other year (data are from the BaFin's annual reports, available at: www.bafin.de).

For instance, although permitted to, just a certain fraction of mutual funds invest in derivatives [e.g., Natter, Rohleder, Schulte and Wilkens, 2016; Rohleder, Schulte and Wilkens, 2017].[8]

Over the years, the regulator strengthens efforts to enforce ownership disclosure rules. The exact timing of these efforts is hard to pin down, but they are likely concentrated around the first disclosure increase. For example, the BaFin, the enforcement agency, reports a decreasing number of incorrect and delayed notifications around the implementation of the 2007 changes.[9]

Figure 1 displays the number of notifications for the sample period, the years 2005–2014, in absolute terms (gray bars) and per firm listed on the regulated market (blue line). Notifications refer to all notifications under the prevailing regime; in 2007, they relate to shares, in later years also to other financial instruments.[10] Figure 1 shows that the disclosure increases in 2007 and 2012 are measurable. From 2006 to 2007 (2011 to 2012), notifications per firm rise from 4.1 to 8.8 (7.1 to 9.7), which the BaFin attributes primarily to the introduction of the 3% threshold and the extension to cash-settled derivatives [BaFin, 2008, 2013*a*]. The outcome thus suggests that investors with holdings close enough to reporting thresholds are at least prompted to (re)consider the costs of ownership disclosure.

The regulatory changes do not take place in isolation. Most importantly, in 2007, the regulator implements an entire bundle of measures from the Transparency Directive (TPD), which applies in the whole European Union (EU). The TPD aims at establishing higher transparency by increasing disclosure requirements not only for investors, but also issuers (e.g., extension of periodic reports, interim reporting) [secs. (1), (2) Directive 2004/104/EC]. An increase in issuer disclosure is usually associated with a decrease in information asymmetry [e.g., Verrecchia, 2001]. In this vein, it can affect incentives for disclosure avoidance and investment strategies. I argue that effects from increased issuer disclosure should be of second order and hence less relevant for investors. However, I cannot directly test my argument. The takeaway for the empirical analysis is that the effect in 2007 is to some degree a joint effect.

## 2.2 Expectations

The disclosure of notifications entails costs and benefits, albeit to a different extent for different investors as investors are heterogeneous [e.g., Cronqvist and Fahlenbrach, 2009; He et al., 2004].

[8] The references describe derivative usage of US mutual funds. Anecdotal evidence suggests similar behavior of German mutual funds. Since 2004, they are allowed to invest in derivatives and do so increasingly over time. See, e.g.: www.wiwo.de/archiv/futures-und-optionen-fonds-voller-derivate-seite-2/5543050-2.html.

[9] The BaFin reports the numbers in its annual reports, available at: www.bafin.de.

[10] Figure 1 includes notifications from trades in the respective periods and excludes notifications about existing holdings, which have to be disclosed following the regulatory changes in 2007 (with the exception of the 3% threshold) and 2012 [sec. 41 (4a), (4d) and (4e) WpHG].

They vary, among other things, with respect to sophistication, trading patterns, organizational structure and existing regulatory constraints.

Disclosure costs arise, at first, because of free riding and front running. Investors who buy or sell stock based on private information give away part of their informational advantage if they disclose trades after the fact [e.g., Fishman and Hagerty, 1995; Huddart et al., 2001]. Other market participants can use the disclosures to either anticipate future trades or intervene in ongoing transactions that are split into smaller trades. If they mimic the trading patterns of the disclosing investors, they can harm them by intensifying price impact. Besides, front running generalizes to other motives. If disclosing investors do not trade based on private information but observable liquidity needs, other market participants can harm them similarly [e.g., Agarwal et al., 2013; Brunnermeier and Pedersen, 2005].

The degree to which investors are exposed to free riding and front running partly depends on the nature of the private information they possess. Institutional investors, as opposed to households, are generally regarded as informed [e.g., Jank, 2011; Kick et al., 2014]. For institutional investors with diversified, high-turnover portfolios and small stakes, private information should mainly translate into superior information processing capabilities. Mutual funds, banks and insurance companies and pension funds fall into this category. They trade frequently and face regulatory upper limits in terms of stake size. For instance, mutual fund families are not permitted to have a stake that is larger than 10% in a single issuer. Prior literature supports the notion that the revelation of their holdings is costly. Frank et al. [2004] and Verbeek and Wang [2013] show that merely replicating mutual funds' portfolios from public disclosures is profitable. Notably, these portfolio disclosures are substantially less timely than notifications. For institutional investors with concentrated, low-turnover portfolios and large stakes, private information should also refer to information about their own intentions. Non-financial corporations, often blockholders, trade less frequently and face no regulatory constraints on stake size. Their disclosures contain implicit information on future monitoring and, in the extreme, takeover plans. Prior studies support the rationale that such implicit information is priced. Clifford [2008] and Brav, Jiang, Partnoy and Thomas [2008] show that issuer announcement returns are higher around the release of ownership disclosures for active (13D) than for passive (13G) investments and that returns prevail even if the holdings per se are already public knowledge.

The degree to which investors are exposed to front running depends additionally on the observability of their liquidity demand. Typical targets for front running are investors in financial distress, under pressure to liquidate assets. Observability of financial distress is given if investors regularly report their capital inflows and outflows. All institutional investors follow some form

of (periodic) reporting and thereby ensure a certain level of transparency about their funds. The literature points in particular to (open-end) mutual funds [e.g., Coval and Stafford, 2007; Shive and Yun, 2013]. In Germany, market participants can derive flows into and out of mutual funds primarily from their portfolio disclosures, which the law requires at a seminannual frequency.

The second component of disclosure costs are agency costs which arise in delegated money management. The requirement to disclose trades in a timely manner can incentivize agents to act according to their principal's current information set and not their own private information for which returns materialize only with delay. In this vein, they meet the expectations of the principal, but lower expected payoffs to the same [e.g., Edmans et al., 2016; Prat, 2005]. Although most institutional investors act on behalf of someone else, mutual funds' organizational structure makes the agency relationship especially pronounced. Fund investors closely monitor fund managers and can easily exit. Fund managers, in turn, seem to undertake costly actions to prevent liquidity outflows, such as window dressing of portfolios or the promotion of short-term goals towards their portfolio firms [e.g., Agarwal et al., 2014, 2017].

Finally, there are direct costs. Investors themselves repeatedly highlight the administrative costs associated with lower reporting thresholds.[11] For instance, the aggregation of holdings from different subsidiaries or fund portfolios to comply with attribution rules can necessitate additional layers of reporting among individual entities. Furthermore, regulators acknowledge the bureaucratic burden from unharmonized national laws for internationally active investors [European Commission, 2008, 2012]. Thus, large investors across all categories of institutional investors might be affected.

Disclosure costs can then be offset by corresponding benefits or become unavoidable due to existing regulation. Benefits lie in improved monitoring in agency relationships [e.g., Ge and Zheng, 2006; Holmström, 1999] or quicker return realization in case of other investors mimicking trades [e.g., Frank et al., 2004; Ge and Zheng, 2006]. The latter represents the upside of free riding. If investors disclose after fully implementing their strategy or as explicit part of it, 'copycat' investors can increase returns. Empirical evidence in this respect is largely confined to short sellers, sophisticated and lightly regulated investors [Ljungqvist and Qian, 2016].

Existing regulation plays two roles. First, it might result in (almost) no incremental costs from the regulatory changes. Ownership disclosure rules are in place before thresholds are lowered; starting with 5%, investors have to disclose their holdings. Hence, for many non-financial

[11] As an example, the BVI criticizes direct costs especially with respect to the introduction of the 3% threshold. See their comments at: http://webarchiv.bundestag.de/archive/2009/0626/ausschuesse/a07/anhoerungen/03-3/Stellungnahmen/04-BVI.pdf.

corporations and other blockholders, the regulatory changes imply no new disclosures, their holdings are already public knowledge. In addition, institutional investors disclose their holdings via other channels, at least in an aggregated fashion. For example, mutual funds, the most transparent type, disclose portfolio positions seminannually. On a voluntary basis, some provide information on selected positions more frequently, usually monthly or quarterly (e.g., their 'top 10 holdings').[12] Notifications differ in their timeliness: the information that a holding crosses a threshold is to be disclosed within a few days. Still, if investors do not rely on precise timing, existing transparency requirements can make notifications redundant. Second, regulatory constraints might prevent investors from acting on their incentives. Insurance companies and pension funds face strict prudent-man principles, but banks are also under substantial regulatory scrutiny [e.g., Del Guercio, 1996; Jank, 2011].

The above discussion demonstrates that it is not obvious which investors face the highest (net) costs upon the introduction of lower reporting thresholds and have the necessary leeway to act on them. Yet, given the heterogeneity across investors, they are most certainly exposed to *different* costs. This heterogeneity is crucial for observing a response at all. If all investors encountered the same costs and subsequently reduced their holdings, price pressure would, in all likelihood, make the adjustment prohibitively costly. On the other hand, if investors encounter different costs, they will respond according to their *relative* costs of disclosure. Investors will reduce their holdings below the threshold as long as their incentives, including the associated price impact, make it worthwile. The remaining investors will take the other side of the trades.[13]

# 3 Research Design and Data

## 3.1 Identification Strategy

I use the regulatory changes in 2007, 2009 and 2012 as treatment events to measure whether investors adjust their holdings in response to lower reporting thresholds. I estimate effects with respect to the effective dates of the written rules, which do not come as a surprise. Investors should not react (long) before, as there are no benefits to anticipatory behavior. A time lag rather supports the rationale of a strategic response, since it allows for the careful assessment of costs and benefits. However, investors might react with some delay after the effective dates awaiting the enforcement of the rules [Bhattacharya and Daouk, 2002]. In favor of this notion,

[12] For an example, see: https://www.allianzglobalinvestors.de/web/b2cdetails?action_id=b2c.FondsDetails.Struktur&l_act_id=b2c.FondsDetails&1180=DE00084 71004.

[13] The discussion further demonstrates that households should face neither costs nor benefits from disclosure. Consequently, they are a candidate for taking the other side of the trades.

enforcement becomes more stringent over time. Further, Christensen, Hail and Leuz [2016] document a gradual effect of the TPD, the directive that includes the ownership disclosure rules in 2007, on liquidity.

In the holdings analysis, I use the segmentation of the German stock market for identification. The regulatory changes affect only holdings in firms listed on the regulated market. Germany has a relatively large unregulated market, governed by stock exchange rules with considerably lower transparency requirements. In particular, holdings in firms in the unregulated market are at no time subject to ownership disclosure rules [Stehle and Schmidt, 2015]. I conduct a difference-in-differences analysis and measure the impact of lower reporting thresholds on holdings in firms in the regulated market (treatment group) relative to holdings in firms in the unregulated market (control group). Firms listed on the unregulated market are on average smaller and younger than firms listed on the regulated market. Besides, they often have the option to list on both segments and make a deliberate choice for the market with lower transparency standards [Vismara et al., 2012]. To address potential differences which might counteract parallel trends, I work with firm (and quarter-year) fixed effects in almost all specifications. In supplemental tests, I add firm-specific linear time trends. While the latter allows for differential trends not only of treatment and control group but individual firms, it is quite demanding. If the impact is of a gradual nature, a linear trend may well capture part of it. To address segment choice and ultimately treatment assignment, I match firms from the regulated market to firms from the unregulated market based on observable, pre-treatment characteristics.

For causal identification, the disclosure increase in 2007 is the strongest and cleanest event. It is strongest because it introduces the 3% threshold for voting rights from shares. It is cleanest because it originates from EU legislation that is passed years in advance, in 2004. The later two disclosure increases are weaker, since they lower thresholds only for investors who invest in both equity and derivative instruments. Moreover, they raise the issue of reverse causality, since they are designed around publicly discussed cases of disclosure avoidance.

In the notifications analysis, I have, by definition, firms that are treated and selected. Investors are required to disclose notifications solely for firms listed on the regulated market and reportable ownership changes do not occur regularly. Instead of exploiting variation across firms, I make use of variation across investor types, mutual funds and all other investors.

## 3.2 Sample

I compile a sample of firms listed on German stock exchanges in the years 2005–2014 and around at least one of the regulatory changes. To do so, I start with the list of all securities contained in

the SHS database of Deutsche Bundesbank. The SHS keeps records for all securities kept in safe custody in Germany, regardless of issuer or holder origin [Bade, Flory and Schönberg, 2016]. It should thus provide the universe of shares listed on the regulated and unregulated market. I then restrict the sample to equities for which I find information in Thomson Reuters Datastream.[14] I drop securities which appear only before March 31, 2007 (Datastream mnemonic: TIME), the quarter-end of the first treatment event, or only after December 31, 2011 (BDATE), the quarter-end before the last treatment event. I also drop securities with a lifetime of less than 365 days or names indicating that they are not common equity (NAME) [Ince and Porter, 2006]. Lastly, I limit the dataset to one security per firm. I keep only major securities, i.e., the most liquid security for a firm with several securities (MAJOR), and primary listings, i.e., the main listing of the major security for a cross-listed firm (ISINID). The resulting sample consists of 1,844 unique firms.

For the holdings analysis, I partition the sample into treatment and control group. I assign a firm to the treatment group if (i) it is listed on the regulated market for at least eight consecutive quarters according to the register maintained by the European Securities and Markets Authority (ESMA) or (ii) it is included in the CDAX in the years 1999–2013; the CDAX is a German stock market index comprising firms in the regulated market.[15] I assign treatment at the firm level, since the CDAX source does not offer time-variant information and the ESMA register does not start before the second quarter of 2007, after the first disclosure increase. I consider the time-invariant treatment definition sufficient, since switching between markets is generally a rare event.[16] I obtain a treatment (control) group with 867 (977) firms.

Requiring holdings data for the main model results in a final sample of 1,549 unique firms with 814 (735) firms in the treatment (control) group. Table 1 displays the number of firms in the last quarter of each sample year, for the whole sample (column 1) and separately for treatment (column 2) and control group (column 3). The table further specifies the percentage of firms listed on the regulated market that I include in the treatment group (column 4). Its

[14] I use ISIN to link the SHS with Datastream. While the SHS contains historical ISIN, Datastream relies on current ISIN. I use several sources to link historical and current ISIN: (i) snapshots of the Datastream universe for the years 2008, 2009, 2010, 2011, 2012, 2013 and 2016; (ii) a list of ISIN changes on the regulated market in Frankfurt for the years 2005–2014, provided by Martin Schmidt and Richard Stehle from Humboldt University of Berlin; and (iii) a translation file for historical and current ISIN of CDAX firms for the years 1999–2013, provided by Erik Theissen, Esad Smajlbegovic and Thomas Johann from the University of Mannheim. After matching the holdings from the SHS to Datastream, I work with the Datastream code (Datastream mnemonic: DSCD) as security identifier.

[15] For details on the ESMA register, see: www.registers.esma.europa.eu. For details on the CDAX source, see Footnote 14.

[16] I rerun all analyses with a modified, time-variant treatment definition. For all quarters starting in or after $2007q2$, I classify a firm-quarter as treated if it is listed in the ESMA register. For all quarters preceding $2007q2$, I use the firm's treatment status in $2007q2$. Results are overall similar.

increase over the years suggests better data coverage while its decrease in the end reflects sample selection criteria; I exclude firms with first-time appearance after 2011. In both treatment and control group, numbers first increase and then decrease. However, changes are more pronounced for the control group, in line with structural breaks in the unregulated market. The introduction of new subsegments in 2005 and 2008, the Entry Standard and the First Quotation Board, leads to more listing options; the closing of the First Quotation Board in 2012 reduces them again.

## 3.3 Data

### 3.3.1 Holdings

I use holdings from the SHS of Deutsche Bundesbank, the German central bank. Since 2005, Deutsche Bundesbank collects quarterly information on security holdings from all banks and investment companies in Germany (reporting institutions).[17] The reporting institutions are required to inform about their own holdings as well as their customers' holdings, i.e., holdings of individuals and entities that have securities deposits with them. For each equity security, they provide information on the number of shares held at quarter-end. To prevent double entries, they are to include only securities which they keep in safe custody for end consumers. To facilitate further analysis, they are to split up end consumers into investor types. By construction, security holders in the SHS are subject to ownership disclosure rules, but they are not necessarily the only ones. The SHS does not cover individuals or entities that are subject to disclosure requirements via rules on attribution [Amann, Baltzer and Schrape, 2012; Bade et al., 2016].

I work with the following investor types. I first define foreign investors ($foreign$) based on (own) holdings of foreign banks and foreign central securities depositories. Since foreign financial institutions are not required to report to Deutsche Bundesbank, I cannot determine the ultimate holder. Prior studies treat them as predominantly institutional investors [Baltzer et al., 2015; Deutsche Bundesbank, 2015]. Among domestic institutional investors, I distinguish between mutual funds ($funds$), banks ($banks$), insurance companies and pension funds ($ins$), non-financial corporations ($corp$), governmental institutions ($gov$) and other financial investors ($other$).[18] Two categories deserve further comment. Mutual funds include hedge funds. I label the category mutual funds because hedge funds account for a tiny percentage of total funds; on average 0.007% for the years 2007–2012.[19] Other financial investors encompass heterogeneous

[17] Deutsche Bundesbank uses the data for regulatory and research purposes. Researchers can apply for access and work with the data on-site.

[18] Strictly speaking, domestic investors include some foreign investors, but their holdings are negligible. For instance, domestic mutual funds have an average stake of 1.262% whereas foreign mutual funds have an average stake of 0.019% in the main sample with 38,330 observations.

[19] The numbers of funds are from the BaFin's annual reports, available at: www.bafin.de.

entities, such as investment companies with their proprietary holdings or bank-type institutions pursuing risky business. I group them together to form a residual category, not because I expect them to behave alike. Lastly, I define domestic households ($hh$) in accordance with the SHS.[20]

In the analysis, I measure the effect on the percentage stake as it determines whether a reporting threshold is crossed. I compute stakes at quarter-end, from the last quarter of 2005, the first available quarter, until the last quarter of 2014. The stake of investor type $\star$ in firm $i$ and quarter $t$, $stake\_\star_{it}$, corresponds to the number of shares held by investor type $\star$ divided by the number of shares outstanding as provided by Datastream (NOSH) and multiplied by 100. I use shares outstanding from Datastream because the SHS does not record foreign holdings for which shares are kept in custody outside of Germany. I also make use of these foreign holdings to refine the sample. In Germany, foreign investors have a clear preference for the largest, index-listed firms [Deutsche Bundesbank, 2014]. At the same time, these firms lack comparable firms in the control group. I hence exclude them by truncating the difference between shares outstanding and total shares reported in the SHS (as percentage of shares outstanding) at the 95$^{\text{th}}$ percentile. Lastly, I clean up the data. I truncate the percentage difference between shares outstanding and total SHS shares at the 5$^{\text{th}}$ percentile to remove negative cases which are obvious errors. I assign the remaining difference to the foreign investors already classified as such. I truncate stakes by type and quarter at the 1$^{\text{st}}$ and 99$^{\text{th}}$ percentile to remove more subtle inconsistencies at the level of the investor type. The final sample contains only firm-quarters with nonmissing $stake\_\star_{it}$ for all investor types.[21]

### 3.3.2 Notifications

I obtain notifications for the period January 20, 2007, to November 30, 2012, from the DGAP, the main service provider for the dissemination of regulated information in Germany. The DGAP distributes a substantial portion of all notifications; the average over the sample years amounts to 62%.[22] Firms in the dataset tend to be larger firms; their average market value exceeds the average market value in the treatment group of the holdings sample. According to informal statements of the DGAP, the client structure remains stable over time.

[20] The SHS classifies security holders according to the European System of Accounts (ESA). I use the following ESA 2010 codes for the investor types. $foreign$: 1225, 1228; $funds$: 1230, 1240; $banks$: 1221, 1222, 1223, 1224; $ins$: 1280, 1290; $corp$: 1100, 1270; $gov$: 1311, 1312, 1313, 1314; $other$: 1251, 1261, 1262; and $hh$: 1400, 1500.

[21] To reduce discretion in data cleaning, I rerun all models with a different version of $stake\_\star_{it}$, the raw stakes truncated at the 1$^{\text{st}}$ and 99$^{\text{th}}$ percentile across all types and the whole sample period. Results are very similar.

[22] The number constitutes a lower bound. I derive it based on the total number of notifications collected by the DGAP and the total number of notifications reported by the BaFin. However, the DGAP and the BaFin define notifications differently. Precisely, one notification of the DGAP often contains several notifications as understood by the BaFin.

The notifications comprise machine-readable, largely unstructured text. The regulator does not prescribe a uniform reporting format which results in some variety in practice; Appendix A.2 presents one example. I use regular expression matching to extract the following items from each notification: issuer ISIN, disclosure and trade date, stake, reporting threshold(s) crossed as well as investor type. For the investor type, I distinguish between mutual funds ($mf$) and other investors. In particular, I search for name matches with the member list of the BVI, the German Investment Funds Association.[23] Hence, I focus on German mutual funds to be consistent with the holdings analysis. At the same time, BVI members are often internationally active, putting more emphasis on the investor type per se. The member list contains 101 different names. The notifications sample for the 3% threshold matches with 33 names, of which a small number clearly dominates (percentage among matches in parentheses): BlackRock (34%), Allianz Global Investors (15%), Universal-Investment (9%), Deka (6%) and Union Investment (5%). Findings are thus driven by a small number of fund families.

In the analysis, I work with the data in two formats. I first examine the subset of notifications for falling below or exceeding the 3% threshold. From 28,270 raw notifications, I drop notifications because they are duplicates due to translations or rules on attribution; they have missing information on the variables of interest; or they relate to other sections of the WpHG. From the 9,082 remaining notifications, 2,596 (2,066) indicate falling below (exceeding) 3% for 375 firms. I then return to the notifications for all reporting thresholds and aggregate them at the firm-quarter level. I investigate the share of notifications for falling below a threshold in mutual funds' notifications in a given firm-quarter; the number of firms declines to 204 as not all firms have notifications submitted by mutual funds. In addition, I investigate mutual funds' share in overall notifications in a given firm-quarter; the number of firms rises to 535.

# 4 Results

## 4.1 Descriptives

Table 2 displays summary statistics.[24] Panel A shows the distributions of stakes, $stake\_\star_{it}$, for the whole sample (line 1) and separately for treatment ($t$, line 2) and control group ($c$, line 3). The columns on the left (right) inform about (the top decile of) equal-weighted percentage stakes and the middle columns about value-weighted stakes, $stake\_\star_{it}$ weighted by market

[23] I collect the names from the member list in the BVI yearbook 2016 (www.bvi.de) and the website of the BVI (www.bvi.de).

[24] I limit summary statistics to mean and standard deviation to conform to the publication guidelines of Deutsche Bundesbank.

value. The sample for value-weighted stakes is slightly smaller, since market value, $mv_{it}$, is not always available. The low number of observations for $stake\_gov_{it}$ in the top decile of the control group reflects a low number of non-zero holdings of governmental institutions in the unregulated market.

Percentage stakes demonstrate particularities of the German capital market. Foreign investors, non-financial corporations and households dominate overall ownership; $stake\_foreign_{it}$, $stake\_corp_{it}$ and $stake\_hh_{it}$ together account for more than 90% with roughly equal proportions. Non-financial corporations' large average holdings represent family holdings and cross-holdings that originate from the traditionally insider-oriented German system [e.g., Baltzer et al., 2015; Franks and Mayer, 2001]. Percentage stakes also illustrate distinct ownership patterns in treatment and control group, which are in part attributable to the investment decision per se. Investors with generally small holdings ($funds$, $banks$, $ins$ and $gov$) have more stakes equal to zero in the unregulated than the regulated market.

Consistent with prior literature, the comparison of equal-weighted with value-weighted stakes depicts a preference of institutional investors for large firms [e.g., Dahlquist and Robertsson, 2001; Ferreira and Matos, 2008]. Due to my sample selection criteria, values differ to some extent from other studies based on the SHS [e.g., Baltzer et al., 2015; Deutsche Bundesbank, 2014]. I eliminate the preferred investment target of mutual funds, banks and insurance companies and pension funds: large and index-listed firms. At the same time, this approach yields treated and control firms that are closer in size. In Panel B, the average market value, $mv_{i(t-1)}$, amounts to 348m € (35m €) in the treated (control) sample.

Finally, average stakes in the top decile of the stake distributions reveal whether reporting thresholds can translate into binding constraints for the different investor types. Among the investors with small holdings, mutual funds and banks reach values well beyond 3%, the newly introduced threshold in 2007; $stake\_funds_{it}$ ($stake\_banks_{it}$) has a mean of almost 10% (5%) in the top decile of the treatment group, suggesting that, also at the individual level, some investors encounter more extensive disclosure requirements. Insurance companies and pension funds and governmental institutions seem to be less affected; $stake\_ins_{it}$ ($stake\_gov_{it}$) has a mean of around 2% (0.2%) in the top decile of the treatment group. Means refer to the whole sample period and therefore include potentially adjusted post-treatment values. In addition, maxima for $stake\_\star_{it}$ in the treatment group clearly exceed 10% for every investor type (untabulated). Still, the low values for insurance companies and pension funds and governmental institutions have implications for the subsequent analysis. If I find evidence consistent with an adjustment of holdings, only very few observations can be responsible for it.

Panel B of Table 2 contains firm characteristics, which are informative about the type of firm that selects into the (un)regulated market. Firms in the regulated market have higher share prices, $up_{i(t-1)}$ has a mean of €26.557 (€17.842) in the treated (control) sample; they have a larger analyst following, $ana_{it}$ amounts on average to 3.317 (0.237) analysts; they are older, $age_{it}$ has a mean of 57.251 (23.938) quarters; and they pay more dividends, $dy_{i(t-1)}$ indicates that dividends amount on average to 1.2% (0.5%) of the share price. Around 83% of treated firms are listed in the CDAX, a stock market index ($cdax_i$). Control firms cannot, by definition, be listed in the CDAX. I provide precise definitions of the variables in Appendix A.1.

### 4.2 Main Analysis

To assess the impact of stricter reporting thresholds on investors' holdings, I estimate the following model (main model):

$$stake\_\star_{it} = \beta_1 post7_t \times treat_i + \beta_2 post9_t \times treat_i + \beta_3 post12_t \times treat_i + \beta_k FE_t + \beta_l FE_i + \epsilon_{it}$$

where $stake\_\star_{it}$ is the percentage stake in firm $i$ and quarter $t$ of foreign investors ($foreign$), mutual funds ($funds$), banks ($banks$), insurance companies and pension funds ($ins$), non-financial corporations ($corp$), governmental institutions ($gov$), other financial investors ($other$) or households ($hh$). $post7_t$ (9) [12] is an indicator variable equal to one if quarter $t$ is larger than $2006q4$ ($2008q4$) [$2011q4$]. $treat_i$ is an indicator variable equal to one if firm $i$ is part of the treatment group, i.e., listed on the regulated market. $FE_t$ and $FE_i$ denote quarter-year and firm fixed effects and subsume the main effects of the $post$ indicators and $treat_i$. Standard errors are two-way clustered by firm and quarter-year to control for both cross-sectional and time-series dependence [Gow, Ormazabal and Taylor, 2010; Petersen, 2009].[25] The coefficients of interest, $\beta_1$, $\beta_2$ and $\beta_3$, capture differential changes in holdings in the treatment group relative to the control group upon the introduction of lower reporting thresholds.

Table 3 shows the results. Mutual funds, banks, insurance companies and pension funds and households reduce their holdings, albeit at different points in time. Mutual funds are the sole investor type with a statistically significant estimate for 2007, the cleanest event for identification. The coefficient indicates a decline in stakes of 38bp or 17% of their pre-treatment level, translating into asset sales of €2,074m or 1% of total market capitalization.[26] Given

[25] As an alternative, I cluster only by firm. Standard errors vary somewhat across specifications, but two-way clustering seems to be the overall more conservative option.

[26] The mean of $stake\_funds_{it}$ in the sample of treated firms in the pre-treatment period is 2.268; the estimate for 2007 ($\beta_1$) is -0.378; and -0.378/2.268 corresponds to a decline of 17%. I multiply the percentage reduction (17%) with the total market value of mututal funds' holdings in treated firms to obtain the monetary value

that the estimate is likely driven by a subset of mutual funds, those that belong to the large fund families, it is economically significant. Moreover, it should constitute a lower bound. In the data, I cannot identify holdings of foreign mutual funds and I do not consider holdings in mutual funds' preferred investment target, the largest, index-listed firms. Mutual funds continue to reduce their holdings in 2009; the coefficient indicates an incremental decline of 29bp. They do not seem to respond to the disclosure increase in 2012; the coefficient is not different from zero. Banks and insurance companies and pension funds reduce their holdings, but the reduction is less clearly linked to the introduction of lower reporting thresholds. Despite altogether negative estimates, only the coefficient(s) for 2009 (and 2012) is (are) statistically significant (in the case of insurance companies and pension funds). As discussed in the previous section, the result for insurance companies and pension funds probably stems from very few observations. Anticipating the results in Table 4, it is also not robust across specifications. Lastly, households reduce their holdings. The finding is limited to the disclosure increase in 2012 and, as discussed below, not robust.

On the other side of the market, non-financial corporations increase their holdings. Estimates are positive throughout and statistically significant for the later two disclosure increases. Estimates for the remaining investor types are not different from zero. Thus, results are consistent with a response to the introduction of lower reporting thresholds. While mutual funds and banks decrease, non-financial corporations increase their holdings. Coefficients for mutual funds and banks do not differ, supporting similar adjustments. However, they differ from the coefficients for non-financial corporations, strengthening the observed, opposite patterns in responses.[27]

I then vary sample and regression model in a number of ways. I first repeat the estimation of the main model with a balanced panel to account for the changing sample composition. For a firm to be in the sample, I require complete quarterly observations for the years 2005–2008. I do not extend the requirement until the end of the sample period to have a reasonable number of control firms for estimation. The resulting sample contains 375 (75) treated (control) firms and 15,117 firm-quarters. By construction, the number of quarterly observations declines towards the end of the sample period.

Second, I add firm-specific linear time trends to the main model to allow for differential trends of individual firms. For a firm to be in the sample, I require observations for the third

(€2,074m). For the total market value of holdings, I use the average of the quarterly values in the pre-treatment period.

[27] Precisely, the $\beta_2$ and $\beta_3$ coefficients of mutual funds and banks differ from the $\beta_2$ and $\beta_3$ coefficients of non-financial corporations at the 5% significance level.

and fourth quarter of 2006, the quarters preceding the first disclosure increase. The resulting sample encompasses 588 (175) treated (control) firms and 23,722 firm-quarters.

Third, I relax the fixed effects structure and include time-varying control variables for investment preferences of investors [e.g., Baltzer et al., 2015; Dahlquist and Robertsson, 2001; Ferreira and Matos, 2008]. The set of control variables, $CONT$, comprises size $[ln(mv)_{i(t-1)}]$, share price $[ln(up)_{i(t-1)}]$, analyst following $[ln(ana)_{it}]$, age $[ln(age)_{it}]$, dividend yield $[dy_{i(t-1)}]$ and index membership $[cdax_i]$. Lagged variables allow for time to respond. I account for industry (and quarter-year) fixed effects instead of firm fixed effects. I add the main effect for the treatment group, $treat_i$, as it is not redundant in this specification. I lose some observations due to missing information on the control variables. The sample contains 798 (714) treated (control) firms and 35,499 firm-quarters.

Fourth, I estimate the main model on subsets of matched firms to address the market segment choice. I start with the balanced panel from above and model whether the firm chooses a listing on the regulated market. I rely on accounting and not capital-market variables as predictors, since the latter are partly the outcome of the segment choice. I use the natural logarithm of total assets, return on assets and sales growth, each as average over the pre-treatment period, the years 2005–2006. I apply propensity score matching (PSM) with replacement and obtain a matched sample with 289 (34) treated (control) firms and 18,368 firm-quarters.[28] In addition, I match firms with the coarsened exact matching (CEM) method [Iacus, King and Porro, 2012]. CEM sorts the predictors into strata and produces exact matches based on the strata. The resulting sample encompasses 137 (35) treated (control) firms and 5,821 firm-quarters. Control observations enter with continuous weights to take into account the size of the strata.[29]

Table 4 shows the results. Overall, mutual funds and banks decrease their holdings. The finding for banks is weaker in that it is less pronounced around the disclosure increase in 2007 and more sensitive to model specification. Non-financial corporations increase their holdings, not necessarily around the disclosure increase in 2007, but seemingly in response to the reduction in holdings of mutual funds and banks. Restricting the sample to a balanced panel leads to statistically significant estimates for all three investor types in 2007. The coefficient for mutual

[28] Specifically, I use one-to-one, nearest neighbor matching and set the caliper to 0.2 × the standard deviation of the propensity score from the logit model (0.031). I verify that differences in the predictors between treated and control firms are not statistically significant after matching.

[29] Under both matching methods, I weight and thus reuse control observations. As an alternative, I repeat the matching, but require that control observations enter only once. The resulting sample for PSM (CEM) contains 76 (34) unique firms, with the same number of treated and control firms, and 2,572 (1,156) firm-quarters. Compared to the estimates in Table 4, results are similar for mutual funds, mixed for banks and weaker for non-financial corporations.

funds in 2007 is comparable in magnitude to the main model; the coefficients for 2009 and 2012 indicate that incremental adjustments in the later years reverse. Introducing firm-specific linear time trends supports the reduction in holdings of mutual funds and hints at a reaction of insurance companies and pension funds. Null results for the other investors come as no surprise as the specification is quite demanding. The pre-treatment period is relatively short and investors might react gradually. Accounting for investment preferences provides further evidence for ownership changes of mutual funds and non-financial corporations and illustrates inconsistency for households; coefficient signs change across specifications. Restricting the sample to matched subsets of firms yields estimates roughly in line with the previous models, but larger in magnitude. In untabulated analyses, I conduct remaining robustness tests. I drop non-German firms and penny stocks and repeat the estimation of the main model. I obtain results very similar to Table 3.

So far, I can show that some investors decrease and others increase their holdings, but I cannot show which actions investors take. In principle, the decrease can result from either the complete sale of the investment or the mere reduction in percentage points. My argument predicts the latter: investors reduce their holdings below the new reporting threshold to avoid disclosure. In Table 5, I provide aggregate statistics to give some insight into the specific actions around the disclosure increase in 2007. For each investor type, it depicts the number of firms in which the investor type is invested, i.e, the number of non-zero investments, in absolute terms and as percentage of all firms in the market; and the market value of the investments, in absolute terms and as percentage of total market capitalization. It contains information separately on the regulated (treatment group, $t$) and unregulated market (control group, $c$) as well as the five quarters preceding (pre-period, $PRE$) and the eight quarters succeeding the disclosure increase in 2007 (post-period, $POST$). The table displays averages of the quarterly values.

Table 5 indicates that, in the regulated market, mutual funds and banks slightly increase the number of firms they are invested in while the market value of their holdings declines. If they sold off completely, the number of non-zero investments should drop. Instead, non-zero investments rise. Mutual funds and banks seem to have less money, but more holdings in the regulated market after the disclosure increase, suggesting on average smaller stakes. Furthermore, percentages of overall firms and total market capitalization are roughly the same in both periods, supporting the notion that it is stricter disclosure rules and not revisions in investment strategies that induces the adjustment of holdings.

This interpretation is only one of several. Alternatively, it could be changes in the number of investors that explain the larger number of non-zero investments. At least for mutual funds,

this does not seem likely. Their number is, if at all, decreasing over the years 2005 until 2008.[30] Moreover, it could be (in part certainly is) stock price changes in the context of the financial crisis that explain the decline in the market value of holdings. German mutual funds experience sizeable outflows during the financial crisis [Deutsche Bundesbank, 2013]. Yet, this does not rule out disclosure avoidance behavior. Some investors, in particular non-financial corporations, increase the market value of their holdings, despite a market-wide declining trend in stock prices. Hence, I acknowledge that the evidence is purely descriptive, but I consider it in line with a mere reduction and not the complete sale of holdings.

Table 5 is informative in other respects. First, it demonstrates that investors with generally small holdings ($funds$, $banks$, $ins$ and $gov$) are invested in just a fraction of the market and a meaningfully smaller fraction in the unregulated market. For instance, mutual funds have holdings in 55% (21%) of all firms in the (un)regulated market in the pre-treatment period. Regression results thus hinge on relatively few (control) observations. Second, trends across time resemble each other in the regulated and unregulated market, which alleviates concerns regarding the suitability of the unregulated market as control.[31]

Taken together, the results of the analyses are consistent with *different* adjustment patterns of investors in response to *different* costs of disclosure. In the extreme, they imply that the introduction of lower reporting thresholds can provoke a shift in ownership structure. The findings suggest that mutual funds and banks are on the sell side of the market. Both are sophisticated investors with small stakes and subject to free riding and front running. Banks' more muted reaction supports binding regulatory constraints. Mutual funds' more pronounced reaction supports the idea that the threat of liquidity outflows is strong. Notably, mutual funds' investment behavior is already fairly transparent. It appears to be the particular disclosure instrument and its timeliness that trigger their response. The findings also suggest that non-financial corporations are on the buy side of the market. They usually have stakes above existing reporting thresholds and trade less frequently. Their reaction seems to be less of a response to the disclosure increases, but more to the disclosure avoidance strategies of others. Finally, the analysis highlights the importance of heterogeneity among investors. Heterogeneity is what makes the adjustment of holdings possible: if some investors want to sell, others need to buy.

[30] To be precise, the number of mutual funds amounts to 6,303 in 2005; 5,884 in 2006; 5,995 in 2007; and 6,031 in 2008. Numbers are from the BaFin's annual reports, available at: www.bafin.de. They refer to all mutual funds and do not differentiate between fund categories, such as equity or fixed income funds.

[31] For example, substitution into the unregulated market could be an alternative story. If investors intend to avoid disclosure, they can shift investments from the regulated into the unregulated market. Under this scenario, holdings in the unregulated market would be affected by the regulatory changes and thus no suitable controls.

At the same time, it often conceals the adjustment of holdings: if different investors are grouped together, their reactions can cancel each other out. The lack of any finding for foreign investors might result precisely from such offsetting.

### 4.3 Mutual Funds

The remaining sections focus on mutual funds. So far, I can link their behavior most robustly to the introduction of lower reporting thresholds. In the following, I use additional test designs and data sources to further link this behavior, the reduction in holdings, to the motivation to avoid disclosure.

I first assess when mutual funds reduce their holdings in response to the introduction of the 3% threshold in 2007. In the difference-in-differences analysis, I measure the effect with respect to the effective date of the new threshold. Yet, enforcement becomes stronger over time, which should result in a gradual adjustment of holdings. Even if mutual funds intend to avoid disclosure, they likely wait to observe the outcome of regulatory oversight. Figure 2 depicts the average quarterly level of mutual funds' holdings, $stake\_funds_{it}$, in the period 2005–2008, separately for treatment and control group. Figure 2a shows averages of the raw stakes for a balanced panel. In the post-treatment period, stakes follow a declining trend in the treatment group while they remain at roughly the same level in the control group. Although this pattern is consistent with a gradual impact over time, it seems to start already in the pre-treatment period. In Figure 2b, I apply the synthetic control method [Abadie, Diamond and Hainmueller, 2015; Abadie and Gardeazabal, 2003]. That is, I report averages for stakes in treated firms and synthetic controls. For each treated firm, I construct a synthetic control, i.e., a weighted combination of control firms. I obtain the weights by identifying the combination of stakes in control firms that most closely tracks the treated firm's stake in the pre-treatment period.[32] By construction, stakes in treated and synthetic control observations follow a parallel trend in the pre-treatment period. In contrast to Figure 2a, stakes in treated firms do not decrease before 2007. They rather start declining and deviating from stakes in synthetic controls when the 3% threshold becomes effective and do so increasingly over the subsequent quarters.

I then explore the notifications data, the subset of publicly observable holdings with information at the level of the individual investor. I start with the distribution of reported stakes around the 3% threshold. The preceding results are in line with mutual funds reducing their holdings

[32]Specifically, I first restrict the sample to a balanced panel and the treatment group to firms with stakes less than or equal to the maximum stake in the control group in the same quarter. I then match each treated firm to a combination of control firms whose combined stake most closely resembles the treated firm's stake in the pre-treatment period.

below the initial reporting threshold and this in part after its introduction. Thus, there should be corresponding notifications and more of them for mutual funds than for other investors. Figure 3 displays the distribution of stakes after crossing the 3% threshold, separately for falling below and exceeding the threshold as well as mutual funds and other investors. Figure 3a and Figure 3c illustrate that mutual funds more frequently report stakes just below the threshold. For them, more than 40% of all stakes fall into the interval next to the threshold while, for other investors, it is around 25%. Similarly, other investors more frequently completely sell off their holdings. For them, almost 30% of all stakes fall into the interval around zero while, for mutual funds, it is less than 10%. Hence, notifications support the disclosure avoidance rationale: compared to other investors, mutual funds accumulate stakes just below the threshold. Figure 3b and Figure 3d show the distribution of stakes after exceeding 3% and demonstrate that, again, mutual funds tend to report stakes closer to the threshold. The difference is less pronounced, but the motivation is also less obvious. It might be the incentive to minimize the information content in the notification. If investors disclose stakes around the threshold, they do not convey much information apart from the threshold that is crossed and has to be disclosed in any case.

Panel A of Table 6 translates the above into a regression framework. The dependent variable is the reported stake, $stake_j$, from notification $j$. The independent variables comprise indicator variables for mutual funds, $mf_j$, and the disclosure increases in 2009 and 2012; $post9_t$ (12) is equal to one if trading day $t$ is after February 28, 2009 (January 31, 2012). Note that I use the exact dates and not only the quarters in which the rules become effective. Of main interest is the coefficient for $mf_j$. If mutual funds report stakes closer to 3%, it will be positive for falling below (columns 1 and 2) and negative for exceeding the threshold (columns 3 and 4). I interact $mf_j$ with the *post* indicators to test for changing behavior across time. I present models with two different fixed effects structures, only firm fixed effects (columns 1 and 3) and fixed effects for firm, quarter-year, month and weekday (columns 2 and 4). Standard errors are two-way clustered by firm and quarter-year.[33] Results strengthen the implications of Figure 3. Mutual funds report stakes closer to 3% and the difference is especially large for falling below the threshold. The coefficient of $mf_j$ equals 0.538 in column 2 and -0.137 in column 4; all coefficients are statistically significant at the 1% level. Coefficients of the interaction terms are not different from zero, indicating no changing behavior over time *conditional* on reporting stakes below or above 3%. I relax this restriction in the following.

If mutual funds reduce their holdings below the threshold in particular around its introduc-

[33] Clustering only by firm changes inferences in column 3; the coefficient of $mj_j$ is statistically significant at the 5% instead of the 1% level.

tion, they will disclose relatively more notifications for falling below in the earlier years. The notion is consistent with an adjustment of holdings to avoid ownership disclosure and, at the same time, a delayed response awaiting the enforcement of the rules. I test it in columns 1 and 2 in Panel B of Table 6. The dependent variable is the share of notifications for falling below a threshold in mutual funds' notifications in a given firm-quarter, $share\_mf\_below_{it}$. Over all years, mutual funds submit notifications for falling below and exceeding in roughly equal proportions; the average of $share\_mf\_below_{it}$ amounts to 0.485. To gain insight into the distribution across time, I divide the sample period into subperiods based on the regulatory changes. I use the *post* indicators for the disclosure increases in 2009 and 2012 and define them again on quarters; $post9_t$ (12) is equal to one if quarter $t$ is larger than $2008q4$ ($2011q4$). I include industry fixed effects in both specifications and the set of control variables $CONT$ in column 2; specifically, I control for size $[ln(mv)_{i(t-1)}]$, share price $[ln(up)_{i(t-1)}]$, analyst following $[ln(ana)_{it}]$, age $[ln(age)_{it}]$, dividend yield $[dy_{i(t-1)}]$ and index membership $[cdax_i]$. Standard errors are two-way clustered by firm and quarter-year.[34] In line with the disclosure avoidance rationale, results show that mutual funds disclose relatively less notifications for falling below in the later subperiods. For instance, the coefficient of $post9_t$ equals -0.089 in column 2, which corresponds to a decline of almost 20% of the dependent variable at its mean.

Finally, I test whether mutual funds' share in overall notifications changes with the disclosure increases in 2009 and 2012. An increase (a decrease) suggests that mutual funds are more (less) affected by the later two regulatory changes, which represent stricter rules only for investors who are active in both equity and derivative markets. In Section 2.1, I argue that mutual funds should, in principle, face more stringent disclosure requirements, since they have the necessary leeway to invest in derivatives (and do so). In this section, I examine whether they exploit the leeway in a manner that leads to more disclosure, compared to all other investors. I report the results in columns 3 and 4 in Panel B of Table 6. The regression models mirror the specifications in columns 1 and 2 except for the dependent variable which is the number of notifications submitted by mutual funds over the total number of notifications disclosed in a given firm-quarter, $share\_mf_{it}$. Over all years, mutual funds submit around one tenth of notifications; the average of $share\_mf_{it}$ amounts to 0.118. Upon the introduction of stricter reporting thresholds, they appear to be disproportionately affected. The results indicate that their share in overall notifications increases significantly over time.

[34] Clustering only by firm results in non-significant coefficients of $post12_t$.

# 5 Liquidity Effects

In this section, I study liquidity consequences of the disclosure increases. I investigate whether the reduction in mutual funds' holdings mitigates the positive relation between disclosure and stock liquidity documented in prior literature. Christensen et al. [2016] show that the implementation of the TPD leads to an increase in liquidity in the EU. They assess the effect, among other things, by comparing firms in the regulated and unregulated market. Agarwal et al. [2015] find that mandatory, more frequent mutual fund disclosures lead to higher liquidity. Both studies argue that more extensive disclosure reduces information asymmetry among market participants and thus increases liquidity. In this vein, both studies shift the focus from disclosing to non-disclosing investors. In Section 4, I concentrate on disclosing investors and their holding adjustments in response to disclosure increases. The results indicate that mutual funds reduce while non-financial corporations add to their holdings. These changes in ownership, however, can have negative effects on liquidity. A smaller fraction of mutual fund investors can lower liquidity because it implies a smaller fraction of investors who frequently trade and, thereby, keep trading costs low [e.g., Demsetz, 1968; Rubin, 2007]. A larger fraction of non-financial corporations can lower liquidity because it implies a larger fraction of informed blockholders [e.g., Heflin and Shaw, 2000; Rubin, 2007].

Empirically, I first provide evidence for the above arguments on a standalone basis in my setting. I start with the ownership structure channel and test whether mutual fund ownership is positively associated with liquidity. I use the notifications for crossing the 3% threshold to identify changes in ownership. I examine whether the change in bid-ask spreads, my proxy for liquidity, differs for changes in mutual fund ownership relative to changes in other investors' ownership. I use other investors as benchmark to abstract from disclosure effects, i.e., confounders primarily driven by the disclosure of notifications and not the underlying changes in ownership. If there is a positive relation between mutual fund ownership and liquidity, the change in bid-ask spreads will be more positive (negative) for decreases (increases) in mutual fund ownership. Panel A of Table 7 displays the results, in column 1 (2) for the sample based on notifications for falling below (exceeding) the threshold, i.e., ownership decreases (increases). The dependent variable is the change of the bid-ask spread from the pre-trade to the post-disclosure period, $\Delta ln(spread)_{it}$; spreads are averages of daily values over three months. The independent variables encompass the indicator variable for mutual funds, $mf_j$, the variable of interest, as well as control variables from prior literature, size [$\Delta ln(mv)_{it}$], turnover [$\Delta ln(to)_{it}$] and return variability [$\Delta ln(sd\_ret)_{it}$] [e.g., Christensen et al., 2016; Leuz and Verrecchia, 2000]. Control variables

enter in changes, derived in a similar fashion as the change of the bid-ask spread.[35] I include industry fixed effects and cluster standard errors by firm. Results suggest a somewhat asymmetric relation between mutual fund ownership and liquidity. Bid-ask spreads increase relatively more if mutual fund ownership decreases, but they do not change differentially if it increases.

I return to the holdings data to have comprehensive information on mutual fund ownership that does not condition on crossing a reporting threshold. The sample includes firm-quarters for the regulated and unregulated market and excludes penny stocks.[36] Following a similar rationale as above, I test whether bid-ask spreads, $ln(spread)_{it}$, are negatively associated with changes in mutual fund ownership, $\Delta stake\_funds_{it}$. I account for size $[ln(mv)_{i(t-4)}]$, turnover $[ln(to)_{i(t-4)}]$ and return variability $[ln(sd\_ret)_{i(t-4)}]$ as well as firm and quarter-year fixed effects.[37] I cluster standard errors by firm and quarter-year. Column 1 in Panel B of Table 7 depicts the results. Spreads are negatively related to changes in mutual fund ownership and the relation is statistically significant at the 1% level. In column 2, I repeat the estimation with a combined ownership category of sophisticated, frequently trading investors (*combined*); mutual funds, banks and insurance companies and pension funds. The coefficient of $\Delta stake\_combined_{it}$ is smaller in magnitude and not statistically significant, suggesting that mutual funds are particularly important as liquidity providers.

In the next step, I address the information asymmetry channel. I analyze liquidity around the introduction of stricter ownership disclosure rules, essentially following the research design of Christensen et al. [2016]. I estimate liquidity responses for firms in the regulated market, the treatment group, relative to firms in the unregulated market, the control group. I include firm-quarters from the holdings analysis with available data for estimation in the sample. Column 1 of Table 8 displays the results. The dependent variable is the bid-ask spread, $ln(spread)_{it}$. The variables measuring the effect of interest comprise the interactions of the indicator variable for the treatment group, $treat_i$, with the *post* indicators; $post7_t$ (9) [12] is equal to one if quarter $t$

[35] For spreads, size and turnover, I compute averages of daily values for the 63 trading days before the trading date of the notification and the 63 trading days after the disclosure date of the notification. For return variability, I compute the standard deviation of daily returns over the same periods. I require a minimum of 33 observations in each period. For all variables, I use the natural logarithm of one plus the variable and calculate the change as the difference between post-disclosure and pre-trade value.

[36] I define a stock as penny stock if its unadjusted price is less than the 5$^{th}$ percentile of the price distribution [Schmidt, von Arx, Schrimpf, Wagner and Ziegler, 2017]. Using the more common cutoff of €1 leads to a loss of observations especially in the unregulated market and likely does not address particularities of the German setting; even among the largest, index-listed firms, prices below €1 are observable [Stehle and Schmidt, 2015]. However, my penny stock definition results in some outliers driving the distributions of the (liquidity) proxies. For instance, the mean of $spread_{it}$ equals 6.2% (Panel B of Table 2).

[37] I define variables on calendar quarters and proceed similarly as above. Specifically, spread and turnover are quarterly averages of the daily values. Return variability is the quarterly standard deviation of daily returns. I require at least 33 (10) observations (for turnover). Size is the market value at the end of the quarter. I lag each control variable four quarters and use the natural logarithm of one plus the variable.

is larger than 2006$q4$ (2008$q4$) [2011$q4$]. I control for size [$ln(mv)_{i(t-4)}$], turnover [$ln(to)_{i(t-4)}$] and return variability [$ln(sd\_ret)_{i(t-4)}$], together subsumed by $CONT$ (see Footnote 37 for the construction of the variables). I account for firm and quarter-year fixed effects and cluster standard errors by firm and quarter-year. Results support a positive relation between disclosure and liquidity. The introduction of stricter reporting thresholds in 2007 and 2009 is associated with lower bid-ask spreads. The further tightening of the disclosure regime in 2012 does not have an incremental impact.

In the final step, I investigate how the ownership structure and information asymmetry channel interact. So far, the evidence is in line with the existence of each in isolation. Since I observe a reduction in mutual funds' holdings around the disclosure increases, I now assess whether they counteract each other when considered jointly. To do so, I extend the model in column 1 of Table 8. Precisely, I add interactions with an indicator variable for high mutual fund ownership in the pre-treatment period, $high_i$. I use pre-treatment values to provide for some degree of exogeneity in the proxy. I single out firms with a high level of mutual fund ownership because I expect them to experience the largest reduction. The variable $high_i$ is equal to one if $stake\_funds_{it}$ is larger than the median in the last quarter of 2006, where the median is defined separately for treatment and control group to take into account generally different ownership levels. I work with two samples: a small sample in which each firm has an observation for the last quarter of 2006 and a broad sample in which I set $high_i$ equal to zero if the firm has no observation. Columns 2 and 3 of Table 8 report the results for the broad and small sample, respectively. Across the disclosure increases in 2007 and 2009 and both samples, there is evidence for an overall offsetting effect. While the treatment effect for firms with low mutual fund ownership is more negative, it is close to zero for firms with high mutual fund ownership. For instance, in column 3, the coefficients of $post7_t \times treat_i$ and $post7_t \times treat_i \times high_i$ are -0.032 and 0.024, respectively. Thus, results are consistent with the reduction in mutual funds' holdings having a negative impact on liquidity and weakening the positive impact through the information asymmetry channel. Lastly, I repeat the estimation for the combined stake of mutual funds, banks and insurance companies and pension funds, *combined*, and report the results in columns 4 and 5 of Table 8. Again, the finding for mutual funds does not generalize.

# 6 Conclusion

In this study, I investigate which investors adjust their holdings around the introduction of lower reporting thresholds for ownership disclosures. I conduct a comprehensive assessment by considering all different investor types in the market in their potential roles of sellers, holders and

buyers of stock. I can do so because I have a dataset of privately reported holdings, encompassing both publicly disclosed and undisclosed investments. For identification, I use three regulatory changes effectively lowering reporting thresholds. I exploit that they affect only holdings in firms in the regulated and not the unregulated market. Results of the difference-in-differences analysis show that mutual funds and, to a lesser extent, banks decrease while non-financial corporations increase their holdings. Results of descriptive analyses support the notion that mutual funds and banks merely reduce their holdings in size and do not completely sell them off, consistent with the avoidance of reporting thresholds.

In supplemental analyses, I turn to the subset of publicly observable holdings from the disclosure instrument that is the subject of the regulatory changes, notifications, and the investor type with the most robust findings, mutual funds. Results again suggest that mutual funds reduce their holdings to avoid reporting thresholds. Relative to all other investors, they more frequently disclose stakes just below the initial reporting threshold. In addition, such activities seem to be concentrated around the first disclosure increase, consistent with a gradual response to stricter enforcement. Relative to all notifications of mutual funds, notifications for falling below a threshold occur less frequently in the later years.

Finally, I reconcile my findings with the extant literature on liquidity effects [e.g., Christensen et al., 2016; Rubin, 2007]. In line with previous work, I illustrate that liquidity is, in general, positively associated with both mutual fund ownership and ownership disclosure. I then show that these two channels almost completely offset each other for firms that are likely most affected by the reduction in mutual funds' holdings.

Taken together, the paper emphasizes that ownership disclosure rules can impact investor behavior and thereby ownership structures. Like prior studies, it stresses the importance of investor heterogeneity [e.g., Cronqvist and Fahlenbrach, 2009; He et al., 2004]. Besides, it underlines that investors' incentives do not arise in isolation, but depend on the investor composition in the market. My findings inform about the ranking of disclosure costs among investors. Even if all investors perceive disclosure as costly ex ante, the market mechanism will ensure that only investors with the *relatively* highest disclosure costs reduce their holdings. Investor heterogeneity hence makes holding adjustments feasible in that it provides for trading partners. Future research faces ample opportunities to explore this diversity further. Settings differ, internationally as well as over time, in investor types and their respective concentrations.

# A Appendix

## A.1 Variable Definitions

| Variable | Description |
|---|---|
| $stake\_\star_{it}$ | The number of shares held by investor type $\star$ divided by the total number of ordinary shares outstanding (NOSH multiplied by 1,000) multiplied by 100.<br>$\star$ is one of the following: foreign investors ($foreign$), mutual funds ($funds$), banks ($banks$), insurance companies and pension funds ($ins$), non-financial corporations ($corp$), governmental institutions ($gov$), other financial investors ($other$) or households ($hh$). |
| $post7_t$ (9) [12] | Indicator variable equal to one if quarter $t$ is larger than $2006q4$ ($2008q4$) [$2011q4$].<br>If the unit of observation is the notification: $post9_t$ (12) is an indicator variable equal to one if trading day $t$ is after February 28, 2009 (January 31, 2012). |
| $treat_i$ | Indicator variable equal to one if firm $i$ is part of the treatment group, i.e., listed on the regulated market. |
| $mv_{it}$ [$ln(mv)_{it}$] | [Natural logarithm of one plus the] market value in millions (MV). |
| $up_{it}$ [$ln(up)_{it}$] | [Natural logarithm of one plus the] unadjusted closing price (UP). |
| $ana_{it}$ [$ln(ana)_{it}$] | [Natural logarithm of one plus the] number of analysts following the firm. |
| $age_{it}$ [$ln(age)_{it}$] | [Natural logarithm of one plus] quarter $t$ less the quarter from which Datastream holds information about the firm (BDATE). |
| $dy_{it}$ | Dividend yield (DY), i.e., dividends per share as percentage of the share price, ranging from 0 to 1. |
| $cdax_i$ | Indicator variable equal to one if firm $i$ is a member of the index CDAX. |
| $spread_{it}$ [$ln(spread)_{it}$] | [Natural logarithm of one plus the] quarterly mean of the daily bid-ask spread. The daily bid-ask spread is based on the daily closing bid (PB) and ask (PA) prices adjusted for capital actions: $= \frac{(PA-PB)}{(PA+PB)/2}$. |
| $ln(to)_{it}$ | Natural logarithm of one plus the quarterly mean of daily share turnover. Daily share turnover is daily unadjusted volume (UVO) over total shares outstanding (NOSH) multiplied by 100. |
| $ln(sd\_ret)_{it}$ | Natural logarithm of one plus the quarterly standard deviation of daily returns. Daily returns are based on the total return index (RI). |
| $high_i$ | Indicator variable equal to one if $stake\_\star_{i(2006q4)}$ is above the median in $2006q4$. |
| $stake_j$ | Stake disclosed in notification $j$. |

*Continued on next page*

*Continued from previous page*

| **Variable** | **Description** |
|---|---|
| $mf_j$ | Indicator variable equal to one if notification $j$ is submitted by a mutual fund. |
| $share\_mf\_below_{it}$ | The number of notifications for falling below submitted by mutual funds over the total number of notifications submitted by mutual funds. |
| $share\_mf_{it}$ | The number of notifications submitted by mutual funds over the total number of notifications disclosed. |

*Notes*: $i$ denotes firm, $j$ notification and $t$ quarter (trading day if the unit of observation is the notification). If applicable, the Datastream or Worldscope mnemonic is displayed in parentheses.

## A.2 Example of a Notification

[redacted]

[redacted]

WKN: [redacted] ISIN: [redacted] Land: Deutschland

Nachricht vom 13.11.2012 | 17:58

**[redacted]: Release according to Article 26, Section 1 of the WpHG [the German Securities Trading Act] with the objective of Europe-wide distribution**

[redacted]

13.11.2012 17:58

Dissemination of a Voting Rights Announcement, transmitted by
DGAP - a company of [redacted].
The issuer is solely responsible for the content of this announcement.

---

On November 12, 2012, [redacted], USA has informed us according to Article 21, Section 1 of the WpHG that via shares its Voting Rights on [redacted], Deutschland, have fallen below the 3% threshold of the Voting Rights on November 06, 2012 and on that day amounted to 2.99% (this corresponds to 9042891 Voting Rights).
According to Article 22, Section 1, Sentence 1, No. 6 in connection with sentence 2 of the WpHG, 2.99% of the Voting Rights (this corresponds to 9042891 Voting Rights) is to be attributed to the company.

13.11.2012 DGAP's Distribution Services include Regulatory Announcements, Financial/Corporate News and Press Releases.
Media archive at www.dgap-medientreff.de and www.dgap.de

---

Language: English
Company: [redacted]
Germany
Internet: www.[redacted].de

End of Announcement DGAP News-Service

## References


Abadie, A., Diamond, A. and Hainmueller, J. [2015], 'Comparative politics and the synthetic control method', *American Journal of Political Science* **59**(2), 495–510.

Abadie, A. and Gardeazabal, J. [2003], 'The economic costs of conflict: A case study of the Basque Country', *The American Economic Review* **93**(1), 113–132.

Agarwal, V., Gay, G. D. and Ling, L. [2014], 'Window dressing in mutual funds', *Review of Financial Studies* **27**(11), 3133–3170.

Agarwal, V., Jiang, W., Tang, Y. and Yang, B. [2013], 'Uncovering hedge fund skill from the portfolio holdings they hide', *The Journal of Finance* **68**(2), 739–783.

Agarwal, V., Mullally, K. A., Tang, Y. and Yang, B. [2015], 'Mandatory portfolio disclosure, stock liquidity, and mutual fund performance', *The Journal of Finance* **70**(6), 2733–2776.

Agarwal, V., Vashishtha, R. and Venkatachalam, M. [2017], 'Mutual fund transparency and corporate myopia', *Working Paper*. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2728969.

Amann, M., Baltzer, M. and Schrape, M. [2012], 'Microdatabase: Securities holdings statistics – a flexible multi-dimensional approach for providing user-targeted securities investment data', *Deutsche Bundesbank: Technical Documentation*. Available at https://www.bundesbank.de/Redaktion/EN/Downloads/Bundesbank/Research_Centre/microdatabase_securities_holdings_statistics.pdf?__blob=publicationFile.

Appel, I. R., Gormley, T. A. and Keim, D. B. [2016], 'Passive investors, not passive owners', *Journal of Financial Economics* **121**(1), 111–141.

Aragon, G. O., Hertzel, M. and Shi, Z. [2013], 'Why do hedge funds avoid disclosure? Evidence from confidential 13F filings', *Journal of Financial and Quantitative Analysis* **48**(05), 1499–1518.

Bade, M., Flory, J. and Schönberg, T. [2016], 'SHS-Base, data report 2016-02 – metadata version 1-1', *Deutsche Bundesbank Research Data and Service Centre: Data Report*. Available at https://www.bundesbank.de/Redaktion/EN/Downloads/Bundesbank/Research_Centre/research_data_micro_data_shsb_data_report.pdf?__blob=publicationFile.

BaFin [2008], *Annual Report 2007*, Bundesanstalt für Finanzdienstleistungsaufsicht, Bonn and Frankfurt am Main.

BaFin [2013*a*], *Annual Report 2012*, Bundesanstalt für Finanzdienstleistungsaufsicht, Bonn and Frankfurt am Main.

BaFin [2013*b*], *Emittentenleitfaden*, 4 edn, Bundesanstalt für Finanzdienstleistungsaufsicht, Bonn and Frankfurt am Main.

Baltzer, M., Jank, S. and Smajlbegovic, E. [2015], 'Who trades on momentum?', *Working Paper*. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2517462.

Bhattacharya, U. and Daouk, H. [2002], 'The world price of insider trading', *The Journal of Finance* **57**(1), 75–108.

Boone, A. L. and White, J. T. [2015], 'The effect of institutional ownership on firm transparency and information production', *Journal of Financial Economics* **117**(3), 508–533.

Brav, A., Jiang, W., Partnoy, F. and Thomas, R. [2008], 'Hedge fund activism, corporate governance, and firm performance', *The Journal of Finance* **63**(4), 1729–1775.

Brunnermeier, M. K. and Pedersen, L. H. [2005], 'Predatory trading', *The Journal of Finance* **60**(4), 1825–1863.

Christensen, H. B., Hail, L. and Leuz, C. [2016], 'Capital-market effects of securities regulation: Prior conditions, implementation, and enforcement', *Review of Financial Studies* **29**(11), 2885–2924.

Clifford, C. P. [2008], 'Value creation or destruction? Hedge funds as shareholder activists', *Journal of Corporate Finance* **14**(4), 323–336.

Coval, J. and Stafford, E. [2007], 'Asset fire sales (and purchases) in equity markets', *Journal of Financial Economics* **86**(2), 479–512.

Cronqvist, H. and Fahlenbrach, R. [2009], 'Large shareholders and corporate policies', *The Review of Financial Studies* **22**(10), 3941–3976.

Dahlquist, M. and Robertsson, G. [2001], 'Direct foreign ownership, institutional investors, and firm characteristics', *Journal of Financial Economics* **59**(3), 413–440.

DDV [2010], *Annual Report 2009*, Deutscher Derivate Verband, Frankfurt am Main and Berlin.

Del Guercio, D. [1996], 'The distorting effect of the prudent-man laws on institutional equity investments', *Journal of Financial Economics* **40**(1), 31–62.

Demsetz, H. [1968], 'The cost of transacting', *The Quarterly Journal of Economics* **82**(1), 33–53.

Deutsche Bundesbank [2013], 'Current developments in the mutual funds market: Demand, structural changes and investor behaviour', *Monthly Report January*.

Deutsche Bundesbank [2014], 'Ownership structure in the German equity market: General trends and changes in the financial crisis', *Monthly Report September*.

Deutsche Bundesbank [2015], 'Securities holdings statistics for analysing holdings of securities in Germany and Europe: Methodology and results', *Monthly Report March*.

Duong, T. X., Huszár, Z. R. and Yamada, T. [2015], 'The costs and benefits of short sale disclosure', *Journal of Banking & Finance* **53**, 124–139.

Edmans, A. [2014], 'Blockholders and corporate governance', *Annual Review of Financial Economics* **6**, 23–50.

Edmans, A., Heinle, M. S. and Huang, C. [2016], 'The real costs of financial efficiency when some information is soft', *Review of Finance* **20**(6), 2151–2182.

European Commission [2008], *Report on more stringent national measures concerning Directive 2004/109/EC on the harmonisation of transparency requirements in relation to information about issuers whose securities are admitted to trading on a regulated market*, Commission Staff Working Document, Brussels.

European Commission [2012], *Report from the Commission: Impact assessment*, Commission Staff Working Paper, Brussels.

Ferreira, M. A. and Matos, P. [2008], 'The colors of investors' money: The role of institutional investors around the world', *Journal of Financial Economics* **88**(3), 499–533.

Fishman, M. J. and Hagerty, K. M. [1995], 'The mandatory disclosure of trades and market liquidity', *Review of Financial Studies* **8**(3), 637–676.

Frank, M. M., Poterba, J. M., Shackelford, D. A. and Shoven, J. B. [2004], 'Copycat funds: Information disclosure regulation and the returns to active management in the mutual fund industry', *The Journal of Law & Economics* **47**(2), 515–541.

Franks, J. and Mayer, C. [2001], 'Ownership and control of German corporations', *Review of Financial Studies* **14**(4), 943–977.

Ge, W. and Zheng, L. [2006], 'The frequency of mutual fund portfolio disclosure', *Working Paper*. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=557186.

Gow, I. D., Ormazabal, G. and Taylor, D. J. [2010], 'Correcting for cross-sectional and time-series dependence in accounting research', *The Accounting Review* **85**(2), 483–512.

He, J., Ng, L. and Wang, Q. [2004], 'Quarterly trading patterns of financial institutions', *The Journal of Business* **77**(3), 493–509.

Heflin, F. and Shaw, K. W. [2000], 'Blockholder ownership and market liquidity', *The Journal of Financial and Quantitative Analysis* **35**(4), 621–633.

Holmström, B. [1999], 'Managerial incentive problems: A dynamic perspective', *The Review of Economic Studies* **66**(1), 169–182.

Hu, H. T. C. [2015], 'Financial innovation and governance mechanisms: The evolution of decoupling and transparency', *Business Lawyer* **70**(2), 347–405.

Hu, H. T. C. and Black, B. [2006], 'The new vote buying: Empty voting and hidden (morphable) ownership', *Southern California Law Review* **79**(4), 811–908.

Huddart, S., Hughes, J. S. and Levine, C. B. [2001], 'Public disclosure and dissimulation of insider trades', *Econometrica* **69**(3), 665–681.

Iacus, S. M., King, G. and Porro, G. [2012], 'Causal inference without balance checking: Coarsened exact matching', *Political Analysis* **20**(1), 1–24.

Ince, O. S. and Porter, R. B. [2006], 'Individual equity return data from Thomson Datastream: Handle with care!', *Journal of Financial Research* **29**(4), 463–479.

Jank, S. [2011], 'Are there disadvantaged clienteles in mutual funds?', *CFR Working Paper Series* 11-02. Available at https://ideas.repec.org/p/zbw/cfrwps/1102.html.

Jank, S., Roling, C. and Smajlbegovic, E. [2016], 'Flying under the radar: The effects of short-sale disclosure rules on investor behavior and stock prices', *Working Paper*. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2783762.

John, K. and Narayanan, R. [1997], 'Market manipulation and the role of insider trading regulations', *The Journal of Business* **70**(2), 217–247.

Jones, C. M., Reed, A. V. and Waller, W. [2016], 'Revealing shorts: An examination of large short position disclosures', *Review of Financial Studies* **29**(12), 3278–3320.

Kick, T., Onali, E., Ruprecht, B. and Schaeck, K. [2014], 'Wealth shocks, credit-supply shocks, and asset allocation', *European Central Bank Working Paper Series No.* 1662. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2408449.

Kyle, A. S. [1985], 'Continuous auctions and insider trading', *Econometrica* **53**(6), 1315–1335.

Lakonishok, J., Shleifer, A., Thaler, R. and Vishny, R. [1991], 'Window dressing by pension fund managers', *The American Economic Review* **81**(2), 227–231.

Leuz, C. and Verrecchia, R. E. [2000], 'The economic consequences of increased disclosure', *Journal of Accounting Research* **38**, 91–124.

Ljungqvist, A. and Qian, W. [2016], 'How constraining are limits to arbitrage?', *Review of Financial Studies* **29**(8), 1975–2028.

Mietzner, M. and Schweizer, D. [2014], 'Hedge funds versus private equity funds as shareholder activists in Germany – differences in value creation', *Journal of Economics and Finance* **38**(2), 181–208.

Natter, M., Rohleder, M., Schulte, D. and Wilkens, M. [2016], 'The benefits of option use by mutual funds', *Journal of Financial Intermediation* **26**, 142–168.

Petersen, M. A. [2009], 'Estimating standard errors in finance panel data sets: Comparing approaches', *The Review of Financial Studies* **22**(1), 435–480.

Prat, A. [2005], 'The wrong kind of transparency', *The American Economic Review* **95**(3), 862–877.

Rohleder, M., Schulte, D. and Wilkens, M. [2017], 'Management of flow risk in mutual funds', *Review of Quantitative Finance and Accounting* **48**(1), 31–56.

Rubin, A. [2007], 'Ownership level, ownership concentration and liquidity', *Journal of Financial Markets* **10**(3), 219–248.

Schmidt, P. S., von Arx, U., Schrimpf, A., Wagner, A. F. and Ziegler, A. [2017], 'On the construction of common size, value and momentum factors in international stock markets: A guide with applications', *Working Paper*. Available at https://papers.ssrn.com/sol3/ papers.cfm?abstract_id=1738315.

Schouten, M. C. [2010], 'The case for mandatory ownership disclosure', *Stanford Journal of Law, Business & Finance* **15**, 127–182.

Shive, S. and Yun, H. [2013], 'Are mutual funds sitting ducks?', *Journal of Financial Economics* **107**(1), 220–237.

Stehle, R. and Schmidt, M. H. [2015], 'Return on German stocks 1954 to 2013', *Credit and Capital Markets* **48**(3), 427–476.

Veil, R., Ruckes, M., Limbach, P. and Doumet, M. [2015], 'Today's or yesterday's news?', *Zeitschrift für Unternehmens- und Gesellschaftsrecht* **44**(5), 709–753.

Verbeek, M. and Wang, Y. [2013], 'Better than the original? The relative success of copycat funds', *Journal of Banking & Finance* **37**(9), 3454–3471.

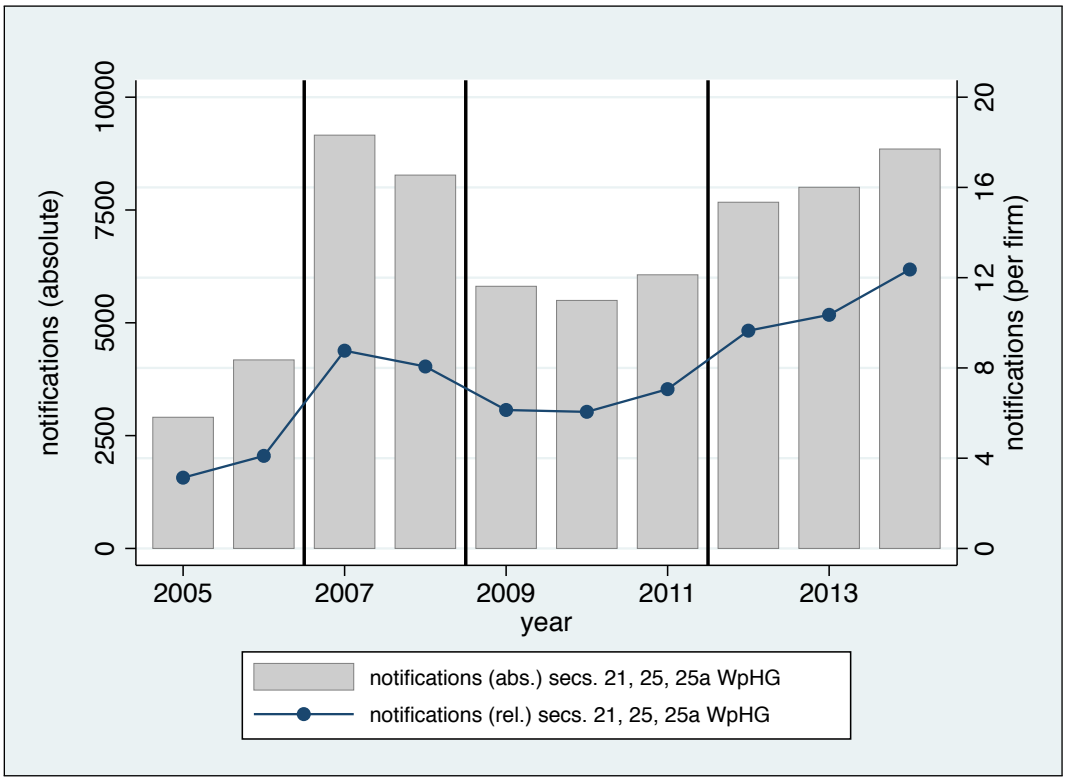
Verrecchia, R. E. [2001], 'Essays on disclosure', *Journal of Accounting and Economics* **32**(1–3), 97–180.

Vismara, S., Paleari, S. and Ritter, J. R. [2012], 'Europe's second markets for small companies', *European Financial Management* **18**(3), 352–388.

# Figures

Figure 1: **Notifications over Time**

*Notifications*



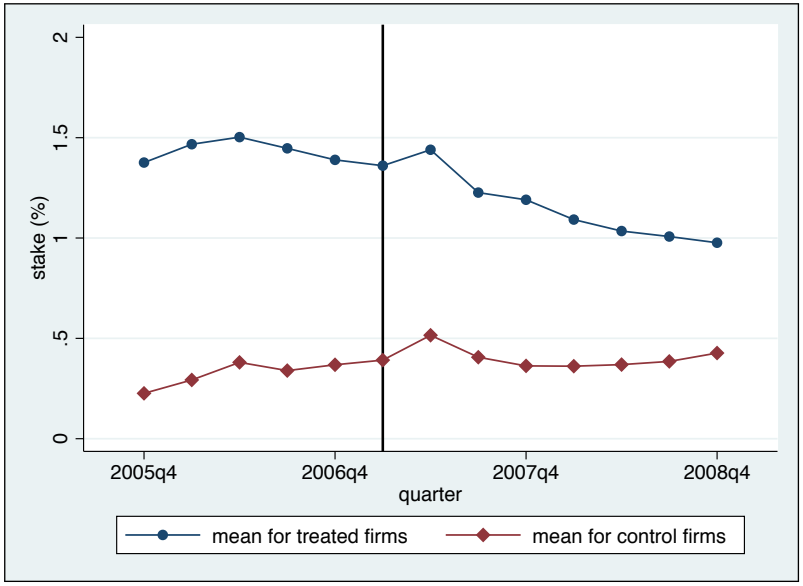

*Notes:* The figure displays the number of notifications for secs. 21, 25 and 25a WpHG per year, in absolute numbers (gray bars) and per firm (blue line). It shows all notifications disclosed in the German market and reported by the BaFin. The number of firms corresponds to the number of firms listed on the regulated market. Source: annual reports of the BaFin, available at: www.bafin.de.

Figure 2: **Mean Stakes around the Disclosure Increase in 2007**

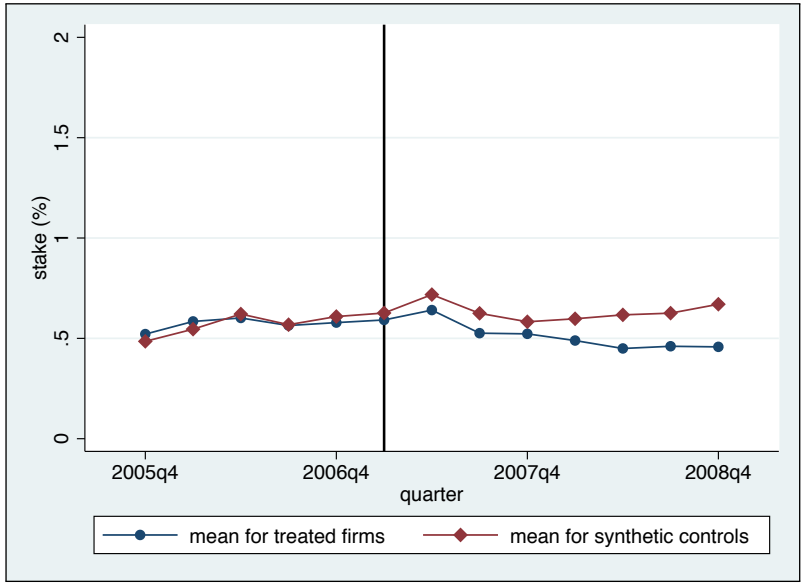
*Holdings*

**Mutual Funds**

(a) *Raw*




(b) *Synthetic Control Method*



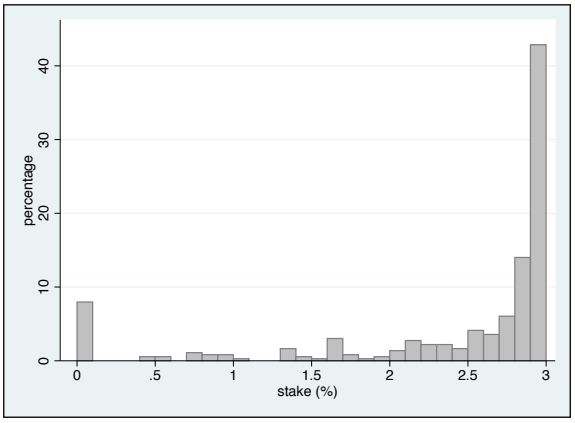

*Notes:* The figure displays the quarterly mean of mutual funds' stakes, $stake\_fund_{it}$, for the period 2005$q4$–2008$q4$. Figure 2a shows the mean of the raw stakes for a balanced panel. The number of treated (control) firms is equal to 375 (75) in each quarter. Figure 2b shows the mean of (raw) stakes for treated firms and synthetic controls. Stakes for synthetic controls are weighted combinations of stakes in control firms. I construct them by matching each treated firm to a combination of control firms whose combined stake most closely tracks the stake in the treated firm in the pre-treatment period. To ensure comparability, I limit the treatment group to firms with stakes less than or equal to the maximum stake in the control group in the same quarter. I further restrict the sample to a balanced panel. The number of treated firms and synthetic controls is equal to 322 in each quarter. I use 75 control firms for the synthetic controls. Source: Research Data and Service Centre (RDSC), Securities Holdings Statistics (SHS-Base) 2005–2008, own calculations.

Figure 3: **Stakes around the 3% Threshold**

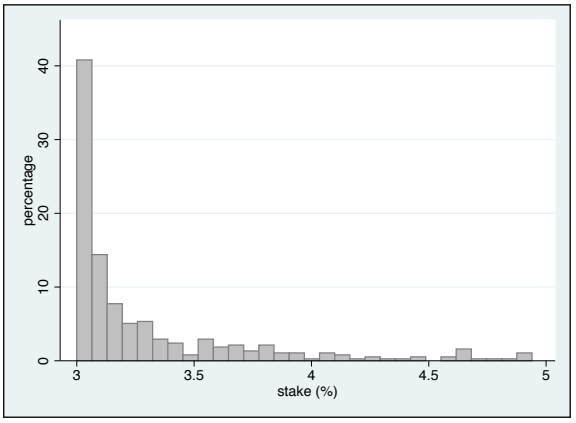
*Notifications*

**Mutual Funds**

(a) *Falling Below*

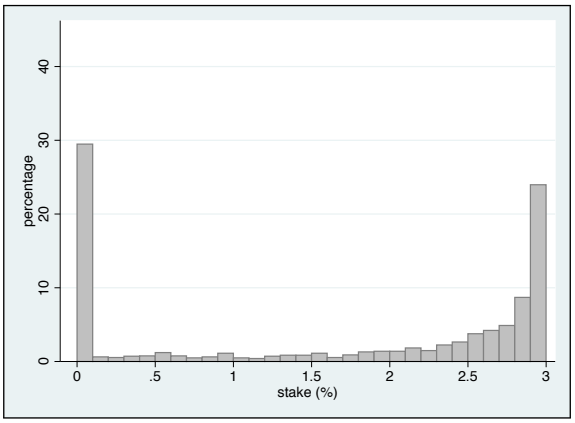



(b) *Exceeding*

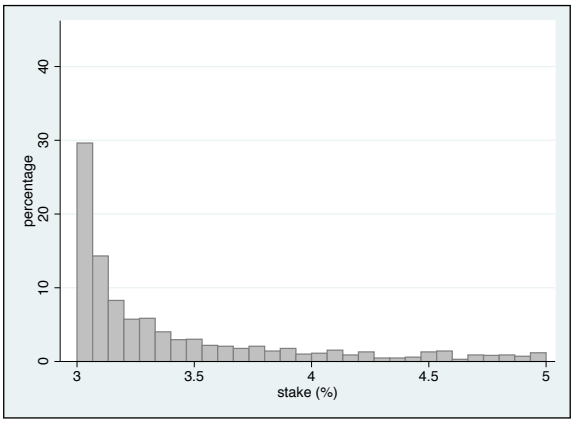



**Other Investors**

(c) *Falling Below*




(d) *Exceeding*




*Notes:* The figure displays the distribution of reported stakes around the 3% threshold, separately for falling below and exceeding the threshold as well as mutual funds and other investors. The number of observations is equal to 364 in Figure 3a, 375 in Figure 3b, 2,232 in Figure 3c and 1,691 in Figure 3d. Source: notifications distributed by the DGAP 2007–2012, own calculations.

# Tables

Table 1: **Firms across Quarters**

*Holdings*

| Quarter | All | Treatment | Control | Treated Firms in % of Firms in the Regulated Market |
|---|---|---|---|---|
| 2005q4 | 701 | 581 | 120 | 63 |
| 2006q4 | 868 | 647 | 221 | 63 |
| 2007q4 | 990 | 659 | 331 | 63 |
| 2008q4 | 1,100 | 677 | 423 | 66 |
| 2009q4 | 1,140 | 670 | 470 | 71 |
| 2010q4 | 1,184 | 663 | 521 | 73 |
| 2011q4 | 1,202 | 654 | 548 | 76 |
| 2012q4 | 1,144 | 632 | 512 | 79 |
| 2013q4 | 1,104 | 619 | 485 | 80 |
| 2014q4 | 814 | 505 | 309 | 71 |
| All years | 1,549 | 814 | 735 | |

*Notes:* The table shows the number of firms in the last quarter of each sample year in the holdings sample, for the whole sample (column 1) and separately for treatment (column 2) and control group (column 3). It further shows the percentage of firms listed on the regulated market that I include in the treatment group (column 4). The total number of firms listed on the regulated market is published in the BaFin's annual reports, available at: www.bafin.de. The last row (All years) depicts the number of unique firms over the entire sample period. Source: Research Data and Service Centre (RDSC), Securities Holdings Statistics (SHS-Base) 2005–2014, own calculations.

Table 2: **Summary Statistics**

Panel A: *Holdings*

| | | | | | Value-weighted | | $10^{th}$ Decile | |
|---|---|---|---|---|---|---|---|---|
| **Variable** | | **N** | **Mean** | **SD** | **N** | **Mean** | **N** | **Mean** |
| $stake\_foreign_{it}$ | | 38,330 | 30.904 | 30.236 | 37,470 | 42.652 | 3,833 | 92.629 |
| | $t$ | 23,363 | 27.303 | 25.758 | 22,890 | 42.142 | 2,336 | 83.059 |
| | $c$ | 14,967 | 36.526 | 35.408 | 14,580 | 50.510 | 1,496 | 96.620 |
| $stake\_funds_{it}$ | | 38,330 | 1.281 | 2.805 | 37,470 | 4.452 | 3,823 | 8.637 |
| | $t$ | 23,363 | 1.810 | 3.235 | 22,890 | 4.667 | 2,336 | 9.954 |
| | $c$ | 14,967 | 0.456 | 1.644 | 14,580 | 1.145 | 1,496 | 4.369 |
| $stake\_banks_{it}$ | | 38,330 | 0.549 | 1.835 | 37,470 | 0.934 | 3,833 | 4.309 |
| | $t$ | 23,363 | 0.674 | 2.009 | 22,890 | 0.973 | 2,336 | 4.943 |
| | $c$ | 14,967 | 0.352 | 1.502 | 14,580 | 0.343 | 1,496 | 3.091 |
| $stake\_ins_{it}$ | | 38,330 | 0.159 | 0.901 | 37,470 | 0.385 | 3,833 | 1.549 |
| | $t$ | 23,363 | 0.215 | 1.029 | 22,890 | 0.395 | 2,336 | 2.027 |
| | $c$ | 14,967 | 0.072 | 0.643 | 14,580 | 0.242 | 1,496 | 0.722 |
| $stake\_corp_{it}$ | | 38,330 | 29.644 | 31.521 | 37,470 | 31.091 | 3,832 | 92.677 |
| | $t$ | 23,363 | 31.069 | 31.035 | 22,890 | 31.213 | 2,336 | 91.819 |
| | $c$ | 14,967 | 27.419 | 32.141 | 14,580 | 29.208 | 1,496 | 93.969 |
| $stake\_gov_{it}$ | | 38,330 | 0.017 | 0.233 | 37,470 | 0.042 | 3,832 | 0.172 |
| | $t$ | 23,363 | 0.023 | 0.287 | 22,890 | 0.044 | 2,330 | 0.228 |
| | $c$ | 14,967 | 0.008 | 0.101 | 14,580 | 0.009 | 724 | 0.167 |
| $stake\_other_{it}$ | | 38,330 | 3.416 | 8.493 | 37,470 | 5.293 | 3,833 | 22.735 |
| | $t$ | 23,363 | 3.382 | 7.608 | 22,890 | 5.444 | 2,336 | 20.766 |
| | $c$ | 14,967 | 3.468 | 9.714 | 14,580 | 2.968 | 1,495 | 25.826 |
| $stake\_hh_{it}$ | | 38,330 | 34.030 | 27.953 | 37,470 | 15.150 | 3,833 | 85.847 |
| | $t$ | 23,363 | 35.523 | 27.752 | 22,890 | 15.123 | 2,336 | 85.918 |
| | $c$ | 14,967 | 31.698 | 28.107 | 14,580 | 15.575 | 1,496 | 85.664 |

*Notes:* The table displays summary statistics for the whole sample (line 1) and separately for treatment ($t$, line 2) and control group ($c$, line 3). The columns on the left (right) contain statistics for (the top decile of) equal-weighted percentage stakes and the middle columns contains value-weighted stakes, i.e., $stake\_\star_{it}$ weighted by market value, $mv_{it}$. $i$ and $t$ denote firm and quarter-year, respectively. $stake\_\star_{it}$ is the number of shares held by investor type $\star$ divided by the number of shares outstanding, multiplied by 100. $\star$ is one of the following: foreign investors ($foreign$), mutual funds ($funds$), banks ($banks$), insurance companies and pension funds ($ins$), non-financial corporations ($corp$), governmental institutions ($gov$), other financial investors ($other$) or households ($hh$). Source: Research Data and Service Centre (RDSC), Securities Holdings Statistics (SHS-Base) 2005–2014, own calculations.

Panel B: *Firm Characteristics*

| **Variable** | | **N** | **Mean** | **SD** | **Variable** | | **N** | **Mean** | **SD** |
|---|---|---|---|---|---|---|---|---|---|
| $post7_t$ | | 38,330 | 0.900 | 0.300 | $post9_t$ | | 38,330 | 0.690 | 0.462 |
| | $t$ | 23,363 | 0.871 | 0.336 | | $t$ | 23,363 | 0.643 | 0.479 |
| | $c$ | 14,967 | 0.945 | 0.228 | | $c$ | 14,967 | 0.763 | 0.425 |
| $post12_t$ | | 38,330 | 0.327 | 0.469 | $treat_i$ | | 38,330 | 0.610 | 0.488 |
| | $t$ | 23,363 | 0.304 | 0.460 | | $t$ | 23,363 | 1.000 | 0.000 |
| | $c$ | 14,967 | 0.362 | 0.481 | | $c$ | 14,967 | 0.000 | 0.000 |
| $mv_{i(t-1)}$ | | 35,499 | 231.361 | 894.180 | $up_{i(t-1)}$ | | 35,499 | 23.308 | 77.668 |
| | $t$ | 22,265 | 348.356 | 1,089.411 | | $t$ | 22,265 | 26.557 | 80.816 |
| | $c$ | 13,234 | 34.529 | 293.764 | | $c$ | 13,234 | 17.842 | 71.734 |
| $ln(mv)_{i(t-1)}$ | | 35,499 | 2.944 | 2.132 | $ln(up)_{i(t-1)}$ | | 35,499 | 1.702 | 1.465 |
| $ana_{it}$ | | 35,499 | 2.169 | 5.742 | $age_{it}$ | | 35,499 | 44.832 | 34.694 |
| | $t$ | 22,265 | 3.317 | 6.962 | | $t$ | 22,265 | 57.251 | 34.450 |
| | $c$ | 13,234 | 0.237 | 0.969 | | $c$ | 13,234 | 23.938 | 23.153 |
| $ln(ana)_{it}$ | | 35,499 | 0.515 | 0.903 | $ln(age)_{it}$ | | 35,499 | 3.473 | 0.945 |
| $dy_{i(t-1)}$ | | 35,499 | 0.009 | 0.019 | $cdax_i$ | | 35,499 | 0.520 | 0.500 |
| | $t$ | 22,265 | 0.012 | 0.020 | | $t$ | 22,265 | 0.828 | 0.377 |
| | $c$ | 13,234 | 0.005 | 0.016 | | $c$ | 13,234 | 0.000 | 0.000 |
| $spread_{it}$ | | 20,324 | 0.062 | 0.095 | | | | | |
| | $t$ | 16,144 | 0.052 | 0.079 | | | | | |
| | $c$ | 4,180 | 0.100 | 0.134 | | | | | |
| $ln(spread)_{it}$ | | 20,324 | 0.057 | 0.077 | | | | | |

*Notes:* The table displays summary statistics for the whole sample (line 1) and separately for treatment ($t$, line 2) and control group ($c$, line 3). $i$ and $t$ denote firm and quarter-year, respectively. $post7_t$ (9) [12] is an indicator variable equal to one if quarter $t$ is larger than $2006q4$ ($2008q4$) [$2011q4$]. $treat_i$ is an indicator variable equal to one if firm $i$ is part of the treatment group, i.e., listed on the regulated market. $mv_{it}$ [$ln(mv)_{it}$] is the [natural logarithm of one plus the] market value. $up_{it}$ [$ln(up)_{it}$] is the [natural logarithm of one plus the] unadjusted closing price. $ana_{it}$ [$ln(ana)_{it}$] is the [natural logarithm of one plus the] number of analysts following the firm. $age_{it}$ [$ln(age)_{it}$] is the [natural logarithm of one plus] firm age in quarters. $dy_{it}$ is dividends as percentage of the share price. $cdax_i$ is an indicator variable equal to one if firm $i$ is a member of the index CDAX. $spread_{it}$ [$ln(spread)_{it}$] is the [natural logarithm of one plus the] quarterly mean of the daily bid-ask spread.

Table 3: **Main Model**

*Holdings*

| | $stake\_\star_{it}$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | $foreign$ | $funds$ | $banks$ | $ins$ | $corp$ | $gov$ | $other$ | $hh$ |
| $post7_t \times treat_i$ | −1.401 | −0.378** | −0.100 | −0.030 | 1.381 | −0.011 | 0.267 | 0.273 |
| | (1.392) | (0.146) | (0.132) | (0.069) | (1.194) | (0.012) | (0.348) | (0.916) |
| $post9_t \times treat_i$ | −0.954 | −0.286* | −0.209** | −0.082** | 2.214** | −0.006 | −0.052 | −0.626 |
| | (0.978) | (0.144) | (0.096) | (0.032) | (1.048) | (0.006) | (0.377) | (0.854) |
| $post12_t \times treat_i$ | −0.828 | 0.097 | −0.022 | −0.072** | 2.436** | −0.004 | 0.385 | −1.991** |
| | (0.888) | (0.101) | (0.062) | (0.028) | (0.917) | (0.005) | (0.385) | (0.831) |
| Quarter-year fixed effects | yes | yes | yes | yes | yes | yes | yes | yes |
| Firm fixed effects | yes | yes | yes | yes | yes | yes | yes | yes |
| adj. $R^2$ | 0.806 | 0.671 | 0.496 | 0.557 | 0.796 | 0.556 | 0.528 | 0.825 |
| adj. $R^2$ within | 0.001 | 0.003 | 0.001 | 0.002 | 0.004 | 0.000 | 0.000 | 0.002 |
| $N$ | 38,330 | 38,330 | 38,330 | 38,330 | 38,330 | 38,330 | 38,330 | 38,330 |

*Notes:* The table displays estimates for the main difference-in-differences analysis. $i$ and $t$ denote firm and quarter-year, respectively. $stake\_\star_{it}$ is the number of shares held by investor type $\star$ divided by the number of shares outstanding, multiplied by 100. $\star$ is one of the following: foreign investors ($foreign$), mutual funds ($funds$), banks ($banks$), insurance companies and pension funds ($ins$), non-financial corporations ($corp$), governmental institutions ($gov$), other financial investors ($other$) or households ($hh$). $post7_t$ (9) [12] is an indicator variable equal to one if quarter $t$ is larger than $2006q4$ ($2008q4$) [$2011q4$]. $treat_i$ is an indicator variable equal to one if firm $i$ is part of the treatment group, i.e., listed on the regulated market. Standard errors are two-way clustered by firm and quarter-year and denoted below the coefficients in parentheses. *, ** and *** indicate significance at the 0.1, 0.05 and 0.01 level, respectively. Source: Research Data and Service Centre (RDSC), Securities Holdings Statistics (SHS-Base) 2005–2014, own calculations.

Table 4: **Main Model Variations**

*Holdings*

| | $stake\_\star_{it}$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | ***foreign*** | ***funds*** | ***banks*** | ***ins*** | ***corp*** | ***gov*** | ***other*** | ***hh*** |
| **(1) Balanced Panel** | | | | | | | | |
| $post7_t \times treat_i$ | −1.410 | −0.351*** | −0.394*** | −0.018 | 2.623** | −0.001 | 0.050 | −0.499 |
| | (1.265) | (0.060) | (0.127) | (0.073) | (1.204) | (0.004) | (0.504) | (0.829) |
| $post9_t \times treat_i$ | 0.207 | −0.406* | −0.226 | −0.075 | 0.108 | 0.003 | 0.137 | 0.252 |
| | (0.959) | (0.214) | (0.228) | (0.049) | (1.380) | (0.006) | (0.834) | (1.543) |
| $post12_t \times treat_i$ | 0.549 | 0.421*** | 0.083 | −0.046 | 2.070 | −0.011* | 0.390 | −3.456* |
| | (1.296) | (0.153) | (0.077) | (0.040) | (1.704) | (0.006) | (0.647) | (1.706) |
| Quarter-year fixed effects | yes | yes | yes | yes | yes | yes | yes | yes |
| Firm fixed effects | yes | yes | yes | yes | yes | yes | yes | yes |
| adj. $R^2$ | 0.795 | 0.652 | 0.520 | 0.588 | 0.839 | 0.662 | 0.471 | 0.863 |
| adj. $R^2$ within | 0.000 | 0.004 | 0.004 | 0.001 | 0.002 | 0.000 | 0.000 | 0.003 |
| | *N = 15,117; 375 (75) treated (control) firms* | | | | | | | |
| **(2) Linear Time Trend** | | | | | | | | |
| $post7_t \times treat_i$ | −0.110 | −0.319** | −0.067 | −0.012 | 1.284 | 0.002 | 0.098 | −0.876 |
| | (1.409) | (0.148) | (0.146) | (0.056) | (1.552) | (0.004) | (0.402) | (1.052) |
| $post9_t \times treat_i$ | 0.507 | −0.274* | −0.244 | −0.068* | −1.186 | 0.003 | −0.141 | 1.403 |
| | (1.043) | (0.146) | (0.159) | (0.037) | (1.185) | (0.005) | (0.663) | (1.131) |
| $post12_t \times treat_i$ | 0.520 | 0.076 | 0.134 | −0.007 | −0.421 | −0.002 | 0.232 | −0.531 |
| | (1.181) | (0.167) | (0.121) | (0.024) | (1.201) | (0.003) | (0.356) | (0.862) |
| Firm fixed effects × trend | yes | yes | yes | yes | yes | yes | yes | yes |
| Quarter-year fixed effects | yes | yes | yes | yes | yes | yes | yes | yes |
| Firm fixed effects | yes | yes | yes | yes | yes | yes | yes | yes |
| adj. $R^2$ | 0.853 | 0.790 | 0.677 | 0.792 | 0.894 | 0.942 | 0.651 | 0.924 |
| adj. $R^2$ within | 0.000 | 0.002 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 | 0.001 |
| | *N = 23,722; 588 (175) treated (control) firms* | | | | | | | |
| **(3) Control Variables** | | | | | | | | |
| $post7_t \times treat_i$ | −2.390 | −0.291* | −0.051 | −0.017 | 0.087 | 0.005 | −0.127 | 2.783** |
| | (1.760) | (0.152) | (0.137) | (0.070) | (1.466) | (0.011) | (0.407) | (1.133) |
| $post9_t \times treat_i$ | −8.363*** | −0.178 | −0.099 | −0.052** | 4.281*** | 0.003 | 0.303 | 4.105*** |
| | (1.668) | (0.133) | (0.083) | (0.025) | (1.304) | (0.005) | (0.425) | (1.075) |
| $post12_t \times treat_i$ | −1.097 | 0.188* | −0.070 | −0.032 | 1.124 | −0.009 | −0.343 | 0.239 |
| | (1.320) | (0.099) | (0.068) | (0.031) | (1.214) | (0.007) | (0.397) | (1.227) |
| $CONT$ and $treat_i$ | yes | yes | yes | yes | yes | yes | yes | yes |
| Quarter-year fixed effects | yes | yes | yes | yes | yes | yes | yes | yes |
| Industry fixed effects | yes | yes | yes | yes | yes | yes | yes | yes |
| adj. $R^2$ | 0.146 | 0.387 | 0.030 | 0.028 | 0.163 | 0.023 | 0.031 | 0.245 |
| adj. $R^2$ within | 0.123 | 0.350 | 0.020 | 0.015 | 0.125 | 0.006 | 0.019 | 0.211 |
| | *N = 35,499; 798 (714) treated (control) firms* | | | | | | | |

| | $stake\_\star_{it}$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | ***foreign*** | ***funds*** | ***banks*** | ***ins*** | ***corp*** | ***gov*** | ***other*** | ***hh*** |
| **(4) Matched Firms: PSM with Replacement** | | | | | | | | |
| $post7_t \times treat_i$ | $-4.978^{*}$ | $-0.770^{**}$ | 0.246 | $-0.051$ | $3.804^{***}$ | $-0.001$ | $-0.766$ | $2.517^{*}$ |
| | (2.510) | (0.316) | (0.529) | (0.087) | (1.128) | (0.001) | (0.739) | (1.273) |
| $post9_t \times treat_i$ | 2.073 | $-0.692$ | $-1.305^{*}$ | $-0.128$ | 2.206 | 0.005 | 1.596 | $-3.754$ |
| | (2.094) | (0.496) | (0.670) | (0.078) | (3.512) | (0.006) | (1.083) | (3.171) |
| $post12_t \times treat_i$ | 0.864 | 0.406 | 0.087 | $-0.005$ | 3.234 | $-0.017^{**}$ | 0.232 | $-4.801$ |
| | (2.299) | (0.283) | (0.339) | (0.037) | (3.791) | (0.008) | (0.776) | (3.000) |
| Quarter-year fixed effects | yes | yes | yes | yes | yes | yes | yes | yes |
| Firm fixed effects | yes | yes | yes | yes | yes | yes | yes | yes |
| adj. $R^2$ | 0.796 | 0.682 | 0.718 | 0.660 | 0.812 | 0.673 | 0.431 | 0.814 |
| adj. $R^2$ within | 0.006 | 0.021 | 0.036 | 0.005 | 0.012 | 0.003 | 0.007 | 0.019 |
| *N = 18,368; 289 (34) treated (control) firms* | | | | | | | | |
| **(5) Matched Firms: CEM with Weights** | | | | | | | | |
| $post7_t \times treat_i$ | 1.410 | $-0.574^{***}$ | $-0.465^{*}$ | $-0.117$ | 0.762 | 0.002 | 0.328 | $-1.348$ |
| | (1.428) | (0.143) | (0.235) | (0.129) | (2.065) | (0.002) | (0.550) | (1.774) |
| $post9_t \times treat_i$ | $4.507^{*}$ | $-0.370$ | $-1.984$ | $-0.137$ | $-5.135$ | 0.005 | 2.109 | 1.004 |
| | (2.435) | (0.240) | (1.546) | (0.109) | (3.087) | (0.010) | (1.716) | (3.421) |
| $post12_t \times treat_i$ | 0.481 | $0.684^{**}$ | $-0.246$ | $-0.010$ | $-0.011$ | $-0.015^{*}$ | 0.325 | $-1.209$ |
| | (1.536) | (0.305) | (0.150) | (0.036) | (3.211) | (0.009) | (0.751) | (2.774) |
| Quarter-year fixed effects | yes | yes | yes | yes | yes | yes | yes | yes |
| Firm fixed effects | yes | yes | yes | yes | yes | yes | yes | yes |
| adj. $R^2$ | 0.801 | 0.582 | 0.608 | 0.599 | 0.810 | 0.608 | 0.374 | 0.828 |
| adj. $R^2$ within | 0.008 | 0.007 | 0.061 | 0.008 | 0.004 | 0.001 | 0.009 | 0.000 |
| *N = 5,821; 137 (35) treated (control) firms* | | | | | | | | |

*Notes:* The table displays regression estimates for variations of the main model: (1) a sample restricted to firms that have complete quarterly observations for the period 2005$q4$–2008$q4$; (2) the main model extended with firm-specific linear time trends; (3) a simplified model with only industry and quarter-year fixed effects, but extended with a set of control variables, $CONT$; (4) a sample of firms from propensity score matching (PSM); and (5) a sample of firms from coarsened exact matching (CEM). Before matching, samples are restricted to firms that have complete quarterly observations for the period 2005$q4$–2008$q4$. $i$ and $t$ denote firm and quarter-year, respectively. $stake\_\star_{it}$ is the number of shares held by investor type $\star$ divided by the number of shares outstanding, multiplied by 100. $\star$ is one of the following: foreign investors ($foreign$), mutual funds ($funds$), banks ($banks$), insurance companies and pension funds ($ins$), non-financial corporations ($corp$), governmental institutions ($gov$), other financial investors ($other$) or households ($hh$). $post7_t$ (9) [12] is an indicator variable equal to one if quarter $t$ is larger than 2006$q4$ (2008$q4$) [2011$q4$]. $treat_i$ is an indicator variable equal to one if firm $i$ is part of the treatment group, i.e., listed on the regulated market. $CONT$ includes the following control variables: the natural logarithm of one plus the market value, $ln(mv)_{i(t-1)}$; the natural logarithm of one plus the unadjusted closing price, $ln(up)_{i(t-1)}$; the natural logarithm of one plus the number of analysts following the firm, $ln(ana)_{it}$; the natural logarithm of one plus firm age in quarters, $ln(age)_{it}$; dividends as percentage of the share price, $dy_{i(t-1)}$; and an indicator variable equal to one if firm $i$ is a member of the index CDAX, $cdax_i$. Standard errors are two-way clustered by firm and quarter-year and denoted below the coefficients in parentheses. *, ** and *** indicate significance at the 0.1, 0.05 and 0.01 level, respectively. Source: Research Data and Service Centre (RDSC), Securities Holdings Statistics (SHS-Base) 2005–2014, own calculations.

Table 5: **Non-zero Investments in the Market**

*Holdings*

| *number of firms*<br>*market value holdings* | | ***foreign*** | | ***funds*** | | ***banks*** | | ***ins*** | | ***corp*** | | ***gov*** | | ***other*** | | ***hh*** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *abs.* | *(%)* | *abs.* | *(%)* | *abs.* | *(%)* | *abs.* | *(%)* | *abs.* | *(%)* | *abs.* | *(%)* | *abs.* | *(%)* | *abs.* | *(%)* | *N* |
| **PRE** | ***t*** | 599 | (99) | 330 | (55) | 493 | (82) | 321 | (53) | 604 | (100) | 158 | (26) | 576 | (95) | 604 | (100) | 5 |
| | | 108,550 | (44) | 12,396 | (5) | 2,877 | (1) | 1,935 | (1) | 73,308 | (30) | 55 | (0) | 9,623 | (4) | 36,244 | (15) | 5 |
| | ***c*** | 156 | (95) | 35 | (21) | 69 | (43) | 27 | (17) | 165 | (100) | 15 | (9) | 144 | (87) | 165 | (100) | 5 |
| | | 4,410 | (40) | 400 | (4) | 85 | (1) | 134 | (1) | 2,781 | (29) | 4 | (0) | 229 | (2) | 2,352 | (24) | 5 |
| **POST** | ***t*** | 660 | (99) | 368 | (55) | 516 | (78) | 332 | (50) | 663 | (100) | 177 | (27) | 636 | (96) | 663 | (100) | 8 |
| | | 106,091 | (44) | 10,289 | (4) | 2,495 | (1) | 1,235 | (1) | 73,996 | (31) | 26 | (0) | 10,862 | (5) | 33,430 | (14) | 8 |
| | ***c*** | 330 | (97) | 66 | (20) | 135 | (39) | 45 | (14) | 332 | (98) | 17 | (5) | 266 | (80) | 336 | 99 | 8 |
| | | 5,583 | (38) | 184 | (1) | 67 | (0) | 43 | (0) | 5,534 | (40) | 1 | (0) | 421 | (3) | 2,406 | (17) | 8 |

*Notes:* The table displays the number of firms with non-zero holdings and the market value of the holdings, separately for treatment ($t$) and control group ($c$), pre-period (PRE) and post-period (POST) as well as investor type $\star$. $\star$ is one of the following: foreign investors ($foreign$), mutual funds ($funds$), banks ($banks$), insurance companies and pension funds ($ins$), non-financial corporations ($corp$), governmental institutions ($gov$), other financial investors ($other$) or households ($hh$). The table provides the number of firms with non-zero holdings and the market value of holdings in absolute terms ($abs.$) and as percentage ($\%$) of overall firms and of total market capitalization, respectively. The pre-period (post-period) includes $2005q4$–$2006q4$ ($2007q1$–$2008q4$). All numbers are averages of the quarterly values; $N$ depicts the number of quarters entering the calculation. Overall, there are 3,022 (824) observations in the treatment (control) group in the pre-period and 5,308 (2,718) in the post-period. Source: Research Data and Service Centre (RDSC), Securities Holdings Statistics (SHS-Base) 2005–2014, own calculations.

Table 6: **Analysis of Notifications**

*Notifications*

Panel A: **Stakes around the 3% Threshold**

| | $stake_j$ | | | |
|---|---|---|---|---|
| | *Falling Below* | | *Exceeding* | |
| $mf_j$ | 0.535*** | 0.538*** | −0.117*** | −0.137*** |
| | (0.070) | (0.074) | (0.035) | (0.037) |
| $post9_t \times mf_j$ | −0.028 | −0.040 | 0.077 | 0.082 |
| | (0.092) | (0.097) | (0.094) | (0.081) |
| $post12_t \times mf_j$ | −0.027 | −0.011 | −0.055 | −0.029 |
| | (0.155) | (0.110) | (0.123) | (0.105) |
| Main effects $post\star$ | yes | yes | yes | yes |
| Firm fixed effects | yes | yes | yes | yes |
| Quarter-year fixed effects | no | yes | no | yes |
| Month fixed effects | no | yes | no | yes |
| Weekday fixed effects | no | yes | no | yes |
| adj. $R^2$ | 0.290 | 0.302 | 0.199 | 0.203 |
| adj. $R^2$ within | 0.029 | 0.028 | 0.005 | 0.006 |
| $N$ | 2,596 | 2,596 | 2,066 | 2,066 |

*Notes:* The table displays estimates for the regression of the disclosed stake, $stake_j$. The sample is restricted to notifications for falling below (exceeding) the 3% threshold in columns 1 and 2 (columns 3 and 4). $j$ and $t$ denote notification and trading day, respectively. $mf_j$ is an indicator variable equal to one if notification $j$ is submitted by a mutual fund. $post9_t$ (12) is an indicator variable equal to one if trading day $t$ is after February 28, 2009 (January 31, 2012). Standard errors are two-way clustered by firm and quarter-year and denoted below the coefficients in parentheses. *, ** and *** indicate significance at the 0.1, 0.05 and 0.01 level, respectively. Source: notifications distributed by the DGAP 2007–2012, own calculations.

Panel B: **Mutual Funds' Shares in Notifications over Time**

| | $share\_mf\_below_{it}$ | | $share\_mf_{it}$ | |
|---|---|---|---|---|
| $post9_t$ | −0.085** | −0.089** | 0.040** | 0.038** |
| | (0.036) | (0.039) | (0.016) | (0.016) |
| $post12_t$ | −0.058** | −0.051** | 0.077*** | 0.067*** |
| | (0.022) | (0.022) | (0.017) | (0.020) |
| $CONT$ | no | yes | no | yes |
| Industry fixed effects | yes | yes | yes | yes |
| adj. $R^2$ | 0.025 | 0.041 | 0.050 | 0.100 |
| adj. $R^2$ within | 0.015 | 0.039 | 0.019 | 0.073 |
| $N$ | 719 | 719 | 3,569 | 3,569 |

*Notes:* The table displays estimates for the regression of mutual funds' shares in notifications. $i$ and $t$ denote firm and quarter-year, respectively. $share\_mf\_below_{it}$ is the share of notifications for falling below a threshold in mutual funds' notifications. $share\_mf_{it}$ is mutual funds' share in overall notifications. $post9_t$ (12) is an indicator variable equal to one if quarter $t$ is larger than $2008q4$ ($2011q4$). $CONT$ includes the following control variables: the natural logarithm of one plus the market value, $ln(mv)_{i(t-1)}$; the natural logarithm of one plus the unadjusted closing price, $ln(up)_{i(t-1)}$; the natural logarithm of one plus the number of analysts following the firm, $ln(ana)_{it}$; the natural logarithm of one plus firm age in quarters, $ln(age)_{it}$; dividends as percentage of the share price, $dy_{i(t-1)}$; and an indicator variable equal to one if firm $i$ is a member of the index CDAX, $cdax_i$. Standard errors are two-way clustered by firm and quarter-year and denoted below the coefficients in parentheses. *, ** and *** indicate significance at the 0.1, 0.05 and 0.01 level, respectively. Source: notifications distributed by the DGAP 2007–2012, own calculations.

Table 7: **Correlation between Liquidity and Mutual Fund Ownership**

Panel A: *Notifications*

| | $\Delta ln(spread)_{it}$ | |
|---|---|---|
| | *Falling Below* | *Exceeding* |
| $mf_j$ | 0.00096** | −0.00039 |
| | (0.00040) | (0.00035) |
| $\Delta ln(mv)_{it}$ | −0.00861*** | −0.00695*** |
| | (0.00167) | (0.00122) |
| $\Delta ln(to)_{it}$ | −0.07159*** | −0.05727*** |
| | (0.01416) | (0.01428) |
| $\Delta ln(sd\_ret)_{it}$ | 0.15246*** | 0.08795*** |
| | (0.03628) | (0.02584) |
| Industry fixed effects | yes | yes |
| adj. $R^2$ | 0.128 | 0.104 |
| adj. $R^2$ within | 0.115 | 0.105 |
| $N$ | 2,225 | 1,833 |

*Notes:* The table displays estimates for the regression of the change in bid-ask spreads. The sample is restricted to notifications for falling below (exceeding) the 3% threshold in column 1 (column 2). $i$, $j$ and $t$ denote firm, notification and post-disclosure period, respectively. $ln(spread)_{it}$ is the three-month mean of daily bid-ask spreads. $mf_j$ is an indicator variable equal to one if notification $j$ is submitted by a mutual fund. $ln(mv)_{it}$ is the natural logarithm of one plus the market value. $ln(to)_{it}$ is the natural logarithm of one plus the three-month mean of daily share turnover. $ln(sd\_ret)_{it}$ is the natural logarithm of one plus the three-month standard deviation of daily returns. $\Delta$ is the difference between the three months post-disclosure and the three months pre-trade. Standard errors are clustered by firm and denoted below the coefficients in parentheses. *, ** and *** indicate significance at the 0.1, 0.05 and 0.01 level, respectively. Source: notifications distributed by the DGAP 2007–2012, own calculations.

Panel B: *Holdings*

| | $ln(spread)_{it}$ | |
|---|---|---|
| | ***funds*** | ***combined*** |
| $\Delta stake\_\star_{it}$ | −0.00047*** | −0.00009 |
| | (0.00015) | (0.00018) |
| $ln(mv)_{i(t-4)}$ | −0.01985*** | −0.01984*** |
| | (0.00165) | (0.00166) |
| $ln(to)_{i(t-4)}$ | −0.05280*** | −0.05287*** |
| | (0.01220) | (0.01221) |
| $ln(sd\_ret)_{i(t-4)}$ | 0.38588*** | 0.38600*** |
| | (0.06338) | (0.06338) |
| Quarter-year fixed effects | yes | yes |
| Firm fixed effects | yes | yes |
| adj. $R^2$ | 0.691 | 0.691 |
| adj. $R^2$ within | 0.112 | 0.112 |
| $N$ | 17,482 | 17,482 |

*Notes:* The table displays estimates for the regression of the bid-ask spread on changes in stakes of mutual funds (column 1) as well as mutual funds, banks and insurance companies and pension funds [combined] (column 2). $i$ and $t$ denote firm and quarter-year, respectively. $stake\_\star_{it}$ is the number of shares held by investor type $\star$ divided by the number of shares outstanding, multiplied by 100. $\star$ is mutual funds ($funds$) in column 1 and the combined category ($combined$) in column 2. $\Delta$ is the difference between quarter $t$ and quarter $(t-1)$. $ln(mv)_{i(t-4)}$ is the natural logarithm of one plus the market value. $ln(to)_{i(t-4)}$ is the natural logarithm of one plus the quarterly mean of daily share turnover. $ln(sd\_ret)_{i(t-4)}$ is the natural logarithm of one plus the quarterly standard deviation of daily returns. Standard errors are two-way clustered by firm and quarter-year and denoted below the coefficients in parentheses. *, ** and *** indicate significance at the 0.1, 0.05 and 0.01 level, respectively. Source: Research Data and Service Centre (RDSC), Securities Holdings Statistics (SHS-Base) 2005–2014, own calculations.

Table 8: **Liquidity Analysis**

*Holdings*

| | $ln(spread)_{it}$ | | | | |
|---|---|---|---|---|---|
| | | *funds* | | *combined* | |
| | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** |
| $post7_t \times treat_i$ | −0.020** | −0.028** | −0.032*** | −0.029** | −0.034** |
| | (0.009) | (0.011) | (0.011) | (0.014) | (0.014) |
| $post7_t \times treat_i \times high_i$ | | 0.021* | 0.024** | 0.013 | 0.018 |
| | | (0.011) | (0.010) | (0.015) | (0.014) |
| $post9_t \times treat_i$ | −0.017** | −0.017** | −0.037* | −0.015** | −0.043* |
| | (0.006) | (0.007) | (0.019) | (0.007) | (0.024) |
| $post9_t \times treat_i \times high_i$ | | 0.023** | 0.045** | 0.005 | 0.033 |
| | | (0.009) | (0.019) | (0.013) | (0.026) |
| $post12_t \times treat_i$ | 0.007 | 0.006 | 0.020 | 0.006 | 0.020 |
| | (0.005) | (0.006) | (0.015) | (0.006) | (0.024) |
| $post12_t \times treat_i \times high_i$ | | 0.003 | −0.011 | 0.009 | −0.005 |
| | | (0.007) | (0.015) | (0.007) | (0.024) |
| $CONT$ | yes | yes | yes | yes | yes |
| Remaining interactions | | yes | yes | yes | yes |
| Quarter-year fixed effects | yes | yes | yes | yes | yes |
| Firm fixed effects | yes | yes | yes | yes | yes |
| adj. $R^2$ | 0.688 | 0.691 | 0.678 | 0.691 | 0.679 |
| adj. $R^2$ within | 0.115 | 0.122 | 0.172 | 0.122 | 0.173 |
| $N$ | 20,324 | 20,324 | 13,334 | 20,324 | 13,334 |

*Notes:* The table displays estimates for the regression of the bid-ask spread on disclosure and ownership proxies. $i$ and $t$ denote firm and quarter-year, respectively. $ln(spread)_{it}$ is the natural logarithm of one plus the quarterly mean of daily bid-ask spreads. $post7_t$ (9) [12] is an indicator variable equal to one if quarter $t$ is larger than $2006q4$ ($2008q4$) [$2011q4$]. $treat_i$ is an indicator variable equal to one if firm $i$ is part of the treatment group, i.e., listed on the regulated market. $high_i$ is an indicator variable equal to one if $stake\_\star_{i(2006q4)}$ is above the median in $2006q4$, where $\star$ is $funds$ ($funds$, $banks$ and $ins$ [$combined$]) in columns 2 and 3 (4 and 5). The sample is restricted to firms with an observation for $2006q4$ in columns 3 and 5. $CONT$ includes the following control variables: the natural logarithm of one plus the market value, $ln(mv)_{i(t-4)}$; the natural logarithm of one plus the quarterly mean of daily share turnover, $ln(to)_{i(t-4)}$; and the natural logarithm of one plus the quarterly standard deviation of daily returns, $ln(sd\_ret)_{i(t-4)}$. Standard errors are two-way clustered by firm and quarter-year and denoted below the coefficients in parentheses. *, ** and *** indicate significance at the 0.1, 0.05 and 0.01 level, respectively. Source: Research Data and Service Centre (RDSC), Securities Holdings Statistics (SHS-Base) 2005–2014, own calculations.

# Accounting Comparability in Mutual Funds' Portfolios*

**Sarah Kröchert**
Humboldt University of Berlin

***Abstract***

The study examines whether accounting comparability matters for mutual funds' portfolio decisions. I measure accounting comparability at the holding level by assessing similarities with portfolio peers, explicitly adopting an investor perspective. Methodologically, I follow De Franco, Kothari and Verdi [2011], extended with a cash flow-based model. I first show that comparability is high and varies predictably with the type of mutual funds. For the same firm, comparability is higher in mutual fund than in analyst portfolios, which I derive from analysts' coverage decisions. Likewise, it is higher in portfolios of active funds. I then provide evidence consistent with comparability arising from the selection of already comparable firms. For the same portfolio, a firm is more likely to be included if it is more comparable to portfolio peers. For the firm that is included, comparability increases until and around inclusion, but not subsequently. The findings are in line with accounting comparability reducing portfolio management costs.

*I am grateful to Janja Brendel, Ulf Brüggemann, Tom Fischer, Joachim Gassen, Maximilian Muhn and Harm Schütt for their helpful feedback. I thank workshop participants at Humboldt University as well as the 2015 EAA Annual Congress, the 2015 RTG 1659 Summer School, the 2015 Trans-Atlantic Doctoral Conference and the 2014 WU-HU International Accounting Research Workshop.

# 1 Introduction

BlackRock, the largest asset manager worldwide and the provider of a wide range of mutual funds, regularly emphasizes the usefulness of comparable accounting information. When engaging in discussions with standard setters, it promotes accounting standards that enhance both clarity and comparability of financial statements. It underlines that its internal analysts, in their role as users of accounting information, spend a substantial amount of time scrutinizing the financial statements of existing and potential portfolio firms. Their main task lies in the adjustment of individual positions to eliminate effects from idiosyncratic choices and, thus, to make them more comparable.[38] However, BlackRock's business model does not primarily rely on manual adjustments, but a combination of human judgement and automated data processing. Many of its mutual funds have large and diversified portfolios and manage risks based on the technology platform Aladdin.[39] As of late, BlackRock officially cuts back on staff specialized in traditional forms of stock picking and expands on computerized alternatives, shifting money to its 'quants'. The rationale, in the words of the head of active equities, is that BlackRock "can more efficiently deliver alpha at a better cost with automated processes".[40] Handing over analysis tasks from humans to machines requires input that is already comparable if the output is to be of an equivalent quality. One source for this input are financial statements. One implication from this input is that the firms BlackRock evaluates and eventually invests in are more comparable in accounting terms. In this study, I investigate whether accounting comparability affects investment decisions of mutual funds in general, not only BlackRock.

Academic research suggests that mutual funds value accounting comparability. The underlying notion is that similar accounting information reduces information processing costs. Mutual funds frequently decide which firms to hold in their portfolios and need standardized information for sound decision-making [e.g., Covrig, DeFond and Hung, 2007; DeFond, Hu, Hung and Li, 2011]. Yet, this literature focuses on comparability benefits in international settings with apparent differences in accounting and auditing standards as well as enforcement structures. I

[38] BlackRock details its views on comparability and analyst activities, for example, when responding to standard setters' proposals through comment letters. See, e.g.: http://www.fasb.org/jsp/FASB/Page/-SectionPage&cid=1218220137090, http://www.fasb.org/jsp/FASB/Page/SectionPage&cid=1218220137090 or http://www.fasb.org/jsp/FASB/Page/SectionPage&cid=1218220137090. BlackRock is not the only institutional investor who argues along these lines. For instance, see: http://www.fasb.org/jsp/FASB/Page/SectionPage&cid=1218220137090, for the position of T. Rowe Price on the accounting for financial instruments.

[39] The use of Aladdin is not restricted to BlackRock. The asset manager distributes its technology to competitors and this to an extent that some discuss the potential for systemic risk arising from too uniform approaches to analysis, see, e.g.: https://www.economist.com or https://www.ft.com.

[40] The quote is from Mark Wiseman, BlackRock's global head of active equities, and included in Bloomberg's report about the cutbacks in the stock-picking department, available at: https://www.bloomberg.com.

concentrate on one country, the United States (US), with a single set of accounting standards, US Generally Accepted Accounting Principles (GAAP). In addition, I examine a country with extensive, well-developed standards for preparers and auditors and an enforcement system that strives for proper implementation of rules and monitoring of compliance [e.g., Dye and Sunder, 2001; Leuz and Wysocki, 2016]. In this setting, investors are familiar with prevailing accounting practices and encounter relatively transparent firms so that comparability does not have to be a key concern. Hence, I first study whether portfolios of mutual funds display high levels of accounting comparability and find that they do.

I then turn to the relevance of accounting comparability for making investment decisions. Extant research provides evidence for a *selection channel*, i.e., investors pick firms whose accounting is aligned with their preferences before investment, and an *influence channel*, i.e., investors attach less importance to the accounting in the initial investment stage and encourage the application of certain practices thereafter [e.g., Bradshaw, Bushee and Miller, 2004; Fang, Maffett and Zhang, 2015]. While, in reality, it likely is a combination of both, one channel might dominate. My findings are consistent with the selection and inconsistent with the influence channel.

I measure accounting comparability following De Franco, Kothari and Verdi [2011]. De Franco et al. propose an output-based construct that centers on how stock returns, the proxy for economic events, map into earnings, the proxy for financial statements. Since the informativeness of returns can depend on factors that are associated with mutual funds' portfolio decisions, I employ an additional model. Precisely, I use the mapping of cash flows, the alternative proxy for economic events, into accruals [Cascino and Gassen, 2015]. Two firms have comparable accounting systems if they produce similar financial statement outcomes for the same economic event. Central to the concept is the set of peers against which comparability is assessed. The accounting system of a firm, in itself, cannot be comparable. It can only be comparable to the accounting systems of other firms. Prior studies largely resort to industry peers [e.g., Brochet, Jagolinzer and Riedl, 2013; De Franco et al., 2011]. I choose portfolio peers as my argument rests on an *investor perspective*. I expect mutual funds to act on comparability because similar financial statement information allows for a thorough evaluation of potential and existing portfolio firms. Since I can only observe existing portfolio firms, I identify similarities with their accounting systems. The resulting proxies are return-based and cash flow-based comparability at the level of the holding. They vary across portfolios, i.e., the same firm has different comparability values for different sets of portfolio peers; they vary across firms, i.e., different firms have different comparability values due to different accounting systems; and they vary within firms, i.e., the same firm experiences changes in its accounting system over time.

I analyze portfolios of US mutual funds, which make up a significant portion of worldwide institutional investment. For example, US domestic equity funds manage 20% of global assets invested in open-end funds at the end of 2016 [ICI, 2017]. Fund managers have the discretion to decide about the inclusion of firms into their portfolios and therefore face potential costs from low comparability of financial statements. I do not consider entire portfolios, but limit observations to the holdings that should be most important to the fund manager. I measure importance with the portfolio weight of the holding, i.e., the percentage of the market value of the portfolio it represents [Fich, Harford and Tran, 2015]. I keep only holdings with a portfolio weight of at least 0.5%. I further limit portfolio firms to the largest, most visible firms to mitigate the influence of mechanical drivers of the comparability proxies (e.g., size or liquidity) that can simultaneously impact portfolio decisions. In the analyses, I work with different samples. I begin with the broadest possible sample and gradually reduce observations to the subset for which accounting comparability should matter most, holdings of actively managed funds. At various stages, I match observations to lower variation in factors I am not interested in. For the same firm, I match holdings from similar portfolios. For the same portfolio, I match holdings from similar firms. I apply the coarsened exact matching (CEM) algorithm that allows to control the degree of (im)balance in matched samples [Iacus, King and Porro, 2012].

To assess whether accounting comparability is high in mutual funds' portfolios, I need a benchmark. I use analysts' coverage decisions, a straightforward, but also ambitious choice. The comparability literature illustrates that analyst outcomes are associated with accounting comparability and explains the findings with lower information processing costs [e.g., De Franco et al., 2011; Neel, 2017], which closely resembles the explanation put forward for mutual funds. I create analyst portfolios from the firms analysts cover and corresponding return-based and cash flow-based comparability at the holding level. Raw differences in comparability are stark. Yet, analyst portfolios differ from mutual fund portfolios in several respects. I account for economic, mechanical drivers of comparability as well as investment preferences of mutual funds and expertise of analysts with a broad range of control variables at the firm and portfolio level. Moreover, I test for differences in comparability in tightly matched samples, comprising, for a single firm, holdings from both analyst and mutual fund portfolios. Results show that comparability in mutual fund portfolios is higher and that differences are large. In the matched samples, the difference in return-based (cash flow-based) comparability translates into 9% (8%) of average comparability in mutual fund portfolios.

To corroborate the above, I exploit cross-sectional variation in mutual funds. Actively managed, nonindexer funds need information to decide about buying, selling or holding stocks.

Within the group of nonindexer funds, stock pickers need even more granular information, in part certainly accounting data, to distinguish undervalued or overvalued stocks. Consequently, the documented comparability effects should originate from their portfolios. I adopt the concept of the active share of Cremers and Petajisto [2009] to proxy for stock picking and use the rough division into (non)indexers to proxy for the existence of investment discretion. I test for differences in comparability along the extent of stock picking and refine the sample by matching, for a single firm, holdings from nonindexer to indexer portfolios. Results show that, overall, comparability is increasing in the extent of stock picking.

So far, findings are in line with a preference for accounting comparability, of mutual funds in general and active funds in particular. This is not to say that the primary concern of mutual funds is the similarity of financial statements. Nevertheless, apart from mutual funds or buy-side analysts, also other intermediaries, such as sell-side analysts and rating agencies, are known to frequently correct reported numbers [e.g., Gu and Chen, 2004; Kim, Kraft and Ryan, 2013]. Furthermore, institutional investors appear to make investment decisions conditional on the accounting information environment [e.g., Bowen, Rajgopal and Venkatachalam, 2014; Bushee and Noe, 2000]. Thus, comparable accounting information might be a secondary, but still necessary requirement to get an understanding of firms.

To examine whether accounting comparability impacts the selection of portfolio firms and, in this vein, mutual funds' investment behavior, I focus on inclusions into portfolios. That is, I study the point in time when mutual funds build up positions in firms. I match each firm that is included to another firm that is neither included nor an existing member of the portfolio. By construction, the matched firm pair shares certain firm characteristics, but only one firm is added to the portfolio. I can show that the probability of inclusion is modestly increasing in accounting comparability, with the same set of portfolio peers. Specifically, a one-standard-deviation increase in return-based (cash flow-based) comparability leads to an increase in the probability of inclusion of 2% (3%). I can further show that the information environment of firms moderates the relation. The impact of accounting comparability is lower for firms with high bid-ask spreads or high return volatility. The finding is consistent with mutual funds valuing comparable accounting information for standardized, automated analyses. Low bid-ask spreads and low return volatility, indicative of stable and predictable operations, alleviate the need for manual adjustments.

To address the relative importance of the selection and influence channel, I track changes in accounting comparability over time, around the inclusion into portfolios. I restrict the sample to newly included holdings that remain in the portfolio for at least four years. I require such

a long holding period mainly because of the measurement of the comparability proxies. At the same time, it should favor the incidence of (in)direct communication and influence activities [McCahery, Sautner and Starks, 2016]. I assess comparability before, around and after inclusion, with the same portfolio. Results reinforce the above findings. Within holdings, comparability clearly increases before inclusion, supporting the selection channel. However, it does not change (much) afterwards, questioning the relevance of the influence channel in my setting.

Hence, the evidence suggests that BlackRock's funds are not an exception. Active mutual funds, on average, seem to value accounting comparability and this by the time they start pursuing investments. These investors regularly rely on automated data processing for which they need comparable information from the financial statements. Even if they apply adjustments, they need information that can be easily standardized, with minimal human involvement [Kim et al., 2013]. Their methods provide a rationale for the findings and, in particular, for the selection channel.

The study comes with the caveat that the comparability proxies might not fully grasp the concept of accounting comparability. They are, by definition, confined to summary metrics of the financial statements [De Franco et al., 2011]. In addition, descriptives illustrate that especially the return-based proxy is subject to economic fluctuations. Nonetheless, descriptives equally reveal that (raw values of) return-based and cash flow-based comparability do not behave alike. Although results are overall weaker for cash flow-based comparability, both proxies yield similar conclusions.

The paper contributes to the literature on institutional investors and accounting comparability. Research frequently works with different accounting regimes and changes therein to operationalize (dis)similarity of financial statements. Bradshaw et al. [2004] consider specific accounting method choices and Yu and Wahid [2014] determine accounting distance in country pairs based on differences in local accounting rules. At a more aggregate level, studies employ the adoption of the International Financial Reporting Standards (IFRS) [e.g., Covrig et al., 2007; DeFond et al., 2011; Florou and Pope, 2012]. Other work shifts the focus to accounting outcomes, such as summary measures of the financial statements or disclosure formats [e.g., Fang et al., 2015; Jung, 2013]. While all of these studies take into account a rich variety of firm and investor characteristics to fill the conceptual notion of accounting comparability, none incorporates characteristics of other portfolio firms. My study attempts to do exactly this by estimating comparability within mutual funds' portfolios. Portfolio holdings are informative about investors' preferences. Certainly, my approach has limitations in that I only observe preferences as revealed through actual choices, i.e., I observe the outcome of the investment process and not

the single stages of analysis. Still, it refines the investor perspective.

Besides, the paper relates to the literature on peer selection. Extant work discusses difficulties in and proposes solutions for identifying industry peers [e.g., Bhojraj, Lee and Oler, 2003; Hoberg and Phillips, 2016] and peers for valuation purposes [e.g., Bhojraj and Lee, 2002; Young and Zeng, 2015]. My findings imply that portfolios of mutual funds, or even institutional investors in general, might be another reference point since these investors select firms in a manner that results in higher accounting comparability. Along similar lines, Ramnath [2002] derives an industry classification from common analyst coverage.

The study proceeds as follows. Section 2 describes the measurement of accounting comparability and Section 3 the data. Section 4 presents the results. Section 5 concludes.

## 2 Measurement of Accounting Comparability

According to the Financial Accounting Standards Board (FASB), "for information to be comparable, like things must look alike and different things must look different" [FASB, 2010, QC23]. De Franco et al. [2011] translate this statement into a measurable construct by assessing the similarity of accounting functions across firms. I closely follow them in the estimation of firm-pair comparability, but differ in the choice of the peer group for the subsequent aggregation into holding-level comparability.

The estimation starts at the firm level. The accounting function, the empirical model for the firm-specific accounting system, maps stock returns, a proxy for economic events, into earnings, a proxy for financial statements. With stock returns, I do not only capture economic events, but also variation in the informativeness of returns about these events. I face a problem for inference if the factors driving this variation simultaneously affect portfolio composition. One candidate is stock liquidity that is associated with both informativeness of returns and mutual funds' holdings [e.g., Falkenstein, 1996; Gassen, Skaife and Veenman, 2016]. To mitigate the influence of these factors, I resort to an additional model for the firm-specific accounting system that maps cash flows, an alternative proxy for economic events, into accruals [Cascino and Gassen, 2015]. I obtain the return-based and cash flow-based accounting functions from estimating the following equations:

$$earnings_{it} = \alpha_{0,i} + \alpha_{1,i} return_{it} + \epsilon_{it}$$

$$accruals_{it} = \beta_{0,i} + \beta_{1,i} cfo_{it} + \epsilon_{it}$$

where $i$ and $t$ denote firm and quarter-year, respectively. $earnings_{it}$ is income before extraordinary items scaled by lagged market value of equity. $return_{it}$ is quarterly stock return. $accruals_{it}$

is the difference between income before extraordinary items and operating cash flow scaled by lagged total assets. $cfo_{it}$ is operating cash flow scaled by lagged total assets. I estimate the equations for each firm $i$ and quarter $t$ over the 16 previous quarters; I relax the number of quarters to 14 if data are unavailable. The estimated coefficients $\hat{\alpha}_{0,i}$ and $\hat{\alpha}_{1,i}$ ($\hat{\beta}_{0,i}$ and $\hat{\beta}_{1,i}$) proxy for the return-based (cash flow-based) accounting function.[41]

To determine the similarity between the accounting functions of firm $i$ and $j$, I derive average absolute forecast errors in the firm pair. For the return-based model, I use the coefficients $\hat{\alpha}_{0,i}$ and $\hat{\alpha}_{1,i}$ and $return_{it}$ to predict firm $i$'s earnings from its own accounting system, $E(earnings)_{iit}$; and I use the coefficients $\hat{\alpha}_{0,j}$ and $\hat{\alpha}_{1,j}$ and $return_{it}$ to predict firm $i$'s earnings from the accounting system of firm $j$, $E(earnings)_{ijt}$. I repeat the procedure for each one of the 16 quarters. For the cash flow-based model, I proceed accordingly to arrive at $E(accruals)_{iit}$ and $E(accruals)_{ijt}$. Accounting comparability is the negative of the average absolute forecast error:

$$compacct_{ijt} = -\frac{1}{16} \times \sum_{t-15}^{t} |E(earnings)_{iit} - E(earnings)_{ijt}|$$

$$compcf_{ijt} = -\frac{1}{16} \times \sum_{t-15}^{t} |E(accruals)_{iit} - E(accruals)_{ijt}|$$

Lower forecast errors, i.e., higher values of $compacct_{ijt}$ and $compcf_{ijt}$, correspond to higher firm-pair accounting comparability. Conceptually, they indicate that both firms depict the same set of underlying events in a similar fashion in their financial statements. The nature of the events is not of interest, they can be idiosyncratic or market-wide. However, the measures incorporate only events experienced by one firm (in the above case, firm $i$), meaning that they are firm-specific.

I then turn to the holding level. To do so, I first have to define the relevant peer group, i.e., which firms $j$ to consider for firm $i$. De Franco et al. [2011] work with industry peers. I choose portfolio peers since my argument rests on an investor perspective. I expect mutual funds to have stakes in firms whose accounting information is similar. To pick up the investor choice, I compute firm-pair comparability for each firm $i$ with all firms $j$ in portfolio $k$. Next, I rank all firm-pair values for firm $i$ and take the mean of the highest ten, resulting in holding-level comparability proxies for the return-based and cash flow-based model: $compacct_{ikt}$ and $compcf_{ikt}$. I focus on ten peers as I require at least eleven firms for every portfolio that is included in the sample.

[41] Although my sample period varies, estimates of the return-based accounting function are very close to De Franco et al. [2011]. De Franco et al. obtain a mean of 0.00 (0.02) with a standard deviation of 0.04 (0.08) for $\alpha_{0,i}$ ($\alpha_{1,i}$) in a sample of 71,295 firm-years for the period 1981–2009. I obtain a mean of 0.00 (0.02) with a standard deviation of 0.03 (0.07) for $\alpha_{0,i}$ ($\alpha_{1,i}$) in a sample of 172,956 firm-quarters for the period 2001–2015.

Relying on ten peers therefore maximizes information content while ensuring an identical number of peers for each firm.[42] The proxies are increasing in comparability and reflect average forecast errors with the most comparable portfolio peers; I multiply the values by 100 so that forecast errors are expressed as percentages.

At a single point in time, the comparability proxies, $compacct_{ikt}$ and $compcf_{ikt}$, vary across portfolios because portfolios differ in their sets of peers; and across firms because firms differ in their accounting functions. The same firm generally has different comparability values if it is part of different portfolios and thus exposed to different peers. Over time, the proxies vary further within firms because accounting functions change. Due to the long estimation period, accounting functions are however sticky. In the analyses, I exploit all of these layers of variation to understand whether accounting comparability is one factor determining the composition of mutual funds' portfolios.

Finally, a remark on my choice of output-based comparability proxies. The literature offers alternative, input-based approaches, exploring accounting method choices [e.g., Bradshaw et al., 2004; Bradshaw, Miller and Serafeim, 2009]. I regard them as less appropriate in my setting. Firms in my study compile their financial statements in accordance with a single set of accounting standards, US GAAP, which contains few explicit accounting options. Yet, firms face numerous implicit options when asked to exercise judgement and to specify assumptions, which are hard to grasp individually. By concentrating on summary measures of the financial statements (earnings and accruals), I intend to capture discretionary choices.

# 3 Data

I examine quarterly data from the first quarter of 2001 to the third quarter of 2015. In this vein, the sample period is long enough to yield, to a certain extent, generalizable results. At the same time, it consists of more recent years with substantial advances in automated data processing, which my argument emphasizes.

I begin by identifying firms for which I can estimate accounting comparability. I gather information from the CRSP/Compustat Merged Database. In accordance with De Franco et al. [2011], I keep only firms with fiscal year-end in March, June, September or December and drop firms whose names indicate that they are holding companies, American Depository Receipts or limited partnerships (e.g., to avoid comparing financial statements within groups). I also

[42] Selecting the ten firms with the highest firm-pair comparability values introduces a mechanical relation with portfolio size. The larger the number of firms in the portfolio, the higher holding-level comparability. I account for the number of peers in the analyses, e.g., by matching on it.

keep only firms with an average stock price above \$5 and average market capitalization above \$100 million. I do so to avoid measuring associations that are driven by differences in firm fundamentals, affecting the comparability proxies as well as mutual funds' portfolio choices. Pre-selecting in this fashion ensures that firms are, in general, more similar. I drop all firm-quarters with missing data on earnings and returns (accruals and cash flows) in the preceding 14 quarters for the estimation of return-based (cash flow-based) comparability. Lastly, I require available information on control variables and membership in portfolios with at least ten peers. Panel A of Table 1 depicts summary statistics of firm characteristics for the resulting universe of firm-quarters. The number of observations differs for control variables of cash flow-based comparability ($acc\_volatility_{it}$ and $cfo\_volatility_{it}$) as data coverage is lower. In total, there are 3,162 unique firms that are part of at least one mutual fund portfolio in at least one quarter. Firms are among the most liquid and largest in the market. For example, the average bid-ask spread is equal to 0.3% ($spread_{it}$) and average market capitalization amounts to \$6,396 million (untabulated; Table 1 contains instead the natural logarithm of market capitalization, $ln(mv)_{it}$).

I then define portfolios of mutual funds. I collect quarterly holdings from the Thomson Reuters S12 file. Thomson Reuters gathers most of its information from statutory filings, quarterly reports on Form N-CSR(S) and N-Q since 2004 and semiannual reports on Form N-30D before. Prior to 2004, mutual funds regularly provide voluntary portfolio disclosures [e.g., Agarwal, Mullally, Tang and Yang, 2015], leading to relatively comprehensive quarterly coverage. I work with the portfolios as available in the database.[43] In merging fund holdings with firm information, I use the calendar quarter of the holdings report date. Report dates do not necessarily coincide with the end of the calendar quarter as statutory filing deadlines refer to fiscal periods and mutual funds have varying fiscal year-ends. However, there is some overlap. In each sample year, mutual funds report more than 60% of all holdings for calendar quarter-ends. Besides, I do not regard time overlap as crucial since the comparability measure per se is rather sticky. I next restrict portfolios to holdings that are likely most important to the mutual fund. In a sense, this is another aspect of adopting an investor perspective. I expect that fund managers assess accounting comparability relative to portfolio firms. I further expect that 'important' portfolio firms receive a larger (than proportional) weight in the assessment. I measure importance to the fund manager with the portfolio weight of the holding, i.e., the fraction of the overall market value of the portfolio that the holding represents [Fich et al., 2015]. I drop all holding-quarters with a portfolio weight below 0.5%. Note that this restriction, by construction, produces com-

[43] I repeat all analyses with holdings from only mandatory portfolio disclosures by dropping observations for the years 2001 until 2004. Results are similar.

parability values that are lower than or equal to the values for entire portfolios. Accounting comparability corresponds to the mean of the ten highest-ranked firm-pair comparability values. Enlarging the set of peers can only alter the mean if firm-pair comparability values are higher with the newly included peers.

I clean up the sample and characterize mutual funds in greater detail with information from the CRSP Mutual Fund Database, linked to Thomson Reuters via Mutual Fund Links (MFLINKS). In doing so, I follow previous studies [e.g., Cremers and Petajisto, 2009; Wermers, 2000]. I drop all mutual funds with investment objectives that indicate that they are not primarily invested in US equities.[44] Similarly, I eliminate mutual funds that invest on average less than 50% of their funds into common stock over the whole sample period and require that the market value of the holdings selected above equals 25% or more of total net assets at a single point in time. CRSP reports total net assets separately for different share classes. I employ the value-weighted average and do not differentiate between share classes. Finally, I limit the country of origin of mutual funds to the US. To describe the mutual fund type, I distinguish, among others, between indexers and nonindexers. I classify index funds based on the variable $index\_fund\_flag$ and name matches with terms related to indexing strategies [e.g., Lines, 2016; Ma, Tang and Gómez, 2016].[45] Panel B of Table 1 depicts summary statistics of fund characteristics for the universe of fund-quarters. A fund-quarter only enters if the portfolio consists of at least eleven firms after applying the filters from above; the variable $peers_{kt}$ informs about the number of peers for each firm in the portfolio. In total, there are 4,219 unique mutual fund portfolios for which holding-level comparability is available in at least one quarter. 14.8% of fund-quarters are classified as indexer ($indexer_{kt}$).

I use information on analyst coverage from IBES and on index constituents of the S&P 500 from Compustat.

In the subsequent sections, I analyze different samples. I start off with the broadest possible set of observations and successively reduce sample size, in part due to a lack of data, but mostly to focus on the question at hand. Sample firms, funds and holdings can differ across specifications. The unit of observation is the holding, i.e., the firm-portfolio combination, in a given quarter.

[44] Specifically, I start with the CRSP objective code and keep funds with domestic equity cap-based, style, mixed or missing objectives. Among the ones with missing information, I keep funds that, according to Thomson Reuters, have an investment objective of aggressive growth, growth, growth and income, balanced or missing. Whenever I use investment objectives in the analyses, I confine mutual funds to the subset for which I have complete information from CRSP.

[45] In particular, I check for matches with the following terms: Index, Ind, Idx, Indx, Mkt, Market, Composite, S&P, SP, Russell, Nasdaq, DJ, Dow, Jones, Wilshire, NYSE, iShares, SPDR, HOLDRs, ETF, Exchange-Traded Fund, PowerShares, StreetTRACKS, 100, 400, 500, 600, 1000, 1500, 2000, 3000, 5000.

# 4 Results

## 4.1 Mutual Funds vs. Analysts

Extant research suggests that institutional investors value comparable accounting information because it can reduce information processing costs [e.g., Bradshaw et al., 2004; Fang et al., 2015]. That is, in managing large portfolios, investors benefit from standardized, clearly defined items in financial statements. However, these studies usually investigate international settings in which the application of different accounting standards and the divergence of local accounting practices hamper accounting comparability in a straightforward manner. I examine the US setting in which the application of a uniform set of accounting standards and more harmonized accounting practices likely result in similar financial statements in general. In addition, domestic institutional investors can regularly resort to alternative information sources, such as local market experts or firm management [e.g., Covrig et al., 2007; DeFond et al., 2011]. Hence, I first assess whether financial statement comparability still matters.

Precisely, I study whether mutual funds display a preference for accounting comparability through their portfolio choices. I contrast comparability in mutual fund portfolios with comparability in analyst portfolios, which are an obvious, but also quite demanding benchmark. The literature discusses potential comparability benefits for analysts and provides supportive evidence [e.g., De Franco et al., 2011; Horton, Serafeim and Serafeim, 2013; Neel, 2017]. The rationale for analysts resembles the rationale for mutual funds to a great extent. Both should experience lower information processing costs from more similar accounting information. Thus, analyst portfolios are an ambitious benchmark in that they may already be tailored towards firms with comparable accounting.

I create analyst portfolios from analysts' coverage decisions. In each quarter, I identify all firms for which a specific analyst issues a forecast and assign them to this analyst's portfolio. For each firm in the analyst portfolio, i.e., the analyst holding, I derive return-based and cash flow-based comparability. The steps are identical to mutual fund portfolios, the only difference is the set of peers. Again, the same firm can be part of several analyst portfolios in which it has different comparability values depending on its peers.

Panel A of Table 2 displays summary statistics of accounting comparability, separately for return-based ($compacct_{ikt}$) and cash flow-based comparability ($compcf_{ikt}$) as well as mutual fund (*mutual funds*) and analyst portfolios (*analysts*). Samples correspond to the first set of samples of the analysis described below. The number of observations shows that mutual fund

holdings exceed analyst holdings. $compacct_{ikt}$ ($compcf_{ikt}$) has 1,136,150 (1,119,375) observations for mutual funds and 251,758 (210,844) observations for analysts. Summary statistics demonstrate that comparability is higher in mutual fund than in analyst portfolios. $compacct_{ikt}$ ($compcf_{ikt}$) has a mean of -0.557 (-0.592) for mutual funds and of -1.234 (-1.015) for analysts. Differences in medians are slightly smaller, but still substantial. The large discrepancies imply that it is, in all likelihood, more than just mutual funds' selection of portfolio firms that drives comparability. Figure 1 gives some insight into potential drivers. Figure 1a (1b) depicts the quarterly mean of $compacct_{ikt}$ ($compcf_{ikt}$), separately for mutual funds and analyst portfolios. The figure illustrates that the comparability proxies are subject to economic fluctuations, the dot-com bubble in the early years and the financial crisis later on. Recall that I construct the proxies from data of the preceding four years so that economic downturns have a lagged and sticky impact.

Descriptives also underline that return-based and cash flow-based comparability proxies capture distinct aspects of the mapping of economic events into financial statements. The correlation between $compacct_{ikt}$ and $compcf_{ikt}$ is positive, but modest. Panel A of Table 2 contains the Pearson (Spearman) correlation, 0.417 (0.414), for the subset of observations with available comparability values from the two models, without differentiating mutual fund or analyst holdings.[46] Moreover, the impact of economic fluctuations appears to be more pronounced for return-based than for cash flow-based comparability. While $compacct_{ikt}$ in Figure 1a varies considerably over time, $compcf_{ikt}$ in Figure 1b seems smoother and almost suggests an increasing trend over the sample period.

To contrast comparability in mutual fund and analyst portfolios after accounting for economic drivers and model-specific factors, I estimate the model:

$$\begin{aligned} comp\star_{ikt} =& \beta_1 mf_{kt} + \beta_2 ln(peers)_{kt} + \beta_3 ln(mv)_{it} + \beta_4 btm_{it} + \beta_5 roa_{it} + \beta_6 \star^1 \_volatility_{it} \\ &+ \beta_7 \star^2 \_volatility_{it} + \beta_8 ear\_predictability_{it} + \beta_9 volume_{it} \\ &+ \beta_{10} spread_{it} + \beta CONT_{kt} + \beta FE_i + \beta FE_t + \epsilon_{ikt} \end{aligned}$$

where $comp\star_{ikt}$ is either $compacct_{ikt}$ or $compcf_{ikt}$. $mf_{kt}$, the variable of interest, is an indicator variable equal to one if the portfolio is a mutual fund portfolio. The remaining variables control for the economic, mechanical drivers of comparability [De Franco et al., 2011]. I rely on the same variables to incorporate preferences of mutual funds and analysts. My interest does not

[46] The overlap between the samples for $compacct_{ikt}$ and $compcf_{ikt}$ is small since the underlying samples are random draws. Matching all observations yields a similar Pearson correlation coefficient (0.402) and a smaller Spearman correlation coefficient (0.289) (N = 4,132,224).

lie in measuring comparability effects that come from pursuing certain investment strategies or being an expert in a particular type of firm. For instance, mutual funds tend to invest in larger firms [e.g., Daniel, Grinblatt, Titman and Wermers, 1997; Falkenstein, 1996]. At the same time, firm size is positively associated with the comparability proxies.

$ln(peers)_{kt}$ is the natural logarithm of the number of portfolio peers. Although I average over an identical number of peers for each firm (ten), the overall number of peers in the portfolio influences comparability values. The more peers, the more options for high firm-pair comparability. In my dataset, the number of peers is strongly associated with comparability. Disregarding other factors, $ln(peers)_{kt}$ has a statistically significant Pearson (Spearman) correlation of 0.229 (0.417) with $compacct_{ikt}$ and of 0.387 (0.489) with $compcf_{ikt}$ (untabulated).

At the firm level, I account for firm size, $ln(mv)_{it}$, the natural logarithm of the market value of equity; book-to-market, $btm_{it}$, the ratio of the book to market value of equity; return on assets, $roa_{it}$, income before extraordinary items over lagged total assets; a proxy for earnings predictability, $ear\_predictability_{it}$, which is the adjusted $R^2$ from a regression of quarterly earnings on lagged quarterly earnings over up to 16 quarters, a period that is identical to the estimation period of comparability; trading volume, $volume_{it}$, the quarterly mean of daily share turnover; and bid-ask spreads, $spread_{it}$, the quarterly mean of daily bid-ask spreads. I further include proxies for the volatility of the summary measures from which I compute comparability. In specifications with return-based comparability, $\star^1\_volatility_{it}$ stands for earnings volatility, $ear\_volatility_{it}$, the standard deviation of quarterly earnings scaled by lagged total assets over up to 16 quarters; and $\star^2\_volatility_{it}$ represents return volatility, $ret\_volatility_{it}$, the standard deviation of monthly returns over up to 48 months. In specifications with cash flow-based comparability, $\star^1\_volatility_{it}$ and $\star^2\_volatility_{it}$ stand for accrual volatility, $acc\_volatility_{it}$, and cash flow volatility, $cfo\_volatility_{it}$, the standard deviation of quarterly accruals and quarterly operating cash flow over up to 16 quarters, respectively.

At the portfolio level, I also control for the mean of each of the firm characteristics listed above, subsumed by $CONT_{kt}$. I proceed in this fashion to incorporate overall investment preferences of mutual funds or fields of specialization of analysts. Finally, I use firm and quarter-year fixed effects, $FE_i$ and $FE_t$. Standard errors are two-way clustered by portfolio and quarter-year [Gow, Ormazabal and Taylor, 2010; Petersen, 2009]. Appendix A.1 lists the variable definitions.

I carefully construct the samples for the analysis. The idea is to exploit the variation across portfolios. That is, I intend to assess differences in comparability between mutual fund and analyst portfolios, ideally for the same firm. Before adjustments, samples are large. There are 4,822,229 (4,220,438) observations for $compacct_{ikt}$ ($compcf_{ikt}$). I work with subsets of observa-

tions, at first the *random* sample. In particular, I randomly select at most ten holdings from each portfolio in each quarter. To ensure that I measure the variation in peers at the firm level, I require that firms appear in at least one mutual fund and one analyst portfolio in each quarter.[47]

In addition, I match holdings of the same firm from mutual fund to analyst portfolios, in the *matched* sample. Mutual fund and analyst portfolios differ in part to a high degree. Notably, mutual fund portfolios have on average more peers. The mean of $ln(peers)_{kt}$ equals 3.681 (3.590) in mutual fund and 2.551 (2.540) in analyst portfolios for $compacct_{ikt}$ ($compcf_{ikt}$); the mean of the raw number of peers equals 44 (40) in mutual fund and 13 (13) in analyst portfolios (untabulated). Therefore, I match holdings to be able to assess differences between portfolios that are similar apart from belonging to mutual funds or analysts. I apply the CEM algorithm [Iacus et al., 2012]. CEM sorts observations into strata which are formed from the predictors, the variables the researcher matches on, and produces exact matches based on these strata. CEM's advantage is that the researcher can, ex ante, fix the allowed range for the predictors. I match on the raw number of peers as well as portfolio means of size, book-to-market and bid-ask spreads (size, accrual and cash flow volatility) for $compacct_{ikt}$ ($compcf_{ikt}$). I choose the portfolio characteristics in line with their (mechanical) relation with the comparability proxies. $compacct_{ikt}$ responds noticeably to capital market factors whereas $compcf_{ikt}$ is more affected by fluctuations in the underlying fundamentals. I resort to CEM's binning algorithm for the portfolio means and explicitly set intervals for the number of peers so that the maximum difference for matched observations cannot be greater than ten. Following this approach, I obtain samples in which each firm is in at least one mutual fund and one analyst portfolio in each quarter. I do not restrict the number of matches, but use all, with corresponding weights in the regression analysis. The resulting matched samples are small: there are 10,925 (12,693) observations for $compacct_{ikt}$ ($compcf_{ikt}$), emphasizing that differences in the random sample may be stark. Since I match on portfolio characteristics, I only include controls for firm characteristics in the regression model. Appendix A.2.1 explains the matching method graphically and shows balance before and after matching.

Table 4 depicts the results. The coefficient of the variable of interest, $mf_{kt}$, is positive throughout, but statistically weaker for cash flow-based comparability. It amounts to 0.163 (0.013) for $compacct_{ikt}$ ($compcf_{ikt}$) in the *random* sample and to 0.083 (0.067) in the *matched*

[47] I consider this approach the cleanest. Nevertheless, it does not hurt the research design (much) to leave all holdings and firms in the sample since I examine within-firm variation. Results for the samples with all observations are similar and stronger. Specifically, the coefficient of $mf_{kt}$ equals 0.205 (0.043) for $compacct_{ikt}$ ($compcf_{ikt}$). For both comparability proxies, the coefficient is statistically significant at the 0.01 level.

sample. Aside from the coefficient for $compcf_{ikt}$ in the random sample, estimates are statistically significant and differences in comparability seem economically meaningful. For example, in the matched sample, the difference translates into 9% (8%) of $compacct_{ikt}$ ($compcf_{ikt}$), evaluated at the mean for mutual fund portfolios.

The majority of the coefficients of the control variables is significant and displays the predicted sign. For instance, the estimate of $ln(peers)_{kt}$ is large, illustrating the positive, mechanical relation between the number of portfolio peers and comparability. The estimates of the volatility variables, $ear\_volatility_{ikt}$ and $ret\_volatility_{ikt}$ for return-based comparability and $acc\_volatility_{ikt}$ and $cfo\_volatility_{ikt}$ for cash flow-based comparability, have a likewise large, but mostly negative impact. More frequent changes in the fundamentals hamper the estimation of the firm-specific accounting system in that the coefficients of the accounting function can only in part reflect the association between the summary measures. Accounting comparability may well be lower, consistent with the negative estimates. The control variables, furthermore, hint at some of the differences in the two comparability models. $compacct_{ikt}$ is negatively related to $spread_{it}$, the average bid-ask spread, which is often interpreted as information asymmetry or illiquidity. I derive the proxy from returns, subject to changes in liquidity and likely responsible for the estimate [e.g., Amihud and Mendelson, 1986, 1989]. $compcf_{ikt}$, on the other hand, reveals similar patterns in the random sample, but weaker relations in the matched sample, which are not completely explained by the matching variables.

Taken together, the analysis shows that accounting comparability is higher in mutual fund than in analyst portfolios. In itself, the analyst benchmark is an ambitious one since analysts face similar incentives as mutual funds to compose their portfolios of firms with comparable accounting. In practice, the analyst benchmark suffers from shortcomings in the empirical implementation since analyst portfolios exhibit characteristics distinct from mutual fund portfolios. However, even after limiting the analysis to portfolios that share relevant characteristics, accounting comparability remains higher in portfolios of mutual funds. In the next sections, I focus on mutual fund portfolios and the role of comparability therein.

## 4.2 Mutual Fund Types

The accounting literature argues that different types of institutional investors make use of accounting information to a different extent [e.g., Bushee and Noe, 2000; D'Souza, Ramesh and Shen, 2010]. Among mutual funds, researchers regularly separate actively managed from index funds [e.g., Aggarwal, Klapper and Wysocki, 2005; Florou and Pope, 2012]. While the former have investment discretion and hence face the need to gather information about potential portfolio firms, the latter merely replicate the composition of an index. Among actively managed funds,

researchers further make a distinction between between stock pickers and those that primarily bet on factors [e.g., Cremers and Petajisto, 2009; Fama, 1972]. Although the implementation of both investment approaches requires information about firms, it should be the stock picker who has a detailed interest in financial statements. Notably, it should be the stock picker who has an interest in the comparability of accounting information of existing and potential portfolio firms. Thus, in this section, I investigate whether accounting comparability varies predictably with the type of mutual funds. In doing so, I exploit the classification into nonindexers and indexers along with the degree of stock picking or stock selection.

I classify mutual funds into nonindexers and indexers with the indicator variable $indexer_{kt}$, which is equal to one if the fund is an indexing investor. I measure the degree of stock picking following the notion of the active share, proposed by Cremers and Petajisto [2009]. Mutual funds normally have a benchmark index against which their performance is assessed. Cremers and Petajisto introduce a metric that captures the share of the portfolio that deviates from the benchmark index. The larger this share, the 'active share', the more active the mutual fund with respect to stock picking. I work with a simplified version by solely studying the share of firms that is not in the index, i.e., I count the overall number of firms in a given portfolio and define the variable $active_{kt}$ as the fraction of firms that is not included in the index.[48] I obtain the fraction by going back to the raw portfolios without imposing any restriction on portfolio weights. Besides, I do not work with the actual benchmarks, but the S&P 500 for all mutual funds. Prior studies similarly concentrate on deviations from the S&P 500 and show that it is a common benchmark with index weights that frequently extend to other indexes [Lines, 2016; Wermers, 2003]. I calculate the share of firms that are not included in the S&P 500 each quarter so that the variable $active_{kt}$ is a time-varying construct that is increasing in the extent of stock picking. Panel B of Table 1 demonstrates that all mutual funds have some portfolio firms that are not in the S&P 500; the minimum of $active_{kt}$ over all fund-quarters amounts to 22%. Moreover, the typical portfolio firm is not a member of the index; the mean (median) of $active_{kt}$ amounts to 89% (94%).

Table 3 presents correlations of $active_{kt}$ with other fund characteristics and return-based (Panel A) as well as cash flow-based accounting comparability (Panel B) for holding-quarters; Pearson (Spearman) correlations in the lower (upper) triangle. Descriptives refer to the random samples which I describe below. $active_{kt}$ is negatively correlated with $indexer_{kt}$. The Pearson

[48] Cremers and Petajisto [2009] are more accurate in that they do not only use the appearance of the firm, but the difference in index and portfolio weights. I implicitly assume that index weights are equal across index members and that, once a firm is in both the index and the portfolio, index and portfolio weights are the same.

(Spearman) correlation is -0.410 (-0.170) in the $compacct_{ikt}$ sample and -0.411 (-0.172) in the $compcf_{ikt}$ sample. On the one hand, this validates the stock-picking measure. By definition, indexers should have more portfolio firms that are index members and consequently a lower value of $active_{kt}$. On the other hand, the magnitude of the correlations underlines that the stock-picking concept is more refined than the partition into nonindexers and indexers. Despite being actively managed, nonindexers have a benchmark against which their investors evaluate their performance and can follow the benchmark composition, also yielding low values of $active_{kt}$. As another validation check, $active_{kt}$ is negatively correlated with portfolio size, $ln(tna)_{kt}$, the natural logarithm of total net assets, and $ln(peers)_{kt}$, as earlier defined. Active funds with stock-picking skills are usually smaller [e.g., Cremers and Petajisto, 2009; Kacpercyzk, Nieuwerburgh and Veldkamp, 2014]. Lastly, $active_{kt}$ is negatively correlated with accounting comparability. The Pearson (Spearman) correlation is -0.063 (-0.232) for $compacct_{ikt}$ and -0.097 (-0.216) for $compcf_{ikt}$. Albeit inconsistent with my expectation, the negative association might be the consequence of other fund characteristics. More active funds are smaller and smaller funds have lower comparability, due to, for instance, a smaller number of peers. Summary statistics of the comparability proxies in Panel B of Table 2, reported separately for nonindexers and indexers, support the negative association. Mean and median of return-based and cash flow-based comparability are lower for nonindexers than for indexers.

To examine the association after accounting for other fund characteristics, I estimate the model:

$$\begin{aligned} comp\star_{ikt} =& \beta_1 active_{kt} + \beta_2 ln(tna)_{kt} + \beta_3 ln(peers)_{kt} \\ &+ \beta CONT_{it} + \beta CONT_{kt} + \beta FE_{IO} + \beta FE_i + \beta FE_t + \epsilon_{ikt} \end{aligned}$$

where $comp\star_{ikt}$ is either $compacct_{ikt}$ or $compcf_{ikt}$. The variable of interest is $active_{kt}$. Controls for other fund characteristics are $ln(tna)_{kt}$, $ln(peers)_{kt}$ and fixed effects for investment objectives, $FE_{IO}$. Data requirements for these variables result in the loss of some mutual funds that are part of the analysis in the previous section. I continue to control for firm characteristics, $CONT_{it}$, portfolio means of firm characteristics, $CONT_{kt}$, and firm and quarter-year fixed effects, $FE_i$ and $FE_t$, as explained in Section 4.1. I cluster standard errors by portfolio and quarter-year.

I construct the samples for the analysis in a corresponding manner to Section 4.1. I still aim for measuring the variation in the comparability proxies that is driven by changing peer firms and to somewhat control for variation within and across firms. This time, I assess differences between stock pickers and other mutual funds. In the regression model, the emphasis lies on the

continuous variable $active_{kt}$. In the sample design, it shifts to the indicator variable $indexer_{kt}$. For each firm and quarter, I choose one observation from a nonindexer and one observation from an indexer portfolio. Initially, there are 4,707,398 (4,287,165) observations for $compacct_{ikt}$ ($compcf_{ikt}$). I first randomly select at most ten holdings from each portfolio in each quarter and require that firms appear in at least one nonindexer and one indexer portfolio (*random* sample).[49] I then match holdings of the same firm from nonindexer to indexer portfolios (*matched* sample). The idea is to constrast accounting comparability in portfolios that are roughly the same apart from the degree of active stock picking. I match, with the CEM algorithm, on the same variables as before: number of peers and portfolio means of size, book-to-market and bid-ask spreads (number of peers and portfolio means of size, accrual and cash flow volatility) for $compacct_{ikt}$ ($compcf_{ikt}$). I specify that the maximum difference in the number of peers for matched observations is not to be greater than ten. Compared to the previous section, there is more overlap in characteristics. Matched samples are larger, with 542,950 (502,510) observations for $compacct_{ikt}$ ($compcf_{ikt}$). Since I match on them, I exclude the number of peers and the portfolio means of firm characteristics in the regression model. Appendix A.2.2 provides further details.

Table 5 depicts the results. The coefficient of the variable of interest, $active_{kt}$, is positive in three out of four specifications and again weaker for cash flow-based comparability. It equals 0.055 (-0.008) for $compacct_{ikt}$ ($compcf_{ikt}$) in the *random* sample and 0.071 (0.032) in the *matched* sample. Except for the coefficient for $compcf_{ikt}$ in the random sample, estimates are statistically significant and suggest an economically mildly significant relation. For example, in the matched sample, a one-standard-deviation increase in $active_{kt}$ leads to an increase of 5% (2%) of $compacct_{ikt}$ ($compcf_{ikt}$), evaluated at its mean. Coefficients of the control variables mostly have the predicted sign. $ln(tna)_{kt}$ loads negatively only for $compacct_{ikt}$ in the random sample, but might be subsumed by the other variables. $ln(peers)_{kt}$, if included, loads positively.

In sum, the findings show that, within mutual fund portfolios, accounting comparability varies with the type of mutual funds. In particular, comparability is higher in portfolios of funds that seem to be stock pickers and thus represent the subset for which similar accounting information should matter most. In the following sections, I restrict mutual funds to nonindexing investors, i.e., I drop all mutual funds that are classified as indexing investor in at least one quarter.

[49] Again, I rerun the regression model on the samples with all observations and obtain similar results. Specifically, the coefficient of $active_{kt}$ equals 0.032 (-0.007) for $compacct_{ikt}$ ($compcf_{ikt}$) and is statistically significant at the 0.01 level for $compacct_{ikt}$.

## 4.3 Portfolio Selection

So far, I present evidence consistent with high levels of accounting comparability in portfolios of mutual funds. Implicitly, I assume that these levels arise because mutual funds act on comparability. Prior literature indicates that, in some settings, accounting information can affect investor behavior. Properties of the accounting information environment are associated with holdings of institutional investors [e.g., Bowen et al., 2014; Bushee and Noe, 2000]. Similarities in the accounting information environment are related to investments by mutual funds [e.g., DeFond et al., 2011; Yu and Wahid, 2014]. The underlying rationale in these studies pertains to portfolio selection. Investors select firms whose accounting coincides with their preferences and expertise. Yet, prior literature also demonstrates that a variety of firm and stock characteristics can explain mutual funds' holdings [e.g., Chen, Jegadeesh and Wermers, 2000; Falkenstein, 1996]. In what follows, I directly test whether mutual funds attach weight to accounting comparability in making portfolio decisions, in addition to other firm and stock attributes.

I start with compiling the sample. I concentrate on the quarter in which mutual funds newly include firms into their portfolios. I define firms as newly included if they are not part of the portfolio in the preceding eight quarters. To avoid misclassifications due to the minimum portfolio weight, I determine inclusion status in the raw portfolios. To stick to the important portion of portfolios, I continue to keep only holdings whose portfolio weight is equal to or exceeds 0.5% (in the quarter of inclusion). I then need firms that are *not* newly included to be able to gauge the impact of accounting comparability on portfolio selection. This analysis, therefore, exploits the variation across firms, i.e., differences in comparability on account of differences in firm-specific accounting systems. To reduce variation in factors aside from comparability, I match firms. Precisely, for each firm and each inclusion, I select another firm, a potential candidate for inclusion, that closely resembles the newly included firm and is neither newly included nor an existing member of the (raw) portfolio. I apply the CEM algorithm and match on size, book-to-market, bid-ask spread, earnings and return volatility (size, book-to-market, bid-ask spread, accrual and cash flow volatility) for $compacct_{ikt}$ ($compcf_{ikt}$). I choose the predictors to incorporate investment preferences of mutual funds (size, book-to-market and spread) as well as mechanical drivers of comparability that should be less informative about the mapping of economic events into financial statements (the volatility variables). In matching, I select exactly one potential candidate for each newly included firm; in the case of several possible matches, the algorithm randomly picks one. Appendix A.2.3 provides further details on the matching.

Resulting samples still contain holdings, i.e., firms can be simultaneously included in different portfolios. For each inclusion, there are two observations (holding-quarters), the newly included

and the potential candidate firm, assigned to the same portfolio and the same quarter. Since I need two previous years to establish inclusion, the sample period begins in the first quarter of 2003. The sample for $compacct_{ikt}$ ($compcf_{ikt}$) comprises 632,046 (578,942) observations.

For the potential candidate firm, I measure accounting comparability *as if* the mutual fund included the firm. That is, I use the same portfolio peers for the potential candidate as for the newly included firm. By this means, I hold portfolios constant and can focus on differences in the two firms.

Panel C of Table 2 displays summary statistics of accounting comparability, separately for newly included (*inclusion*) and potential candidate holdings. Raw differences are in line with a positive association between comparability and portfolio inclusion, but they are not large. $compacct_{ikt}$ ($compcf_{ikt}$) has a mean of -0.507 (-0.533) for newly included and of -0.556 (-0.561) for potential candidate holdings. Medians are even closer.

In the analysis, I take into account alternative determinants of portfolio selection and estimate the model:

$$\begin{aligned} inclusion_{ikt} =& \beta_1 comp\star_{ikt} + \beta_2 ln(peers)_{kt} + \beta_3 ln(mv)_{it} + \beta_4 btm_{it} + \beta_5 retq1_{it} + \beta_6 rety1_{it} \\ &+ \beta_7 ret\_sd_{it} + \beta_8 volume_{it} + \beta_9 spread_{it} + \beta FE_{IO} + \beta FE_i + \beta FE_t + \epsilon_{ikt} \end{aligned}$$

where $inclusion_{ikt}$ is an indicator variable equal to one if the firm is newly included. By construction, $inclusion_{ikt}$ is equal to one for one half of the sample. The variabel of interest, $comp\star_{ikt}$, is either $compacct_{ikt}$ or $compcf_{ikt}$. The remaining variables proxy for investment preferences and characteristics of mutual funds [e.g., Falkenstein, 1996; Ke and Petroni, 2004; Shive and Yun, 2013]. I control for the number of peers in the portfolio, $ln(peers)_{kt}$; firm size, $ln(mv)_{it}$; book-to-market, $btm_{it}$; two factors for momentum trading, $retq1_{it}$ and $rety1_{it}$, the buy-and-hold return from quarter $(t-4)$ to quarter $(t-1)$ and from quarter $(t-1)$ to quarter $t$, respectively; return volatility, $ret\_sd_{it}$, the quarterly standard deviation of daily returns; trading volume, $volume_{it}$; and bid-ask spreads, $spread_{it}$.[50] I use fixed effects for investment objectives of mutual funds, firms and quarters, $FE_{IO}$, $FE_i$ and $FE_t$. I cluster standard errors by firm and quarter-year.[51] Appendix A.1 lists the variable definitions.

Column 1 and 4 in Table 6 depict the results. The coefficients of the comparability variables are positive and statistically significant. The estimate of $compacct_{ikt}$ ($compcf_{ikt}$) equals 0.020

[50] The list of control variables is not exhaustive. I extend the above set of controls with, for example, index membership, analyst following, return on assets, dividend yield, leverage ratio, earnings-to-price ratio and sales growth [e.g., Bushee and Noe, 2000; Florou and Pope, 2012]. Results are very similar.

[51] Alternatively, I cluster standard errors by portfolio and quarter-year. Results are very similar.

(0.058). Economically, the impact is modest. A one-standard-deviation increase in $compacct_{ikt}$ ($compcf_{ikt}$) corresponds to an increase in the probability of inclusion of around 2% (3%). Thus, accounting comparability does seem to play a role in investment decisions, even if only a secondary one.

Coefficients of the control variables largely display predicted signs. Note that I already match on some of them, which likely weakens their associations. For instance, results are consistent with mutual funds trading on return momentum, $retq1_{it}$ and $rety1_{it}$ load positively; and with mutual funds selecting more liquid and frequently traded stocks, $volume_{it}$ is also positive.

Moreover, I make use of cross-sectional variation in firms. I expect that mutual funds value accounting comparability especially if firms have stable and predictable information environments. My argument rests in part on the increasing usage of automated data processing techniques. Reliance on standardized computer routines makes information necessary that is equally standardized. Hence, in selecting portfolio firms, mutual funds might pay particular attention to comparability when firms are not subject to idiosyncratic shocks or other disruptive events. Along these lines, D'Souza et al. [2010] show that dissemination speed of financial statement information from Compustat is positively related to trading volume and negatively related to idiosyncratic risk. In their interpretation, dissemination speed is the outcome of institutional demand for (standardized) accounting information and trading volume and idiosyncratic risk characterize this demand.

I use bid-ask spreads and return volatility, each measured over the same time period as accounting comparability, to split up observations. Specifically, I create the indicator variable $SPREAD_{it}$, which is equal to one if the mean of monthly bid-ask spreads is above the median; and the indicator variable $RET_{it}$, which is equal to one if the standard deviation of monthly returns is above the median. Firms with high bid-ask spreads or high return volatility should experience more unexpected events and should have less predictable information environments. To test for a differential impact of accounting comparability on portfolio selection for these firms, I interact $compacct_{ikt}$ and $compcf_{ikt}$ in the above equation with $SPREAD_{it}$ and $RET_{it}$. Columns 2–3 and 5–6 in Table 6 depict the results. Coefficients of the interaction terms are negative in three out of four specifications. $SPREAD_{it}$ weakens the impact of return-based comparability, but has no effect for cash flow-based comparability. $RET_{it}$ weakens the impact of both comparability proxies.

To sum up, high levels of accounting comparability can arise through the channel of portfolio selection. Findings are more pronounced for firms with more predictable and less uncertain information environments, offering insight into possible forms of analysis that are based on

similar financial statements.

## 4.4 Selection vs. Influence

The previous section provides evidence in line with mutual funds selecting firms conditional on the comparability of their accounting with other portfolio firms. In addition to the selection of ex ante comparable firms, mutual funds can also influence accounting outcomes ex post, after including firms in their portfolio. Influence, in this context, does not always need to take the form of direct communication. Mutual funds can express their preferences through various channels, such as portfolio choices or public disclosures. Extant research supports both channels, suggesting that it regularly is a combination. Institutional investors in general and mutual funds in particular seem to respond to firms' disclosures and, at the same time, to induce changes in disclosure patterns [e.g., Boone and White, 2015; Fang et al., 2015; Yu and Wahid, 2014]. In the final analysis, I examine how accounting comparability evolves throughout the holding period to understand whether mutual funds encourage portfolio firms to align their accounting practices with their preferences.

I test the relative importance of the selection against the influence channel by studying accounting comparability around the inclusion of firms into portfolios. I begin with constructing the samples and focus on firms that are included. The idea is to mainly exploit within-firm variation, i.e., changes in comparability that are driven by changes in the firm-specific accounting system over time. As in Section 4.3, I define newly included firms as firms that are not part of the portfolio in the preceding eight quarters and collect the information from the raw portfolios of nonindexer funds. In constrast to Section 4.3, I track firms through the holding period and relax the requirement on portfolio weights. Investors often build up positions successively so that stocks can enter the portfolio before the portfolio weight reaches 0.5%. Likewise, investors can reduce positions successively so that stocks remain in the portfolio after the portfolio weight falls below 0.5%. I only require that the portfolio weight is larger than or equal to 0.5% in at least one quarter in the period after inclusion.

In the next step, I restrict samples to observations with a holding period of 16 quarters or more. I intend to assess the possibility of mutual funds inducing changes in the firm-specific accounting system of their portfolio firms. Therefore, I need estimates of the accounting system that are derived exclusively from the period during which the mutual fund has a stake in the firm. In any given quarter, these estimates, $\hat{\alpha}_{0,i}$ and $\hat{\alpha}_{1,i}$ ($\hat{\beta}_{0,i}$ and $\hat{\beta}_{1,i}$) for the return-based (cash flow-based) accounting function (see Section 2), depend on the prior 16 quarters. The restriction thus ensures that there is at least one quarter in which the estimates originate from

accounting practices carried out in their entirety during the holding period, quarter 16. Until quarter 16, the weight of these accounting practices continuously increases. If mutual funds resort to the influence channel, accounting comparability will increase similarly. I balance the samples around inclusion, i.e., for each holding, there are 24 quarters with 8 (16) quarters pre (post) to inclusion. Due to the overall sample period, inclusions into portfolios occur between the first quarter of 2003 and the fourth quarter of 2011. The resulting sample for $compacct_{ikt}$ ($compcf_{ikt}$) contains 76,944 (64,824) observations, 24 observations for each of the 3,206 (2,701) holdings.

For the period before inclusion, I measure accounting comparability *as if* the holding was already part of the portfolio. That is, I create pseudo portfolios consisting of the actual portfolios of mutual funds and the firms that they subsequently add. In this vein, I can observe how comparable, in accounting terms, firms are with the portfolio choices of mutual funds, at every single point in time.

Samples are not representative of the earlier samples, but should favor the incidence of influence activities. Panel D of Table 2 shows that $compacct_{ikt}$ ($compcf_{ikt}$) has a mean of -0.392 (-0.411) and a median of -0.214 (-0.332), clearly exceeding the respective values in the other samples. Besides, the minimum holding period of 16 quarters does not generalize to the typical holding. In the underlying samples, the holding period in a mutual fund, nonindexer portfolio amounts on average to 3 quarters, with a median of 2 quarters (untabulated).[52] However, a longer investment horizon appears to make the (in)direct communication of preferences more likely, strengthening the influence channel [McCahery et al., 2016].

Figure 2 illustrates changes in accounting comparability around inclusion in portfolios. Figure 2a (2b) depicts the quarterly mean of $compacct_{ikt}$ ($compcf_{ikt}$) in each of the 24 quarters. I redefine quarters relative to the quarter of inclusion and set the quarter of inclusion equal to zero. In the quarters leading up to inclusion, comparability increases noticeably for either proxy, complementing the findings of Section 4.3. Accounting comparability is not persistently high, but rises steadily before inclusion and hence hints at a timing component in portfolio selection. The firm fits into the mutual fund portfolio at the point in time when it is added. The magnitude of the increase suggests that the simultaneous selection of other, more similar peers contributes to the level of comparability at inclusion. In the quarters following inclusion, comparability increases moderately for the return-based proxy and does not markedly change

[52] Mean and median are lower bounds. They refer to the underlying samples which I use in Section 4.2 and in which I require a portfolio weight of at least 0.5%.

for the cash flow-based proxy. Consequently, the influence channel does not seem to dominate at first glance. To account for alternative factors impacting comparability, I estimate the model:

$$\begin{aligned} comp\star_{ikt} =& \beta_1 pre1_{ikt} + \beta_2 post1_{ikt} + \beta_3 post2_{ikt} + \beta_4 post3_{ikt} + \beta_5 post4_{ikt} \\ &+ \beta_6 ln(peers)_{kt} + \beta CONT_{it} + \beta CONT_{kt} + \beta FE_{IO} + \beta FE_i + \beta FE_t + \epsilon_{ikt} \end{aligned}$$

where $comp\star_{ikt}$ is either $compacct_{ikt}$ or $compcf_{ikt}$. The variables of interest are indicator variables for subperiods relative to the second year before portfolio inclusion, which serves as reference period. $pre1_{ikt}$ is equal to one starting in the year, i.e., the four quarters, preceding inclusion. $post1_{ikt}$ ($post2_{ikt}$) [$post3_{ikt}$] {$post4_{ikt}$} is equal to one starting in the first (second) [third] {fourth} year subsequent to inclusion. All indicator variables are equal to one until the end of the sample period, i.e., they measure incremental changes in the respective subperiods. I control for the number of peers in the portfolio, $ln(peers)_{kt}$, as well as firm characteristics and portfolio means of firm characteristics, $CONT_{it}$ and $CONT_{kt}$, as explained in Section 4.1. I include fixed effects for investment objectives of mutual funds, firms and quarters, $FE_{IO}$, $FE_i$ and $FE_t$. I cluster standard errors by portfolio and quarter-year.[53]

Table 7 displays the results. In the quarters before and around inclusion, comparability increases, reinforcing the impression from the graphical representation. The coefficient of $pre1_{ikt}$ amounts to 0.021 (0.011) and the coefficient of $post1_{ikt}$ amounts to 0.013 (0.011) for $compacct_{ikt}$ ($compcf_{ikt}$). The coefficients are statistically significant and translate into small, but noticeable changes in comparability. For instance, the estimate of $pre1_{ikt}$ implies an increase of 5% (2%) of $compacct_{ikt}$ ($compcf_{ikt}$), evaluated at its mean for the preceding four quarters. In the following quarters, comparability does not change further or even reverts to earlier levels, again confirming the graphical representation. The coefficients of $post2_{ikt}$ and $post3_{ikt}$ are not different from zero for both comparability proxies. The coefficient of $post4_{ikt}$ is not different from zero for $compacct_{ikt}$ and indicates a statistically significant decrease for $compcf_{ikt}$; it amounts to -0.008.

Accordingly, the findings are consistent with accounting comparability being one factor that mutual funds consider in selecting portfolio firms. Section 4.3 demonstrates that the probability of inclusion into portfolios is increasing in comparability. This section emphasizes that the level of comparability is particularly high at the time of inclusion. On the other hand, the findings are not consistent with accounting comparability being a factor that mutual funds target after initial investment, even for a subset of holdings for which influence activities are more likely. As

[53] Alternatively, I cluster standard errors by (i) portfolio and quarter relative to inclusion; (ii) firm and quarter-year; and (iii) firm and quarter relative to inclusion. Results are similar.

it is still a very limited subset, I do not intend to generalize any results. For this subset, though, selection appears to be relatively more prevalent.

# 5 Conclusion

In this study, I investigate whether mutual funds' portfolio choices reflect a preference for accounting comparability. I argue that investors with diversified portfolios need standardized financial statement information to identify investment opportunities and conduct ongoing evaluations of their holdings. I examine portfolios of mutual funds, investors directly engaging in data analysis and stock selection. I measure accounting comparability following De Franco et al. [2011] and extend the concept in two respects. In addition to stock returns, I resort to operating cash flow to proxy for economic events. In determining the relevant set of peers, I adopt an investor perspective and assess comparability with portfolio peers. Two portfolio firms are comparable in their accounting if they produce similar financial statements for the same economic events. I thus derive accounting comparability for holdings and first study variation in its level. I can show that comparability is higher in mutual fund portfolios than in analyst portfolios, which I construct from analysts' coverage decisions. Moreover, comparability is higher in portfolios of actively managed mutual funds whose portfolio composition suggests that they carefully single out individual stocks. While these findings support the notion that mutual funds benefit from accounting comparability, they do not inform about a potential impact on their investment behavior. In the second step, I hence analyze comparability around the inclusion of firms into portfolios. I can show that the probability of inclusion is moderately increasing in accounting comparability. The relation is stronger for firms with more liquid stocks and less volatile returns, in line with a preference for similar accounting information when firms' information environment allows for automated data processing. Furthermore, comparability is increasing until and around inclusion, but does not change much afterwards, consistent with investment selection conditional on comparability and inconsistent with the subsequent influence of accounting practices.

Altogether, the study provides evidence in line with mutual funds relying on comparable accounting information for the selection of portfolio firms. The findings underline that accounting comparability is not a primary concern. Yet, it does affect mutual funds' investment decisions. As the use of automated data processing techniques is becoming more common, the relevance of similar financial statements might increase as well. As BlackRock's CEO Larry Fink puts it: "A year from now, [...] they [people]'ll just have different skill sets. It'll be more data analysis, there will be more model producers. We are not saying active is dead. We think active can be

more alive, just using different insights".[54]

[54]The quote comes from an interview with Larry Fink in April 2017, availabe at: https://www.bloomberg.com/features/2017-blackrock-larry-fink-interview/.

# A Appendix

## A.1 Variable Definitions

| Variable | Description |
|---|---|
| **Holding-level variables** | |
| $compacct_{ikt}$ | Mean of the ten highest firm-pair comparability values for firm $i$ in portfolio $k$, based on the mapping of returns into earnings; multiplied by 100. |
| $compcf_{ikt}$ | Mean of the ten highest firm-pair comparability values for firm $i$ in portfolio $k$, based on the mapping of operating cash flows into accruals; multiplied by 100. |
| $inclusion_{ikt}$ | Indicator variable equal to one if firm $i$ is newly included in portfolio $k$ in quarter $t$. |
| $pre1_{ikt}$ | Indicator variable equal to one starting in the year, i.e., the four quarters, preceding inclusion of firm $i$ in portfolio $k$. |
| $post1_{ikt}$ ($post2_{ikt}$) [$post3_{ikt}$] {$post4_{ikt}$} | Indicator variable equal to one starting in the first (second) [third] {fourth} year subsequent to inclusion of firm $i$ in portfolio $k$. |
| **Firm-level variables** | |
| $ln(mv)_{it}$ | Natural logarithm of the market value of equity. |
| $btm_{it}$ | Ratio of book to market value of equity. |
| $roa_{it}$ | Income before extraordinary items over lagged total assets; multiplied by 100. |
| $ear\_volatility_{it}$ | Standard deviation of quarterly earnings scaled by lagged total assets over the 16 previous quarters; multiplied by 100. I require a minimum of 14 quarters. |
| $ret\_volatility_{it}$ | Standard deviation of monthly returns over the previous 48 months; multiplied by 100. I require a minimum of 36 months. |
| $acc\_volatility_{it}$ | Standard deviation of quarterly accruals scaled by lagged total assets over the 16 previous quarters; multiplied by 100. I require a minimum of 14 quarters. Accruals are the difference between income before extraordinary items and operating cash flow. |
| $cfo\_volatility_{it}$ | Standard deviation of quarterly operating cash flow scaled by lagged total assets over the previous 16 quarters; multiplied by 100. I require a minimum of 14 quarters. |

*Continued on next page*

*Continued from previous page*

| Variable | Description |
|---|---|
| $ear\_predictability_{it}$ | Adjusted $R^2$ from a regression of quarterly earnings on lagged quarterly earnings over 16 quarters. I require a minimum of 14 quarters. Earnings are scaled by lagged market value of equity. |
| $volume_{it}$ | Quarterly mean of daily turnover; multiplied by 100. I require a minimum of 32 observations. Turnover is daily volume over the number of shares outstanding. |
| $spread_{it}$ | Quarterly mean of daily bid-ask spreads; multiplied by 100. I require a minimum of 32 observations. The bid-ask spread is the difference between the ask and bid price over the midpoint. |
| $SPREAD_{it}$ | Indicator variable equal to one if the mean of monthly bid-ask spreads over the previous 48 months is above the median in quarter $t$. I require a minimum of 36 months. |
| $RET_{it}$ | Indicator variable equal to one if $ret\_volatility_{it}$ is above the median in quarter $t$. |
| $retq1_{it}$ | Buy-and-hold return from quarter $(t-1)$ to quarter $t$. |
| $rety1_{it}$ | Buy-and-hold return from quarter $(t-4)$ to quarter $(t-1)$. |
| $ret\_sd_{it}$ | Quarterly standard deviation of daily returns. I require a minimum of 32 observations. |
| **Portfolio-level variables** | |
| $active_{kt}$ | The share of firms in portfolio $k$ that is not in the S&P 500. |
| $indexer_{kt}$ | Indicator variable equal to one if portfolio $k$ is classified as indexing investor in quarter $t$. |
| $ln(tna)_{kt}$ | Natural logarithm of total net assets in portfolio $k$. |
| $peers_{kt}$ [$ln(peers)_{kt}$] | [Natural logarithm of the] number of portfolio peers. |
| $mf_{kt}$ | Indicator variable equal to one (zero) if portfolio $k$ is a mutual fund (analyst) portfolio. |

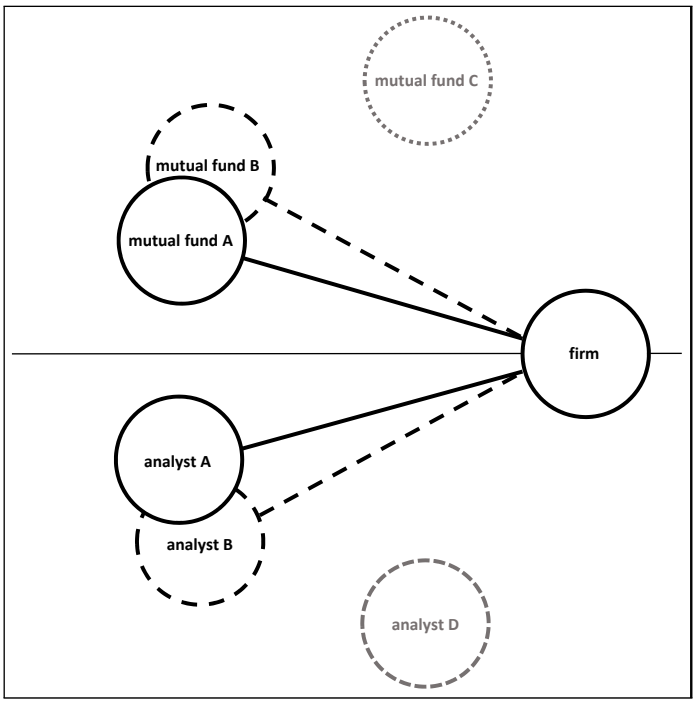
*Notes*: $i$, $k$ and $t$ denote firm, portfolio and quarter-year, respectively.

## A.2 Matching Approach

### A.2.1 Mutual Funds vs. Analysts

The idea is to contrast comparability in mutual fund and analyst portfolios. To be able to focus on the variation across portfolios, I use portfolios for the *same* firm. To control for other portfolio characteristics, I match on predictor variables.

The figure shows the matching approach for a single firm. The firm is in three mutual fund portfolios (of type A, B and C) and three analyst portfolios (of type A, B and D). After matching, I retain only holdings, i.e., firm-portfolio combinations, for which the portfolio type appears among both mutual funds and analysts. That is, I retain holdings with portfolios of type A and B and drop holdings with portfolios of type C and D.



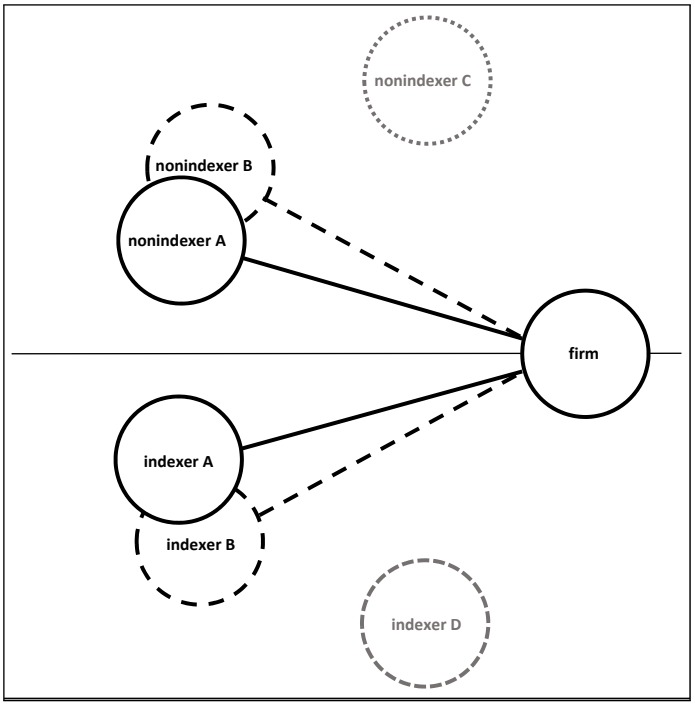

I characterize the type with the following predictors for $compacct_{ikt}$ ($compcf_{ikt}$): number of peers and portfolio means of size, book-to-market and bid-ask spreads (size, accrual and cash flow volatility). The table shows balance before and after matching. Note that CEM allows to specify the tolerable range of the predictors ex ante so that balance checking is not necessary.

| | $\boldsymbol{compacct_{ikt}}$ | | | | | $\boldsymbol{compcf_{ikt}}$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Mean before** | | **Mean after** | | | **Mean before** | | **Mean after** | |
| **Variable** | *fund* | *analyst* | *fund* | *analyst* | **Variable** | *fund* | *analyst* | *fund* | *analyst* |
| $peers_{kt}$ | 50.338 | 14.240 | 19.180 | 15.800 | $peers_{kt}$ | 46.337 | 14.016 | 19.375 | 15.976 |
| $mean\_ln(mv)_{kt}$ | 9.286 | 8.222 | 8.704 | 8.785 | $mean\_ln(mv)_{kt}$ | 9.268 | 8.212 | 8.696 | 8.733 |
| $mean\_btm_{kt}$ | 0.439 | 0.533 | 0.475 | 0.483 | $mean\_cfo\_volatility_{kt}$ | 2.169 | 2.128 | 2.043 | 1.885 |
| $mean\_spread_{kt}$ | 0.138 | 0.163 | 0.145 | 0.139 | $mean\_acc\_volatility_{kt}$ | 1.976 | 1.871 | 1.912 | 1.760 |

### A.2.2 Mutual Fund Types

The idea is to contrast comparability in nonindexer and indexer portfolios. To be able to focus on the variation across portfolios, I use portfolios for the *same* firm. To control for other portfolio characteristics, I match on predictor variables.

The figure shows the matching approach for a single firm. The firm is in three nonindexer portfolios (of type A, B and C) and three indexer portfolios (of type A, B and D). After matching, I retain only holdings, i.e., firm-portfolio combinations, for which the portfolio type appears among both nonindexers and indexers. That is, I retain holdings with portfolios of type A and B and drop holdings with portfolios of type C and D.



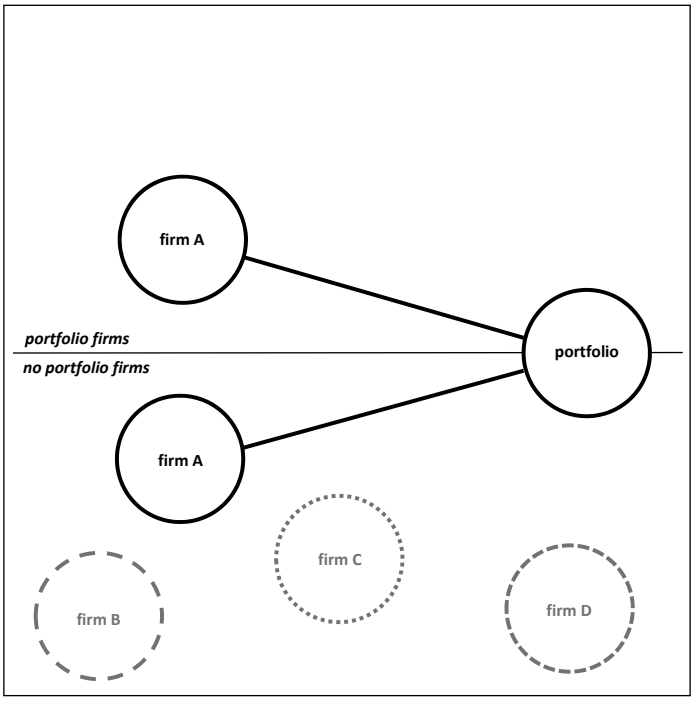

I characterize the type with the following predictors for $compacct_{ikt}$ ($compcf_{ikt}$): number of peers and portfolio means of size, book-to-market and bid-ask spreads (size, accrual and cash flow volatility). The table shows balance before and after matching. Note that CEM allows to specify the tolerable range of the predictors ex ante so that balance checking is not necessary.

| | $\boldsymbol{compacct_{ikt}}$ | | | | | $\boldsymbol{compcf_{ikt}}$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Mean before** | | **Mean after** | | | **Mean before** | | **Mean after** | |
| **Variable** | *non* | *indexer* | *non* | *indexer* | **Variable** | *non* | *indexer* | *non* | *indexer* |
| $peers_{kt}$ | 50.382 | 50.960 | 46.885 | 45.995 | $peers_{kt}$ | 46.156 | 47.178 | 42.575 | 41.895 |
| $mean\_ln(mv)_{kt}$ | 9.117 | 9.616 | 10.345 | 10.390 | $mean\_ln(mv)_{kt}$ | 9.094 | 9.594 | 10.287 | 10.351 |
| $mean\_btm_{kt}$ | 0.442 | 0.444 | 0.419 | 0.414 | $mean\_cfo\_volatility_{kt}$ | 2.237 | 2.060 | 1.918 | 1.910 |
| $mean\_spread_{kt}$ | 0.151 | 0.116 | 0.116 | 0.120 | $mean\_acc\_volatility_{kt}$ | 2.039 | 1.883 | 1.747 | 1.741 |

### A.2.3 Portfolio Selection

The idea is to compare newly included firms with potential candidates for inclusion. To be able to focus on the variation across firms, I create holdings for the *same* portfolio. To control for differences in firm characteristics, I match on predictor variables.

The figure shows the matching approach for a single portfolio. The newly included firm is of type A. Among firms that are neither newly included nor existing members of the portfolio, there are firms of type A, B, C and D. I retain only pairs of holdings, i.e., pairs of firm-portfolio combinations, for which I can match a firm that is of the same type as the newly included firm. That is, I keep both holdings with firms of type A and drop potential holdings with firms of type B, C and D.




I characterize the type with the following predictors for $compacct_{ikt}$ ($compcf_{ikt}$): size, book-to-market, bid-ask spread, earnings and return volatility (size, book-to-market, bid-ask spread, accrual and cash flow volatility). The table shows the mean of the predictors after matching.

| | $\boldsymbol{compacct_{ikt}}$ | | | $\boldsymbol{compcf_{ikt}}$ | |
|---|---|---|---|---|---|
| | **Mean after** | | | **Mean after** | |
| **Variable** | *inclusion* | | **Variable** | *inclusion* | |
| $ln(mv)_{it}$ | 8.437 | 8.392 | $ln(mv)_{it}$ | 8.437 | 8.392 |
| $btm_{it}$ | 0.430 | 0.433 | $btm_{it}$ | 0.430 | 0.433 |
| $spread_{it}$ | 0.113 | 0.121 | $spread_{it}$ | 0.113 | 0.121 |
| $ear\_volatility_{it}$ | 1.203 | 1.192 | $acc\_volatility_{it}$ | 1.203 | 1.192 |
| $ret\_volatility_{it}$ | 10.249 | 10.189 | $cfo\_volatility_{it}$ | 10.249 | 10.189 |

## References


Agarwal, V., Mullally, K. A., Tang, Y. and Yang, B. [2015], 'Mandatory portfolio disclosure, stock liquidity, and mutual fund performance', *The Journal of Finance* **70**(6), 2733–2776.

Aggarwal, R., Klapper, L. and Wysocki, P. D. [2005], 'Portfolio preferences of foreign institutional investors', *Journal of Banking & Finance* **29**(12), 2919–2946.

Amihud, Y. and Mendelson, H. [1986], 'Asset pricing and the bid-ask spread', *Journal of Financial Economics* **17**(2), 223–249.

Amihud, Y. and Mendelson, H. [1989], 'The effects of beta, bid-ask spread, residual risk, and size on stock returns', *The Journal of Finance* **44**(2), 479–486.

Bhojraj, S. and Lee, C. M. C. [2002], 'Who is my peer? A valuation-based approach to the selection of comparable firms', *Journal of Accounting Research* **40**(2), 407–439.

Bhojraj, S., Lee, C. M. C. and Oler, D. K. [2003], 'What's my line? A comparison of industry classification schemes for capital market research', *Journal of Accounting Research* **41**(5), 745–774.

Boone, A. L. and White, J. T. [2015], 'The effect of institutional ownership on firm transparency and information production', *Journal of Financial Economics* **117**(3), 508–533.

Bowen, R. M., Rajgopal, S. and Venkatachalam, M. [2014], 'Is Warren Buffett's commentary on accounting, governance, and investing practices reflected in the investment decisions and subsequent influence of Berkshire Hathaway?', *The Accounting Review* **89**(5), 1609–1644.

Bradshaw, M. T., Bushee, B. J. and Miller, G. S. [2004], 'Accounting choice, home bias, and U.S. investment in non-U.S. firms', *Journal of Accounting Research* **42**(5), 795–841.

Bradshaw, M. T., Miller, G. S. and Serafeim, G. [2009], 'Accounting method heterogeneity and analysts' forecasts', *Working Paper*. Available at https://www.researchgate.net/profile/-Mark_Bradshaw2.

Brochet, F., Jagolinzer, A. D. and Riedl, E. J. [2013], 'Mandatory IFRS adoption and financial statement comparability', *Contemporary Accounting Research* **30**(4), 1373–1400.

Bushee, B. J. and Noe, C. F. [2000], 'Corporate disclosure practices, institutional investors, and stock return volatility', *Journal of Accounting Research* **38**(Supplement: Studies on Accounting Information and the Economics of the Firm), 171–202.

Cascino, S. and Gassen, J. [2015], 'What drives the comparability effect of mandatory IFRS adoption?', *Review of Accounting Studies* **20**(1), 242–282.

Chen, H.-L., Jegadeesh, N. and Wermers, R. [2000], 'The value of active mutual fund management: An examination of the stockholdings and trades of fund managers', *Journal of Financial and Quantitative Analysis* **35**(3), 343–368.

Covrig, V. M., DeFond, M. L. and Hung, M. [2007], 'Home bias, foreign mutual fund holdings, and the voluntary adoption of International Accounting Standards', *Journal of Accounting Research* **45**(1), 41–70.

Cremers, K. J. M. and Petajisto, A. [2009], 'How active is your fund manager? A new measure that predicts performance', *The Review of Financial Studies* **22**(9), 3329–3365.

Daniel, K., Grinblatt, M., Titman, S. and Wermers, R. [1997], 'Measuring mutual fund performance with characteristic-based benchmarks', *The Journal of Finance* **52**(3), 1035–1058.

De Franco, G., Kothari, S. P. and Verdi, R. S. [2011], 'The benefits of financial statement comparability', *Journal of Accounting Research* **49**(4), 895–931.

DeFond, M., Hu, X., Hung, M. and Li, S. [2011], 'The impact of mandatory IFRS adoption on foreign mutual fund ownership: The role of comparability', *Journal of Accounting and Economics* **51**(3), 240–258.

D'Souza, J. M., Ramesh, K. and Shen, M. [2010], 'The interdependence between institutional ownership and information dissemination by data aggregators', *The Accounting Review* **85**(1), 159–193.

Dye, R. A. and Sunder, S. [2001], 'Why not allow FASB and IASB standards to compete in the U.S.?', *Accounting Horizons* **15**(3), 257–271.

Falkenstein, E. G. [1996], 'Preferences for stock characteristics as revealed by mutual fund portfolio holdings', *The Journal of Finance* **51**(1), 111–135.

Fama, E. F. [1972], 'Components of investment performance', *The Journal of Finance* **27**(3), 551–567.

Fang, V. W., Maffett, M. and Zhang, B. [2015], 'Foreign institutional ownership and the global convergence of financial reporting practices', *Journal of Accounting Research* **53**(3), 593–631.

FASB [2010], 'Statement of financial accounting concepts no. 8', Conceptual Framework for Financial Reporting.

Fich, E. M., Harford, J. and Tran, A. L. [2015], 'Motivated monitors: The importance of institutional investors' portfolio weights', *Journal of Financial Economics* **118**(1), 21–48.

Florou, A. and Pope, P. F. [2012], 'Mandatory IFRS adoption and institutional investment decisions', *The Accounting Review* **87**(6), 1993–2025.

Gassen, J., Skaife, H. A. and Veenman, D. [2016], 'Illiquidity and the measurement of stock price synchronicity', *Working Paper*. Available at https://papers.ssrn.com/sol3/papers.cfm?-abstract_id=2405465.

Gow, I. D., Ormazabal, G. and Taylor, D. J. [2010], 'Correcting for cross-sectional and time-series dependence in accounting research', *The Accounting Review* **85**(2), 483–512.

Gu, Z. and Chen, T. [2004], 'Analysts' treatment of nonrecurring items in street earnings', *Journal of Accounting and Economics* **38**(1), 129–170.

Hoberg, G. and Phillips, G. [2016], 'Text-based network industries and endogenous product differentiation', *Journal of Political Economy* **124**(5), 1423–1465.

Horton, J., Serafeim, G. and Serafeim, I. [2013], 'Does mandatory IFRS adoption improve the information environment?', *Contemporary Accounting Research* **30**(1), 388–423.

Iacus, S. M., King, G. and Porro, G. [2012], 'Causal inference without balance checking: Coarsened exact matching', *Political Analysis* **20**(1), 1–24.

ICI [2017], '*Investment company fact book*', Investment Company Institute (ICI).

Jung, M. J. [2013], 'Investor overlap and diffusion of disclosure practices', *Review of Accounting Studies* **18**(1), 167–206.

Kacpercyzk, M., Nieuwerburgh, S. V. and Veldkamp, L. [2014], 'Time-varying fund manager skill', *The Journal of Finance* **69**(4), 1455–1484.

Ke, B. and Petroni, K. [2004], 'How informed are actively trading institutional investors? Evidence from their trading behavior before a break in a string of consecutive earnings increases', *Journal of Accounting Research* **42**(5), 895–927.

Kim, S., Kraft, P. and Ryan, S. G. [2013], 'Financial statement comparability and credit risk', *Review of Accounting Studies* **18**(3), 783–823.

Leuz, C. and Wysocki, P. D. [2016], 'The economics of disclosure and financial reporting regulation: Evidence and suggestions for future research', *Journal of Accounting Research* **54**(2), 525–622.

Lines, A. [2016], 'Do institutional incentives distort asset prices?', *Working Paper*. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2873739.

Ma, L., Tang, Y. and Gómez, J.-P. [2016], 'Portfolio manager compensation and mutual fund performance', *Working Paper*. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2024027.

McCahery, J. A., Sautner, Z. and Starks, L. T. [2016], 'Behind the scenes: The corporate governance preferences of institutional investors', *The Journal of Finance* **71**(6), 2905–2932.

Neel, M. [2017], 'Accounting comparability and economic outcomes of mandatory IFRS adoption', *Contemporary Accounting Research* **34**(1), 658–690.

Petersen, M. A. [2009], 'Estimating standard errors in finance panel data sets: Comparing approaches', *The Review of Financial Studies* **22**(1), 435–480.

Ramnath, S. [2002], 'Investor and analyst reactions to earnings announcements of related firms: An empirical analysis', *Journal of Accounting Research* **40**(5), 1351–1376.

Shive, S. and Yun, H. [2013], 'Are mutual funds sitting ducks?', *Journal of Financial Economics* **107**(1), 220–237.

Wermers, R. [2000], 'Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses', *The Journal of Finance* **55**(4), 1655–1695.

Wermers, R. [2003], 'Are mutual fund shareholders compensated for active management "bets"?', *Working Paper*. Available at https://www.researchgate.net/profile/Russ_Wermers.

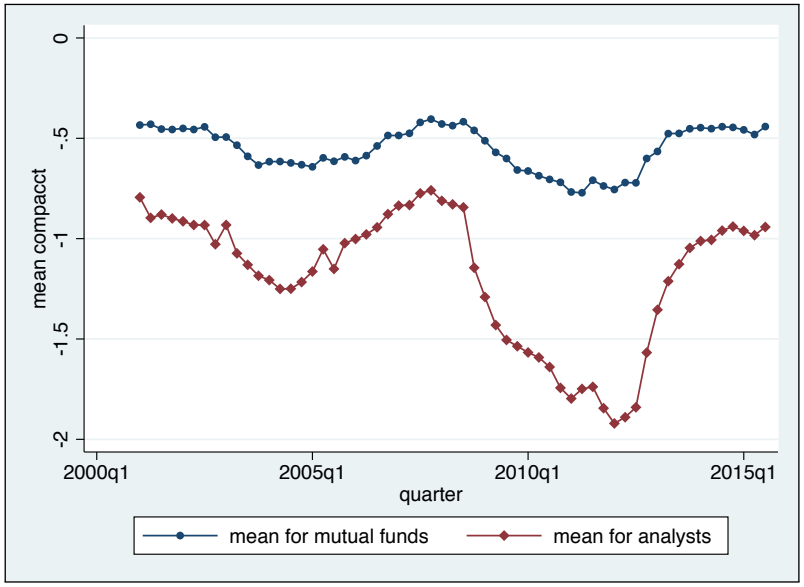
Young, S. and Zeng, Y. [2015], 'Accounting comparability and the accuracy of peer-based valuation models', *The Accounting Review* **90**(6), 2571–2601.

Yu, G. and Wahid, A. S. [2014], 'Accounting standards and international portfolio holdings', *The Accounting Review* **89**(5), 1895–1930.

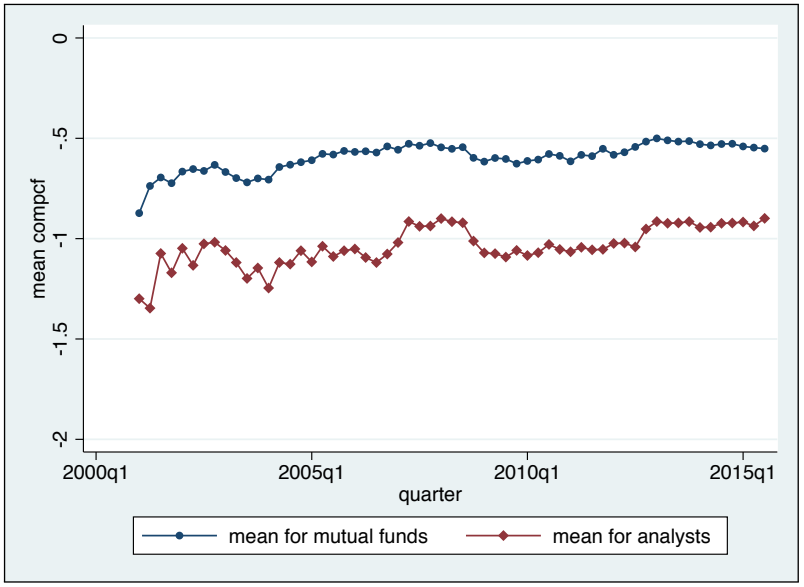
# Figures

Figure 1: **Mutual Funds vs. Analysts**

(a) $\boldsymbol{compacct_{ikt}}$



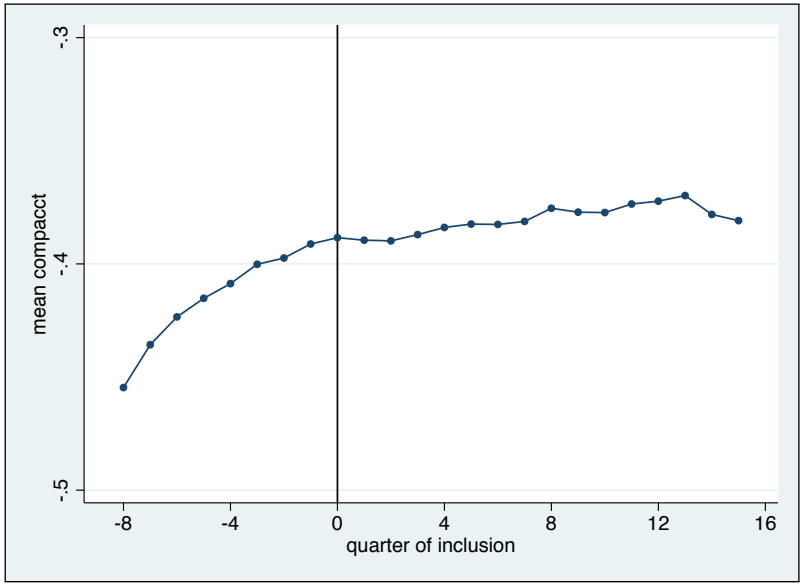

(b) $\boldsymbol{compcf_{ikt}}$



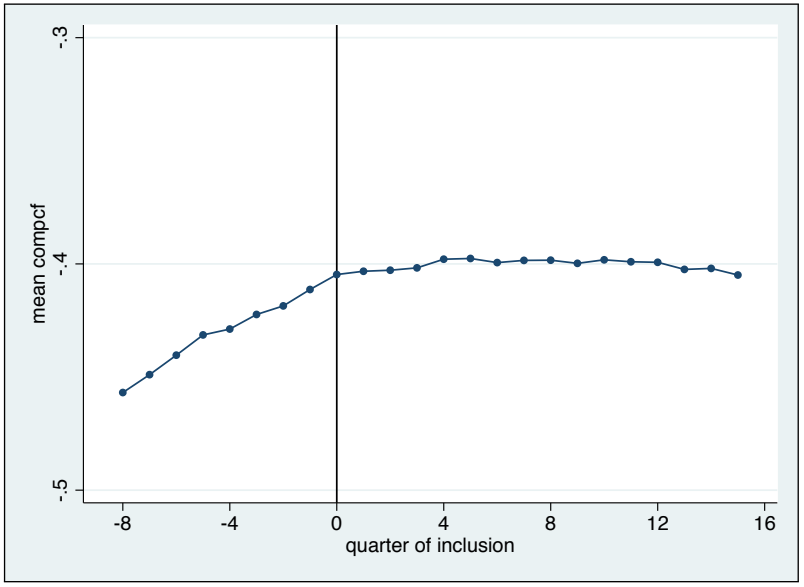

*Notes:* The figure displays the quarterly mean of accounting comparability for mutual fund (blue line) and analyst portfolios (red line). $i$, $k$ and $t$ denote firm, portfolio and quarter-year, respectively. $compacct_{ikt}$ ($compcf_{ikt}$) is the mean of the ten highest firm-pair comparability values for firm $i$ in portfolio $k$, based on the mapping of returns into earnings (cash flows into accruals). The sample for Figure 1a (Figure 1b) comprises 1,387,908 (1,330,219) observations, 1,136,150 (1,119,375) for mutual fund and 251,758 (210,844) for analyst portfolios.

Figure 2: **Selection vs. Influence**

(a) $\boldsymbol{compacct_{ikt}}$




(b) $\boldsymbol{compcf_{ikt}}$




*Notes:* The figure displays the quarterly mean of accounting comparability around the inclusion in mutual fund portfolios. $i$, $k$ and $t$ denote firm, portfolio and quarter-year, respectively. Before taking the mean, I limit mutual funds to nonindexers. I further restrict observations to holdings that are not part of portfolio $k$ for eight quarters, are then newly included and remain in portfolio $k$ for at least four years. For each holding, the sample contains 24 observations, two years before and four years following inclusion. Quarter 0 is the quarter of inclusion. $compacct_{ikt}$ ($compcf_{ikt}$) is the mean of the ten highest firm-pair comparability values for firm $i$ in portfolio $k$, based on the mapping of returns into earnings (cash flows into accruals). The sample for Figure 2a (Figure 2b) comprises 3,206 (2,701) holdings.

# Tables

Table 1: **Summary Statistics**

***Sample***

Panel A: ***Firms***

| **Variable** | **N** | **Mean** | **SD** | **Min** | **25$^{th}$** | **Med** | **75$^{th}$** | **Max** |
|---|---|---|---|---|---|---|---|---|
| $ln(mv)_{it}$ | 92,370 | 7.173 | 1.614 | 2.583 | 5.951 | 6.991 | 8.171 | 13.290 |
| $btm_{it}$ | 92,370 | 0.563 | 0.392 | −0.498 | 0.300 | 0.485 | 0.731 | 2.887 |
| $roa_{it}$ | 92,370 | 0.812 | 2.817 | −22.728 | 0.210 | 0.930 | 2.016 | 9.166 |
| $ear\_volatility_{it}$ | 92,370 | 1.798 | 2.473 | 0.024 | 0.457 | 0.970 | 2.126 | 31.253 |
| $ret\_volatility_{it}$ | 92,370 | 12.109 | 5.845 | 2.462 | 7.940 | 10.873 | 14.905 | 47.235 |
| $acc\_volatility_{it}$ | 83,095 | 2.770 | 2.094 | 0.180 | 1.342 | 2.236 | 3.589 | 28.194 |
| $cfo\_volatility_{it}$ | 83,095 | 2.464 | 1.801 | 0.166 | 1.264 | 2.016 | 3.171 | 22.258 |
| $ear\_predictability_{it}$ | 92,370 | 0.066 | 0.175 | −0.071 | −0.061 | −0.013 | 0.129 | 0.744 |
| $volume_{it}$ | 92,370 | 0.853 | 0.716 | 0.007 | 0.383 | 0.666 | 1.101 | 12.521 |
| $spread_{it}$ | 92,370 | 0.304 | 0.477 | −0.041 | 0.070 | 0.140 | 0.322 | 10.319 |

Panel B: ***Mutual Funds***

| **Variable** | **N** | **Mean** | **SD** | **Min** | **25$^{th}$** | **Med** | **75$^{th}$** | **Max** |
|---|---|---|---|---|---|---|---|---|
| $active_{kt}$ | 117,696 | 0.893 | 0.161 | 0.220 | 0.889 | 0.941 | 0.978 | 1.000 |
| $indexer_{kt}$ | 117,696 | 0.148 | | | | | | |
| $ln(tna)_{kt}$ | 117,696 | 5.345 | 2.052 | −2.303 | 3.991 | 5.352 | 6.746 | 12.909 |
| $ln(peers)_{kt}$ | 117,696 | 3.676 | 0.452 | 2.303 | 3.401 | 3.738 | 4.007 | 4.949 |
| $peers_{kt}$ | 117,696 | 43.329 | 17.568 | 10.000 | 30.000 | 42.000 | 55.000 | 141.000 |

*Notes:* The table displays summary statistics for the universe of firms and mutual funds from which I construct the samples for the analyses. Panel A depicts summary statistics for firms and Panel B for mutual funds. $i$, $k$ and $t$ denote firm, portfolio and quarter-year, respectively. I limit firm-quarters to quarters in which the firm appears in at least one mutual fund portfolio. $ln(mv)_{it}$ is the natural logarithm of the market value of equity. $btm_{it}$ is the ratio of book to market value of equity. $roa_{it}$ is income before extraordinary items over lagged total assets. $ear\_volatility_{it}$ is the standard deviation of quarterly earnings scaled by lagged total assets over 16 quarters. $ret\_volatility_{it}$ is the standard deviation of monthly returns over 48 months. $acc\_volatility_{it}$ is the standard deviation of quarterly accruals scaled by lagged total assets over 16 quarters. $cfo\_volatility_{it}$ is the standard deviation of quarterly operating cash flow scaled by lagged total assets over 16 quarters. $ear\_predictability_{it}$ is the adjusted $R^2$ from a regression of quarterly earnings on lagged quarterly earnings over 16 quarters. $volume_{it}$ is the quarterly mean of daily turnover. $spread_{it}$ is the quarterly mean of daily bid-ask spreads. I multiply $roa_{it}$, $ear\_volatility_{it}$, $ret\_volatility_{it}$, $acc\_volatility_{it}$, $cfo\_volatility_{it}$, $volume_{it}$ and $spread_{it}$ by 100. $active_{kt}$ is the share of firms in portfolio $k$ that is not in the S&P 500. $indexer_{kt}$ is an indicator variable equal to one if portfolio $k$ is classified as indexing investor in quarter $t$. $ln(tna)_{kt}$ is the natural logarithm of total net assets in portfolio $k$. $peers_{kt}$ [$ln(peers)_{kt}$] is [the natural logarithm of] the number of portfolio peers. There are 3,162 (4,219) unique firms (mutual funds) in Panel A (B).

Table 2: **Summary Statistics**

***Accounting Comparability***

Panel A: ***Mutual Funds vs. Analysts***

| **Variable** | **N** | **Mean** | **SD** | **Min** | **25$^{th}$** | **Med** | **75$^{th}$** | **Max** |
|---|---|---|---|---|---|---|---|---|
| $compacct_{ikt}$ | 1,387,908 | −0.680 | 1.172 | −25.618 | −0.665 | −0.305 | −0.170 | −0.028 |
| *mutual funds* | 1,136,150 | −0.557 | 1.034 | −24.719 | −0.506 | −0.256 | −0.154 | −0.028 |
| *analysts* | 251,758 | −1.234 | 1.539 | −25.618 | −1.414 | −0.730 | −0.404 | −0.048 |
| $compcf_{ikt}$ | 1,330,219 | −0.659 | 0.641 | −18.495 | −0.799 | −0.476 | −0.294 | −0.008 |
| *mutual funds* | 1,119,375 | −0.592 | 0.563 | −15.092 | −0.705 | −0.438 | −0.284 | −0.008 |
| *analysts* | 210,844 | −1.015 | 0.874 | −18.495 | −1.307 | −0.827 | −0.484 | −0.013 |

Correlation between $compacct_{ikt}$ and $compcf_{ikt}$

| | | |
|---|---|---|
| Pearson | 490,062 | $0.417^{***}$ |
| Spearman | 490,062 | $0.414^{***}$ |

Panel B: ***Mutual Fund Types***

| **Variable** | **N** | **Mean** | **SD** | **Min** | **25$^{th}$** | **Med** | **75$^{th}$** | **Max** |
|---|---|---|---|---|---|---|---|---|
| $compacct_{ikt}$ | 956,225 | −0.514 | 0.965 | −21.911 | −0.472 | −0.246 | −0.150 | −0.025 |
| *nonindexers* | 788,229 | −0.515 | 0.960 | −21.735 | −0.477 | −0.248 | −0.151 | −0.025 |
| *indexers* | 167,996 | −0.510 | 0.987 | −21.911 | −0.450 | −0.237 | −0.142 | −0.030 |
| $compcf_{ikt}$ | 963,288 | −0.580 | 0.535 | −19.693 | −0.694 | −0.435 | −0.282 | −0.012 |
| *nonindexers* | 795,790 | −0.587 | 0.537 | −13.375 | −0.705 | −0.439 | −0.283 | −0.015 |
| *indexers* | 167,498 | −0.546 | 0.524 | −19.693 | −0.643 | −0.417 | −0.277 | −0.012 |

Correlation between $compacct_{ikt}$ and $compcf_{ikt}$

| | | |
|---|---|---|
| Pearson | 259,255 | $0.356^{***}$ |
| Spearman | 259,255 | $0.304^{***}$ |

Panel C: ***Portfolio Selection***

| Variable | | N | Mean | SD | Min | 25$^{th}$ | Med | 75$^{th}$ | Max |
|---|---|---|---|---|---|---|---|---|---|
| $compacct_{ikt}$ | | 632,046 | −0.531 | 0.875 | −19.845 | −0.514 | −0.261 | −0.157 | −0.029 |
| | *inclusion* | 316,023 | −0.507 | 0.822 | −19.845 | −0.495 | −0.254 | −0.155 | −0.029 |
| | | 316,023 | −0.556 | 0.925 | −19.379 | −0.533 | −0.268 | −0.160 | −0.031 |
| $compcf_{ikt}$ | | 578,942 | −0.547 | 0.539 | −15.304 | −0.628 | −0.393 | −0.263 | −0.008 |
| | *inclusion* | 289,471 | −0.533 | 0.499 | −14.418 | −0.618 | −0.389 | −0.261 | −0.008 |
| | | 289,471 | −0.561 | 0.576 | −15.304 | −0.640 | −0.398 | −0.264 | −0.012 |

Panel D: ***Selection vs. Influence***

| Variable | N | Mean | SD | Min | 25$^{th}$ | Med | 75$^{th}$ | Max |
|---|---|---|---|---|---|---|---|---|
| $compacct_{ikt}$ | 76,944 | −0.392 | 0.597 | −11.912 | −0.396 | −0.214 | −0.135 | −0.028 |
| $compcf_{ikt}$ | 64,824 | −0.411 | 0.297 | −4.802 | −0.493 | −0.332 | −0.226 | −0.029 |

*Notes:* The table displays summary statistics of accounting comparability for the different samples of the analyses. Panel A refers to the sample in Table 4; Panel B to the sample in Table 5; Panel C to the sample in Table 6; and Panel D to the sample in Table 7. $i$, $k$ and $t$ denote firm, portfolio and quarter-year, respectively. $compacct_{ikt}$ ($compcf_{ikt}$) is the mean of the ten highest firm-pair comparability values for firm $i$ in portfolio $k$, based on the mapping of returns into earnings (cash flows into accruals). Panel A and Panel B depict the correlations between $compacct_{ikt}$ and $compcf_{ikt}$ for the observations that appear in both samples. The overlap is small since the underlying samples are random draws. $^{*}$, $^{**}$ and $^{***}$ indicate significance at the 0.1, 0.05 and 0.01 level, respectively.

Table 3: **Correlations**

***Mutual Fund Types***

Panel A: $\boldsymbol{compacct_{ikt}}$

| | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** |
|---|---|---|---|---|---|
| **(1)** $compacct_{ikt}$ | 1.000 | $-0.232^{***}$ | $0.024^{***}$ | $0.060^{***}$ | $0.256^{***}$ |
| **(2)** $active_{kt}$ | $-0.063^{***}$ | 1.000 | $-0.170^{***}$ | $-0.222^{***}$ | $-0.357^{***}$ |
| **(3)** $indexer_{kt}$ | $0.002^{*}$ | $-0.410^{***}$ | 1.000 | $-0.001$ | $0.011^{***}$ |
| **(4)** $ln(tna)_{kt}$ | $0.029^{***}$ | $-0.200^{***}$ | $0.010^{***}$ | 1.000 | $0.164^{***}$ |
| **(5)** $ln(peers)_{kt}$ | $0.094^{***}$ | $-0.140^{***}$ | $0.035^{***}$ | $0.171^{***}$ | 1.000 |

Panel B: $\boldsymbol{compcf_{ikt}}$

| | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** |
|---|---|---|---|---|---|
| **(1)** $compcf_{ikt}$ | 1.000 | $-0.216^{***}$ | $0.031^{***}$ | $0.096^{***}$ | $0.469^{***}$ |
| **(2)** $active_{kt}$ | $-0.097^{***}$ | 1.000 | $-0.172^{***}$ | $-0.220^{***}$ | $-0.324^{***}$ |
| **(3)** $indexer_{kt}$ | $0.029^{***}$ | $-0.411^{***}$ | 1.000 | 0.000 | $0.011^{***}$ |
| **(4)** $ln(tna)_{kt}$ | $0.076^{***}$ | $-0.199^{***}$ | $0.011^{***}$ | 1.000 | $0.165^{***}$ |
| **(5)** $ln(peers)_{kt}$ | $0.354^{***}$ | $-0.114^{***}$ | $0.033^{***}$ | $0.170^{***}$ | 1.000 |

*Notes:* The table displays Pearson (Spearman) correlations for accounting comparability and mutual fund characteristics in the lower (upper) triangle. Panel A depicts correlations for accounting comparability based on the mapping of returns into earnings, $compacct_{ikt}$. Panel B depicts correlations for accounting comparability based on the mapping of cash flows into accruals, $compcf_{ikt}$. $i$, $k$ and $t$ denote firm, portfolio and quarter-year, respectively. $compacct_{ikt}$ ($compcf_{ikt}$) is the mean of the ten highest firm-pair comparability values for firm $i$ in portfolio $k$. $active_{kt}$ is the share of firms in portfolio $k$ that is not in the S&P 500. $indexer_{kt}$ is an indicator variable equal to one if portfolio $k$ is classified as indexing investor in quarter $t$. $ln(tna)_{kt}$ is the natural logarithm of total net assets in portfolio $k$. $ln(peers)_{kt}$ is the natural logarithm of the number of portfolio peers. The number of observations is 956,225 (963,288) in Panel A (B). $^{*}$, $^{**}$ and $^{***}$ indicate significance at the 0.1, 0.05 and 0.01 level, respectively.

Table 4: **Mutual Funds vs. Analysts**

| | $compacct_{ikt}$ | | $compcf_{ikt}$ | |
|---|---|---|---|---|
| *Sample* | *random* | *matched* | *random* | *matched* |
| $mf_{kt}$ | 0.163*** | 0.083*** | 0.013 | 0.067*** |
| | (0.016) | (0.018) | (0.009) | (0.012) |
| $ln(peers)_{kt}$ | 0.280*** | | 0.403*** | |
| | (0.008) | | (0.005) | |
| $ln(mv)_{it}$ | 0.255*** | 0.257*** | 0.029*** | 0.024 |
| | (0.023) | (0.061) | (0.008) | (0.028) |
| $btm_{it}$ | 0.313*** | 0.544*** | 0.087*** | 0.070 |
| | (0.053) | (0.115) | (0.013) | (0.048) |
| $roa_{it}$ | 0.014*** | 0.043*** | 0.010*** | 0.005 |
| | (0.004) | (0.012) | (0.002) | (0.005) |
| $\star^1\_volatility_{it}$ | −0.237*** | −0.269*** | −0.116*** | −0.213*** |
| | (0.011) | (0.035) | (0.010) | (0.016) |
| $\star^2\_volatility_{it}$ | −0.089*** | −0.155*** | −0.024*** | 0.064** |
| | (0.005) | (0.013) | (0.008) | (0.028) |
| $ear\_predictability_{it}$ | 0.021 | 0.115 | −0.082*** | −0.083* |
| | (0.023) | (0.148) | (0.015) | (0.044) |
| $volume_{it}$ | 0.004 | −0.003 | −0.014* | −0.004 |
| | (0.014) | (0.050) | (0.007) | (0.021) |
| $spread_{it}$ | −0.171*** | −0.491*** | 0.021 | 0.017 |
| | (0.063) | (0.179) | (0.028) | (0.034) |
| Portfolio means of firm characteristics in portfolio $k$ | yes | no | yes | no |
| Firm and quarter-year fixed effects | yes | yes | yes | yes |
| adj. $R^2$ | 0.625 | 0.764 | 0.698 | 0.772 |
| adj. $R^2$ within | 0.335 | 0.349 | 0.371 | 0.120 |
| $N$ | 1,387,908 | 10,925 | 1,330,219 | 12,693 |
| Holdings | 475,658 | 7,436 | 449,431 | 8,282 |
| Firms | 2,807 | 1,219 | 2,613 | 1,174 |
| Mutual funds | 4,218 | 940 | 4,190 | 1,083 |
| Analysts | 2,189 | 1,084 | 1,909 | 1,000 |

*Notes:* The table displays estimates for the association between accounting comparability and membership in mutual fund portfolios, relative to analyst portfolios. $i$, $k$ and $t$ denote firm, portfolio and quarter-year, respectively. $compacct_{ikt}$ ($compcf_{ikt}$) is the mean of the ten highest firm-pair comparability values for firm $i$ in portfolio $k$, based on the mapping of returns into earnings (cash flows into accruals). In column 1 and 3, I draw a random sample of at most 10 holdings per portfolio and quarter, while ensuring that each firm appears in at least one mutual fund and one analyst portfolio (*random*). In column 2 and 4, I match holdings of the same firm in the same quarter from mutual fund to analyst portfolios (*matched*). I apply the coarsened exact matching (CEM) algorithm and match on portfolio means of size, book-to-market and spread (size, accrual and cash flow volatility) in column 2 (4). I further require that the difference in the number of portfolio peers is at most 10. $mf_{kt}$ is an indicator variable equal to one (zero) if portfolio $k$ is a mutual fund (analyst) portfolio. $ln(peers)_{kt}$ is the natural logarithm of the number of portfolio peers. $ln(mv)_{it}$ is the natural logarithm of the market value of equity. $btm_{it}$ is the ratio of book to market value of equity. $roa_{it}$ is income before extraordinary items over lagged total assets. $\star^1\_volatility_{it}$ is $ear\_volatility_{it}$ ($acc\_volatility_{it}$) in column 1 and 2 (column 3 and 4), the standard deviation of quarterly earnings (accruals) scaled by lagged total assets over 16 quarters. $\star^2\_volatility_{it}$ is $ret\_volatility_{it}$ ($cfo\_volatility_{it}$) in column 1 and 2 (column 3 and 4), the standard deviation of monthly returns (quarterly operating cash flow scaled by lagged total assets) over 48 months (16 quarters). $ear\_predictability_{it}$ is the adjusted $R^2$ from a regression of quarterly earnings on lagged quarterly earnings over 16 quarters. $volume_{it}$ is the quarterly mean of daily turnover. $spread_{it}$ is the quarterly mean of daily bid-ask spreads. Standard errors are clustered by portfolio $k$ and quarter $t$. *, ** and *** indicate significance at the 0.1, 0.05 and 0.01 level, respectively.

Table 5: **Mutual Fund Types**

| | $compacct_{ikt}$ | | $compcf_{ikt}$ | |
|---|---|---|---|---|
| *Sample* | *random* | *matched* | *random* | *matched* |
| $active_{kt}$ | 0.055*** | 0.071*** | −0.008 | 0.032*** |
| | (0.014) | (0.015) | (0.009) | (0.011) |
| $ln(tna)_{kt}$ | −0.002** | 0.000 | −0.001 | 0.001 |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| $ln(peers)_{kt}$ | 0.238*** | | 0.395*** | |
| | (0.006) | | (0.006) | |
| Firm characteristics of firm $i$ | yes | yes | yes | yes |
| Portfolio means of firm characteristics in portfolio $k$ | yes | no | yes | no |
| Investment objective fixed effects | yes | yes | yes | yes |
| Firm and quarter-year fixed effects | yes | yes | yes | yes |
| adj. $R^2$ | 0.602 | 0.622 | 0.684 | 0.663 |
| adj. $R^2$ within | 0.303 | 0.264 | 0.313 | 0.146 |
| $N$ | 956,225 | 542,950 | 963,288 | 502,510 |
| Holdings | 382,144 | 154,223 | 374,608 | 153,357 |
| Firms | 2,485 | 2,268 | 2,410 | 2,173 |
| Mutual funds | 3,981 | 3,405 | 3,972 | 3,435 |

*Notes:* The table displays estimates for the association between accounting comparability and mutual fund types. $i$, $k$ and $t$ denote firm, portfolio and quarter-year, respectively. $compacct_{ikt}$ ($compcf_{ikt}$) is the mean of the ten highest firm-pair comparability values for firm $i$ in portfolio $k$, based on the mapping of returns into earnings (cash flows into accruals). In column 1 and 3, I draw a random sample of at most 10 holdings per portfolio and quarter, while ensuring that each firm appears in at least one nonindexer and one indexer portfolio (*random*). In column 2 and 4, I match holdings of the same firm in the same quarter from nonindexer to indexer portfolios (*matched*). I apply the coarsened exact matching (CEM) algorithm and match on portfolio means of size, book-to-market and spread (size, accrual and cash flow volatility) in column 2 (4). I further require that the difference in the number of portfolio peers is at most 10. $active_{kt}$ is the share of firms in portfolio $k$ that is not in the S&P 500. $ln(tna)_{kt}$ is the natural logarithm of total net assets in portfolio $k$. $ln(peers)_{kt}$ is the natural logarithm of the number of portfolio peers. Firm characteristics comprise the natural logarithm of the market value of equity, $ln(mv)_{it}$; the ratio of book to market value of equity, $btm_{it}$; income before extraordinary items over lagged total assets, $roa_{it}$; the standard deviation of quarterly earnings (accruals) scaled by lagged total assets over 16 quarters in column 1 and 2 (column 3 and 4), $ear\_volatility_{it}$ ($acc\_volatility_{it}$); the standard deviation of monthly returns (quarterly operating cash flow scaled by lagged total assets) over 48 months (16 quarters) in column 1 and 2 (column 3 and 4), $ret\_volatility_{it}$ ($cfo\_volatility_{it}$); the adjusted $R^2$ from a regression of quarterly earnings on lagged quarterly earnings over 16 quarters, $ear\_predictability_{it}$; the quarterly mean of daily turnover, $volume_{it}$; and the quarterly mean of daily bid-ask spreads, $spread_{it}$. Standard errors are clustered by portfolio $k$ and quarter $t$. *, ** and *** indicate significance at the 0.1, 0.05 and 0.01 level, respectively.

Table 6: **Portfolio Selection**

| | $\boldsymbol{inclusion_{ikt}}$ | | | | | |
|---|---|---|---|---|---|---|
| *Comparability proxy* | $\boldsymbol{compacct_{ikt}}$ | | | $\boldsymbol{compcf_{ikt}}$ | | |
| *Indicator variable* | | $SPREAD$ | $RET$ | | $SPREAD$ | $RET$ |
| $comp\star_{ikt}$ | $0.020^{***}$ | $0.037^{***}$ | $0.051^{***}$ | $0.058^{***}$ | $0.055^{***}$ | $0.097^{***}$ |
| | (0.003) | (0.006) | (0.009) | (0.004) | (0.006) | (0.010) |
| $indicator_{it}$ | | 0.003 | $-0.049^{***}$ | | $0.012^{***}$ | $-0.021^{***}$ |
| | | (0.004) | (0.005) | | (0.004) | (0.006) |
| $comp\star_{ikt} \times indicator_{it}$ | | $-0.019^{***}$ | $-0.035^{***}$ | | 0.004 | $-0.048^{***}$ |
| | | (0.005) | (0.008) | | (0.006) | (0.009) |
| $ln(peers)_{kt}$ | −0.003 | $-0.005^{**}$ | $-0.006^{**}$ | $-0.018^{***}$ | $-0.018^{***}$ | $-0.023^{***}$ |
| | (0.002) | (0.002) | (0.002) | (0.003) | (0.003) | (0.003) |
| $ln(mv)_{it}$ | $0.031^{***}$ | $0.032^{***}$ | $0.028^{***}$ | $0.035^{***}$ | $0.036^{***}$ | $0.036^{***}$ |
| | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) |
| $btm_{it}$ | $0.056^{***}$ | $0.056^{***}$ | $0.063^{***}$ | $0.022^{*}$ | $0.022^{*}$ | $0.023^{*}$ |
| | (0.013) | (0.013) | (0.013) | (0.012) | (0.012) | (0.012) |
| $retq1_{it}$ | $0.152^{***}$ | $0.151^{***}$ | $0.156^{***}$ | $0.150^{***}$ | $0.150^{***}$ | $0.149^{***}$ |
| | (0.011) | (0.011) | (0.011) | (0.010) | (0.010) | (0.010) |
| $rety1_{it}$ | $0.100^{***}$ | $0.099^{***}$ | $0.101^{***}$ | $0.089^{***}$ | $0.089^{***}$ | $0.089^{***}$ |
| | (0.007) | (0.007) | (0.007) | (0.006) | (0.006) | (0.006) |
| $ret\_sd_{it}$ | $0.739^{*}$ | $0.736^{*}$ | $1.097^{**}$ | $1.396^{***}$ | $1.381^{***}$ | $1.373^{***}$ |
| | (0.416) | (0.417) | (0.418) | (0.507) | (0.509) | (0.504) |
| $volume_{it}$ | $0.067^{***}$ | $0.067^{***}$ | $0.069^{***}$ | $0.059^{***}$ | $0.060^{***}$ | $0.059^{***}$ |
| | (0.005) | (0.005) | (0.005) | (0.006) | (0.006) | (0.006) |
| $spread_{it}$ | −0.031 | $-0.034^{*}$ | $-0.041^{**}$ | $-0.066^{**}$ | $-0.067^{**}$ | $-0.065^{**}$ |
| | (0.019) | (0.019) | (0.020) | (0.027) | (0.027) | (0.027) |
| Investment objective fixed effects | yes | yes | yes | yes | yes | yes |
| Firm and quarter-year fixed effects | yes | yes | yes | yes | yes | yes |
| adj. $R^2$ | 0.045 | 0.045 | 0.046 | 0.054 | 0.054 | 0.055 |
| adj. $R^2$ within | 0.009 | 0.009 | 0.010 | 0.010 | 0.010 | 0.011 |
| $N$ | | 632,046 | | | 578,942 | |
| Holdings | | 527,562 | | | 487,950 | |
| Firms | | 2,880 | | | 2,740 | |
| Mutual funds | | 2,904 | | | 2,906 | |

*Notes:* The table displays estimates for the determinants of portfolio selection. $i$, $k$ and $t$ denote firm, portfolio and quarter-year, respectively. I limit mutual funds to nonindexers. I restrict observations to (i) holdings that are not part of portfolio $k$ in the preceding eight quarters and are newly included in quarter $t$; and (ii) potential holdings that I match based on firm characteristics: size, book-to-market, spread, earnings and return volatility (size, book-to-market, spread, accrual and cash flow volatility) in columns 1–3 (4–6). I apply the coarsened exact matching (CEM) algorithm and match exactly one potential holding to each new holding in the quarter of inclusion. The dependent variable is $inclusion_{ikt}$, an indicator variable equal to one if firm $i$ is newly included portfolio $k$ in quarter $t$. $comp\star_{ikt}$ is $compacct_{ikt}$ ($compcf_{ikt}$), the mean of the ten highest firm-pair comparability values for firm $i$ in portfolio $k$, based on the mapping of returns into earnings (cash flows into accruals). $SPREAD_{it}$ is an indicator variable equal to one if the four-year mean of monthly bid-ask spreads is above the median in quarter $t$. $RET_{it}$ is an indicator variable equal to one if the four-year standard deviation of monthly returns is above the median in quarter $t$. $ln(peers)_{kt}$ is the natural logarithm of the number of portfolio peers. $ln(mv)_{it}$ is the natural logarithm of the market value of equity. $btm_{it}$ is the ratio of book to market value of equity. $retq1_{it}$ is the return from quarter $(t-1)$ to quarter $t$. $rety1_{it}$ is the return from quarter $(t-4)$ to quarter $(t-1)$. $ret\_sd_{it}$ is the quarterly standard deviation of daily returns. $volume_{it}$ is the quarterly mean of daily turnover. $spread_{it}$ is the quarterly mean of daily bid-ask spreads. Standard errors are clustered by firm $i$ and quarter $t$. $^{*}$, $^{**}$ and $^{***}$ indicate significance at the 0.1, 0.05 and 0.01 level, respectively.

Table 7: **Selection vs. Influence**

| | $\boldsymbol{compacct_{ikt}}$ | $\boldsymbol{compcf_{ikt}}$ |
|---|---|---|
| $pre1_{ikt}$ | $0.021^{**}$ | $0.011^{***}$ |
| | (0.009) | (0.004) |
| $post1_{ikt}$ | $0.013^{**}$ | $0.011^{***}$ |
| | (0.006) | (0.003) |
| $post2_{ikt}$ | 0.003 | 0.002 |
| | (0.006) | (0.003) |
| $post3_{ikt}$ | 0.003 | $-0.002$ |
| | (0.005) | (0.003) |
| $post4_{ikt}$ | 0.001 | $-0.008^{**}$ |
| | (0.007) | (0.003) |
| Number of peers in portfolio $k$ | yes | yes |
| Firm characteristics of firm $i$ | yes | yes |
| Portfolio means of firm characteristics in portfolio $k$ | yes | yes |
| Investment objective fixed effects | yes | yes |
| Firm and quarter-year fixed effects | yes | yes |
| adj. $R^2$ | 0.628 | 0.678 |
| adj. $R^2$ within | 0.217 | 0.267 |
| $N$ | 76,944 | 64,824 |
| Holdings | 3,206 | 2,701 |
| Firms | 775 | 688 |
| Mutual Funds | 551 | 534 |

*Notes:* The table displays estimates for the change in comparability upon inclusion into mutual fund portfolios. $i$, $k$ and $t$ denote firm, portfolio and quarter-year, respectively. I limit mutual funds to nonindexers. I restrict observations to holdings that are not part of portfolio $k$ for eight quarters, are then newly included and remain in portfolio $k$ for at least four years. For each holding, the sample contains 24 observations, two years before and four years following inclusion. $compacct_{ikt}$ ($compcf_{ikt}$) is the mean of the ten highest firm-pair comparability values for firm $i$ in portfolio $k$, based on the mapping of returns into earnings (cash flows into accruals). I define $pre1_{ikt}$ and $post\star_{ikt}$ relative to the quarter of inclusion into portfolio $k$. $pre1_{ikt}$ is an indicator variable equal to one starting in the year preceding inclusion. $post1_{ikt}$ ($post2_{ikt}$) [$post3_{ikt}$] {$post4_{ikt}$} is an indicator variable equal to one in the first (second) [third] {fourth} starting in the year subsequent to inclusion. Number of peers is the natural logarithm of the number of portfolio peers, $ln(peers)_{kt}$. Firm characteristics comprise the natural logarithm of the market value of equity, $ln(mv)_{it}$; the ratio of book to market value of equity, $btm_{it}$; income before extraordinary items over lagged total assets, $roa_{it}$; the standard deviation of quarterly earnings (accruals) scaled by lagged total assets over 16 quarters in column 1 (column 2), $ear\_volatility_{it}$ ($acc\_volatility_{it}$); the standard deviation of monthly returns (quarterly operating cash flow scaled by lagged total assets) over 48 months (16 quarters) in column 1 (column 2), $ret\_volatility_{it}$ ($cfo\_volatility_{it}$); the adjusted $R^2$ from a regression of quarterly earnings on lagged quarterly earnings over 16 quarters, $ear\_predictability_{it}$; the quarterly mean of daily turnover, $volume_{it}$; and the quarterly mean of daily bid-ask spreads, $spread_{it}$. Standard errors are clustered by portfolio $k$ and quarter $t$. $^{*}$, $^{**}$ and $^{***}$ indicate significance at the 0.1, 0.05 and 0.01 level, respectively.

# Temporary Migration within Multinational Corporations: Evidence from the Accounting Industry*

**Ulf Brüggemann**
Humboldt University of Berlin

**Joachim Gassen**
Humboldt University of Berlin

**Sarah Kröchert**
Humboldt University of Berlin

***Abstract***

This research note examines the determinants of temporary migration within multinational corporations. We examine a comprehensive dataset on international assignments within a large accounting firm. Our analysis yields two findings that are new to the literature. First, we document that country-level factors established in prior studies on long-term migration (e.g., economic ties or cultural proximity between home and host country) are also associated with migration decisions in our setting. Second, we provide some evidence that regulatory variation across countries at the occupational level (e.g., differences in accounting standards) reduces the extent of temporary migration within multinational corporations.

*The study is based on proprietary data provided by a large accounting firm. We thank our contacts in this firm for supporting this project by making their data available to us. We also thank the members of the Research Training Group (RTG) 1659 at Humboldt University of Berlin and the audience of Humboldt's accounting brown bag seminar for valuable comments.

# 1 Introduction

Temporary migration of skilled labor is a highly relevant phenomenon around the world and particularly pronounced in the accounting industry. The four largest accounting firms have an aggregate workforce of more than 500,000 employees located in more than 150 countries. Of these, roughly 11,000 to 16,000 (2.2% to 3.2%) participate in formal international mobility programs [Beaverstock, 2017]. These statistics support the notion that temporary mobility[55] within multinational corporations is a widespread phenomenon in the accounting industry and, more generally, in industries that offer skilled services. Yet, we know very little about the determinants of this phenomenon. This research note addresses this gap by presenting large-scale empirical evidence based on a proprietary dataset on international assignments within a large accounting firm.

The firm operates in three lines of services (LoS) or occupations as is common in the accounting industry. The first LoS (auditing) combines typical auditing and assurance-oriented activities. The second LoS (taxation) comprises tax filing preparation and tax-related consulting activities. The third LoS (consulting) includes general consulting services, which are usually related to the accounting activities of client firms.

Our empirical analysis consists of two parts. In the first part, we use a sample of 50,060 possible country-pair observations across the three LoS. We consider a pair possible if both the host and the home country operate in the respective LoS. Our outcome variable is an indicator equal to one if there is at least one individual from the respective home country working in the respective host country. The analysis shows that the likelihood of temporary migration is higher if host and home country have strong economic ties, share a similar culture or are geographically close to each other. These findings support the view that country-level determinants established in prior literature on long-term migration across firms [e.g., Beine, Docquier and Özden, 2011; Mayda, 2010] also matter for temporary migration within multinational corporations.

In the second part of the empirical analysis, we examine whether regulatory differences across countries affect migration decisions in certain occupations. This analysis is based on the following arguments: Employees in auditing need to have detailed knowledge of the relevant financial accounting and auditing rules that their clients are subject to. Prior work documents that these rules vary significantly across countries [e.g., Bae, Tan and Welker, 2008]. This cross-country variation can have two opposing effects on temporary migration. On the one

[55] In line with prior literature on temporary movements of labor [Beaverstock, 2017], we use the terms 'migration' and 'mobility' interchangeably throughout this note.

hand, regulatory differences between host and home country may impose a cost of migration if the employee has to be trained in the host country's regulatory environment. On the other hand, temporary migration may be enhanced to the extent that the regulatory expertise of the individual is demanded by the host country (e.g., to provide services to a subsidiary of a home country's parent operating in the host country). Similar arguments hold for employees in taxation. In contrast, consulting skills are more universal and, thus, less likely to be affected by regulatory differences across countries.

We test for these potential associations at the occupational level using the share of mobile employees as the outcome variable (i.e., we focus on the extent instead of the existence of temporary migration). The analysis is restricted to the subset of 1,291 country pairs for which we observe positive mobility. We document that the number of auditors and tax experts on international assignments is higher for host countries with accounting and tax regimes that resemble the regulatory environment in the home country. These results suggest that regulatory differences at the occupational level generate migration costs that are higher than potential benefits sparked by demand for expertise in foreign rules. We also find that migration decisions by consultants are positively associated with the similarity of tax rules between home and host country. This finding is harder to reconcile with our arguments above, illustrating that our empirical evidence on occupation-level migration determinants is not conclusive.

Extant literature on temporary migration within multinational corporations focuses on descriptive statistics and qualitative aspects of this important phenomenon [e.g., Beaverstock, 1991, 1996; Millar and Salt, 2008; Peixoto, 2001; Tzeng, 1995; Beaverstock, 2017, for a short overview]. We contribute to this literature by providing two new findings. First, we use insights from prior migration literature to show that established country-level factors (e.g., economic ties or cultural proximity between home and host country) are also associated with migration decisions in our setting. Second, we find that regulatory variation across countries at the occupational level hampers temporary migration within firms. While the results on occupation-level determinants are inconclusive, we provide one of the first pieces of evidence that regulatory variation across countries has the potential to impose barriers to labor market migration and competition [see also Bloomfield, Brüggemann, Christensen and Leuz, 2017].

## 2 Migration Data

Our study is based on a proprietary dataset on international assignments within a large accounting firm.[56] Individuals in this dataset are employees of the firm in the home country and remain employees of the firm when working on an international assignment in the host country. The employees change their workplace only temporarily; the average (median) length of an assignment is 1.9 (2.0) years. At the end of their stay abroad, they return to their home office.

The raw dataset includes information on the LoS (auditing, taxation or consulting), the home country and the host country for all individuals that were on an international assignment during the period 2005 to 2008. Since this period is too short to exploit variation across time, we use this dataset for cross-sectional analyses only. The key variable is a mobility measure. This measure is a ratio with the number of mobile employees in the host country (i.e., those on an international assignment) in the numerator and the total number of employees in the respective home country in the denominator. We compute the mobility measure by LoS for each possible pair of host and home country. We consider a pair possible if both the host and the home country operate in the respective LoS. Since the firm offers consulting services in 116 countries, this procedure yields 13,340 possible country-pair observations ($= 116 \times 115$) for this LoS. The firm operates in the other two LoS in 136 countries leading to 18,360 possible country pairs for both auditing and taxation.[57] Hence, our final dataset comprises 50,060 possible country-pair observations across three LoS. For comparison, the studies by Beaverstock [1991, 1996] focus on the U.K. as home or host country and do not distinguish between different LoS.

Our dataset is subject to two caveats: First, there is a mismatch between the numerator and the denominator of our mobility measure. While the numerator is based on yearly averages of international assignments over the period 2005 to 2008, the denominator uses data on the total number of employees as of 2016 because data for earlier periods are not available to us. This mismatch could affect our results to the extent that the number of employees across home countries has changed systematically over the last decade. Second, our dataset contains little usable information on individual assignments (e.g., demographic data, wage data). We therefore focus our empirical analysis on country-level determinants of temporary migration within the accounting firm.

[56] We are not allowed to disclose the firm's name or any other information that may reveal its identity (e.g., total number of employees or total number of mobile employees).

[57] The firm offers auditing services in three countries without offering taxation services, and vice versa. The final dataset therefore comprises 139 countries in total.

# 3 Migration Determinants

## 3.1 Existence of Temporary Migration

The labor migration literature indicates that prospects of higher future wages or, in general, higher living standards provide incentives to migrate to host countries. It further argues that economic ties and cultural as well as geographic proximity between home and potential host countries strengthen these incentives by reducing migration costs. Existing business relations and common forms of understanding decrease information asymmetries prior to migration and facilitate the familiarization with a new environment after arrival in the host country. Short distances imply lower direct costs of moving and are usually positively correlated with the extent of economic and personal exchange [e.g., Sjaastad, 1962]. We investigate whether the mobility patterns in our dataset are associated with these country-level determinants. Ex ante, it is not obvious that such associations exist. First, we focus on temporary migration, which might weaken the role of cultural factors because employees return to their home country in the near future [e.g., Dustmann and Görlach, 2016]. Second, we focus on within-firm migration, which might lessen the importance of economic factors in the host country because employment is guaranteed and firm-specific characteristics such as career development programs or temporary vacancies play into the migration decision [e.g., Edström and Galbraith, 1977].

We begin the empirical analysis by assessing the likelihood of observing temporary migration between possible pairs of host and home countries. The dependent variable is an indicator, $mobile_{ijk}$, which is equal to one if our mobility measure is larger than zero, i.e., if there is at least one individual from LoS $k$ in home country $i$ that works in host country $j$ during the period 2005 to 2008. The indicator is set to zero if the mobility measure equals zero. We estimate the following regression model:

$$\begin{aligned} mobile_{ijk} =& \beta_1 gdp\_high_{ij} + \beta_2 taxhaven\_host_{ij} + \beta_3 crosslisting_{ij} + \beta_4 tradeflow_{ij} \\ &+ \beta_5 comlang_{ij} + \beta_6 colony_{ij} + \beta_7 distance\_high_{ij} + \beta_m FE_i + \beta_n FE_j + \epsilon_{ijk} \end{aligned}$$

The set of independent variables is inspired by prior literature that examines long-term migration decisions outside current employments [e.g., Beine et al., 2011; Mayda, 2010]. These variables are defined as follows (see Appendix A.1 for further details): $gdp\_high_{ij}$ is an indicator equal to one if the per capita GDP of host country $j$ is larger than the per capita GDP of home country $i$, and equal to zero otherwise. We use this variable as a proxy for income differences between home and host country. The variable is based on GDP data from the World Bank's World Development Indicators. $taxhaven\_host_{ij}$ is an indicator equal to one if host country $j$

is among the five tax havens that are closest in geographic distance to home country $i$, and equal to zero otherwise. We use this variable to capture high demand for tax consulting. We compute the variable using the list of tax havens provided by Dyreng and Lindsey [2009]. Our sample of 139 countries comprises 22 tax havens according to this list. $crosslisting_{ij}$ is an indicator equal to one if firms from home country $i$ are cross-listed in host country $j$, and zero otherwise. This variable is a proxy for capital market interdependencies between home and host country. We compute it using data on cross-listed firms from Sarkissian and Schill [2004, 2009, 2016].[58] $tradeflow_{ij}$ is an indicator equal to one if exports from home country $i$ to host country $j$ are above the median for home country $i$, and zero otherwise. It is based on trade flow data from CEPII [Fouquin and Hugot, 2016]. We use this variable to quantify the importance of the host country as a recipient of domestic products from the perspective of the home country. The remaining independent variables characterize cultural and geographic proximity between home and host country based on data from CEPII [Mayer and Zignago, 2011]. $comlang_{ij}$ and $colony_{ij}$ are indicators equal to one if both countries share the same language and a colonial history, respectively. $distance\_high_{ij}$ is an indicator equal to one if the distance between the home and host country is larger than the sample median, and equal to zero otherwise.

We estimate the model as a cross-sectional OLS regression with separate fixed effects for home and host countries, $FE_i$ and $FE_j$. These fixed effects ensure that the coefficient estimates on the independent variables are less susceptible to the influence of individual countries with extreme home or host characteristics (e.g., high economic development, strict immigration policies). Standard errors are adjusted for heteroskedasticity and two-way clustering by home and host country. To explore potential differences in migration determinants across LoS, we estimate the regression separately for auditing, taxation and consulting.

Table 1 provides summary statistics for all variables. The statistics for the dependent variable $mobile_{ijk}$ show that there is no temporary migration between most country pairs. Moreover, migration patterns differ substantially across LoS: $mobile_{ijk}$ is equal to one for 4.5% of all country pairs in auditing, 2.3% in taxation and 3.4% in consulting. The statistics for the independent variables illustrate that cross-listings are rare (2.0% of all country pairs) and that 13.3% and 1.5% of all country pairs share the same language and a colonial history, respectively.

Table 2 shows correlations between all variables. The correlation estimates between the dependent variable and the independent variables are statistically significant throughout and often economically meaningful. For instance, $mobile_{ijk}$ has a Pearson correlation coefficient of

[58] Sergei Sarkissian makes the data available on his website at: http://sergei-sarkissian.com/data.html.

0.402 with $crosslisting_{ij}$ and of 0.174 with $colony_{ij}$. The correlations among the independent variables are also high. As an example, $tradeflow_{ij}$ has a Pearson correlation coefficient of 0.205 with $gdp\_high_{ij}$, of 0.131 with $crosslisting_{ij}$ and of -0.161 with $distance\_high_{ij}$.

Table 3 displays the regression results. In contrast to the univariate correlations, not all independent variables are significantly associated with the dependent variable. For example, the coefficient on $gdp\_high_{ij}$ is not statistically significant in any specification. This finding suggests that income differences between home and host country play little role in the decision to temporarily migrate within the accounting firm once other determinants are controlled for. In contrast, the coefficient estimate on $crosslisting_{ij}$ is statistically significant and large in magnitude (between 0.318 and 0.354) for all LoS. While this result indicates a strong influence of capital market interdependencies on migration decisions in our setting, we cannot rule out that the cross-listing proxy captures part of the explanatory power of other independent variables that are highly correlated with it (see Table 2). The coefficient estimate on $tradeflow_{ij}$ is (weakly) statistically significant for all LoS but low in magnitude (between 0.004 and 0.007), in line with a modestly positive association between the likelihood of migration and trade ties. Similarly, the coefficient estimate on $taxhaven\_host_{ij}$ is weakly statistically significant for auditing and taxation, providing mild evidence that geographically close tax havens create demand for services in these LoS. Finally, the regression results suggest that employees in the accounting firm are more likely to migrate to host countries that, relative to their home country, have the same language, share the same colonial history and are geographically close. Specifically, the coefficient estimates on $comlang_{ij}$, $colony_{ij}$ and $distance\_high_{ij}$ are statistically significant in almost all specifications with absolute magnitudes ranging from 0.009 ($distance\_high_{ij}$ for taxation) to 0.144 ($colony_{ij}$ for auditing). Untabulated statistics show that the documented associations between dependent and independent variables do not vary systematically across LoS with a few exceptions (e.g., the coefficient estimate on $taxhaven\_host_{ij}$ is significantly larger for auditing and taxation relative to consulting).

The bottom of Table 3 depicts the three home and host countries with the largest fixed effects coefficients. The United States (US) and the United Kingdom (GB) are among the most popular home and host countries in all LoS. Australia (AU) and Canada (CA) are popular host countries. These statistics suggest that mobility is particularly common between large, English-speaking countries.

Taken together, the results of the first part of the empirical analysis support the notion that country-level determinants established in prior migration literature [e.g., Beine et al., 2011; Mayda, 2010] also matter in our specific context of temporary migration within firms.

### 3.2 Extent of Temporary Migration

In the second part of the empirical analysis, we investigate whether migration decisions in our setting also depend on LoS-specific factors. This analysis is based on the following argument: Auditors and tax experts regularly have to turn to accounting, auditing and tax rules to carry out their duties. These rules can be international standards, such as the IFRS, or national law with differing degrees of harmonization, such as most tax regulation. Similarity in these rules across borders facilitates the transition into a new work environment abroad and, thus, lowers the cost for employees to migrate [e.g., Bloomfield et al., 2017]. Consultants, on the other hand, apply procedures that are not subject to national regulations and, thus, should have little impact on migration decisions within this LoS. We test these expectations by examining whether migration decisions depend on similarity in accounting and tax regimes as LoS-specific factors. We benchmark auditing and taxation against consulting, which should respond less to these LoS-specific factors. In addition, we benchmark associations with similarity in accounting and tax regimes against assocations with a well established migration determinant, income differences. Monetary incentives, such as higher compensation, are usually regarded as first-order determinants [e.g., Borjas, 1987; Grogger and Hanson, 2011; Roy, 1951] and could also matter in our setting. Within-firm mobility normally implies additional compensation for relocation and potential adjustments for higher living expenses. Anecdotal evidence from the accounting firm confirms that employees in our sample can indeed only be better off financially by migrating abroad.

In contrast to the first part of the analysis, we examine the extent instead of the existence of temporary migration and resort to finer measures informing about relative numbers of mobile employees. Figure 1 illustrates that relative numbers vary across LoS. It displays, for each country, flows to the host countries to which mobile employees are sent (in the color of the country) as well as flows from home countries from which it receives mobile employees (in the colors of the respective home countries). Flows are expressed relative to the overall number of mobile employees and countries correspond to the sample countries for the analysis below. The figure demonstrates, for instance, that South Africa sends relatively more employees in auditing; that India is a relatively more important home country in consulting; and that assignments between countries often associated with favorable tax regimes (e.g., Netherlands, Luxembourg and Switzerland) play a larger role in taxation.

In the analysis, we use our mobility measure in continuous form ($mobile\_share_{ijk}$) as dependent variable (see Section 2). We restrict the analysis to the subset of country pairs for which we observe positive mobility, i.e., where $mobile_{ijk}$ is equal to one. As Figure 2a illustrates, raw

values are positively skewed because each LoS has a few outliers that mostly belong to host countries with few employees. To reduce the influence of these outliers, we use the mobility measure in logarithmic form ($ln(mobile\_share)_{ijk}$) in the regression analysis. Figure 2b shows that the distribution of this logged measure roughly follows the normal distribution. We estimate the following regression model:

$$\begin{aligned} ln(mobile\_share)_{ijk} =& \beta_1 acc\_sim_{ij} + \beta_2 tax\_sim_{ij} + \beta_3 gdp\_high_{ij} \\ & [+\beta_l CONTROLS_{ij}] + \beta_m FE_i + \beta_n FE_j + \epsilon_{ijk} \end{aligned}$$

The independent variables are defined as follows (see Appendix A.1 for further details): $acc\_sim_{ij}$ is our proxy for similarity in accounting regimes. We follow Bae, Tan and Welker [2008] and count differences in accounting rules for country pairs. Bae et al. [2008] provide information on whether accounting rules agree with IAS for 49 countries based on a survey study summarized in Street [2002]. We add around 30 countries that adopted IFRS until 2007 and compute similarity scores in line with Cascino and Gassen [2015].[59] $acc\_sim_{ij}$ is computed as the number of differences in accounting rules for the respective country pair divided by the maximum difference in the sample. We multiply this score by -1 so that the resulting proxy is increasing in accounting similarity. $tax\_sim_{ij}$ is our proxy for similarity in tax regimes. This proxy is based on data from the tax attractiveness index described in Schanz, Dinkel, Fritz, Grosselfinger and Keller [2017] and Keller and Schanz [2013] and applied in previous studies [e.g., Beuselinck, Deloof and Vanstraelen, 2015; Schanz, Dinkel and Keller, 2017]. The index aggregates information about the existence and design of various tax rules in 100 countries.[60] We focus on ten items that allow simple comparisons across countries using values for the year 2007.[61] $tax\_sim_{ij}$ is the number of differences in tax rules for each country pair divided by the maximum difference in the sample. We multiply the score by -1 so that the resulting proxy is increasing in tax similarity. $gdp\_high_{ij}$ captures income differences between home and host country (see Section 3.1) and, thus, reflects the potential for wage increases. $CONTROLS_{ij}$ includes the remaining migration determinants used in the previous Section: $taxhaven\_host_{ij}$, $crosslisting_{ij}$, $tradeflow_{ij}$,

[59] We use the term 'IFRS' to refer to both the International Financial Reporting Standards (IFRS) and the International Accounting Standards (IAS). We have details on IFRS adoption years from the jurisdiction profiles provided by the IFRS Foundation, available at: http://www.ifrs.org/use-around-the-world/use-of-ifrs-standards-by-jurisdiction/.

[60] The data are publicly available through the website http://www.taxindex.de.

[61] Specifically, we focus on the following ten items: anti-avoidance rules, CFC rules, EU member state, group taxation regime, holding tax climate, loss carryback, loss carryforward, R&D tax incentives, thin capitalization rules and transfer pricing rules. All of these items are either coded 0/1 or 0/0.5/1 and hence allow straightforward comparisons in country pairs.

$comlang_{ij}$, $colony_{ij}$ and $distance\_high_{ij}$.[62]

As in Section 3.1, we estimate the model separately for each LoS and as cross-sectional OLS regressions with separate fixed effects for home and host countries, $FE_i$ and $FE_j$. Standard errors are again adjusted for heteroskedasticity and two-way clustering by home and host country.

The summary statistics in Table 1 show that, on average, between 0.708% (taxation) and 1.295% (auditing) of all employees in the home country are on an international assignment. The proxy for accounting similarity, $acc\_sim_{ij}$, has a mean (median) of -0.295 (-0.250). These values are higher than those reported in related literature [e.g., Cascino and Gassen, 2015] due to the large number of IFRS adopters in our sample. The proxy for tax similarity, $tax\_sim_{ij}$, has a mean (median) of -0.504 (-0.500) indicating that about half of the tax rules are the same in the average country pair.

The correlations in Table 2 for $ln(mobile\_share)_{ijk}$, $acc\_sim_{ij}$ and $tax\_sim_{ij}$ are based on the sample used in the subsequent regressions (1,291 observations). The correlations between $ln(mobile\_share)_{ijk}$ and the country-level migration determinants are largely similar to those reported in Section 3.1. For instance, the mobility measure has a Pearson correlation coefficient of 0.182 with $crosslisting_{ij}$ and of 0.172 with $colony_{ij}$. The proxies for accounting and tax similarity are higher in country pairs with stronger economic ties and shorter distances. For example, $acc\_sim_{ij}$ ($tax\_sim_{ij}$) has a Pearson correlation coefficient of 0.062 (0.068) with $crosslisting_{ij}$ and of -0.133 (-0.068) for $distance\_high_{ij}$.[63]

Table 4 shows the regression results for the second part of our empirical analysis. The results are consistent with LoS-specific factors playing a role for migration in our setting. Higher accounting similarity is positively associated with the extent of migration in auditing and taxation, suggesting that the proxy also captures characteristics of tax regimes. For auditing, the coefficient estimate on $acc\_sim_{ij}$ is statistically significant, even after controlling for the country-level migration determinants used in the first part. The association is economically meaningful as the coefficient estimate of 0.800 corresponds to a 28% increase in the mobility measure for a one-standard-deviation increase in accounting similarity. For taxation, the coefficient estimate on $acc\_sim_{ij}$ is weakly statistically significant once the control variables enter and smaller in magnitude (0.401). Economically, it translates into a 14% increase in the mobility measure for

[62] Data on some independent variables are not available for all countries. We limit the sample in this analysis to country pairs with non-missing information on accounting and tax similarity scores. For all other variables, we set missing values to zero.

[63] However, correlations for accounting similarity and the proxies for economic ties are weaker. $acc\_sim_{ij}$ has a Pearson correlation coefficient of -0.079 with $tradeflow_{ij}$ and statistically insignificant Spearman correlation coefficients with both $crosslisting_{ij}$ and $tradeflow_{ij}$.

a one-standard-deviation increase in accounting similarity.[64] Higher similarity in tax rules is positively associated with the extent of migration in taxation and consulting. For taxation, the association is not robust to the inclusion of the control variables; the coefficient estimate on $tax\_sim_{ij}$ becomes insignificant. For consulting, it becomes even stronger with the inclusion of the control variables. This finding is somewhat contrary to our expectation that consultants should respond less to LoS-specific factors. Finally, income differences do not seem to be associated with the extent of migration in all three LoS. For auditing and taxation, the coefficient estimates are not different from zero throughout. For consulting, the coefficient estimate is weakly significant, but only without further control variables. In the middle part of Table 4, we compare the coefficient estimates on the key independent variables across LoS. The coefficient estimates on $acc\_sim_{ij}$ are larger and statistically different for auditors relative to the other two LoS. Supporting the rather weak findings above, the coefficient estimates on $tax\_sim_{ij}$ are not statistically different for tax experts relative to auditors and consultants. In a similar vein, the coefficient estimates on $gdp\_high_{ij}$ are not statistically different for consultants relative to auditors and tax experts.

Coefficient estimates for the other migration determinants are weaker, but largely consistent with the results in the first part of the analysis. Common cultural links with the host country (captured by $comlang_{ij}$ and $colony_{ij}$) are significantly associated with the share of mobile employees in all LoS. However, we do not have an explanation for the significantly negative coefficient estimate on $taxhaven\_host_{ij}$ in consulting.

The bottom of Table 4 shows the three home and host countries with the largest fixed effects coefficients. The most popular home countries tend to have relatively few employees, such as Macedonia (MK) or Zimbabwe (ZW). The most popular host countries are primarily English-speaking countries with large offices, such as the United States (US), the United Kingdom (GB) and Australia (AU).

In summary, our empirical findings are in line with LoS-specific factors shaping migration decisions. Specifically, the number of auditors and tax experts on international assignments is higher for host countries with accounting (and tax) regimes that resemble the regulatory environment in the home country. Yet, our empirical findings are not clear. Despite greater flexibility, consultants appear to respond to similarity in tax regimes and this to a higher degree than to income differences.

[64] The standard deviation of $acc\_sim_{ij}$ for the *auditing* (*taxation*) subsample amounts to 0.308 (0.327).

# 4 Conclusion

This research note examines the determinants of temporary migration within multinational corporations. Based on a comprehensive and proprietary dataset provided by a large accounting firm, we document that country-level as well as occupation-level determinants are associated with the likelihood and extent of temporary migration. While the associations with country-level determinants (e.g., proxies for economic ties or geographic proximity between home and host country) are in line with prior literature on long-term migration across firms, we are, to our knowledge, the first to document that these determinants also matter for migration decisions within firms. In addition, we find some evidence that occupation-related regulatory variation across countries (e.g., differences in accounting standards) reduces the extent of temporary migration within firms.

Our analysis is explorative in nature and some of the evidence is not conclusive. Nonetheless, we believe that the analysis yields interesting insights on what drives temporary migration within multinational corporations. Thus far, these insights have been elusive to researchers due to a lack of appropriate data. Further research in this area is warranted, in particular with regard to how regulatory differences across countries affect the international market for skilled labor. We hope that this research note encourages such research.

# A Appendix

## A.1 Variable Definitions

| Variable | Description |
|---|---|
| $mobile_{ijk}$ | Indicator variable equal to one if at least one individual from LoS $k$ in home country $i$ works in host country $j$ in the years 2005–2008. Line of service (LoS) is *auditing*, *taxation* or *consulting*. The migration data pertains to mobility within a large accounting firm. |
| $mobile\_share_{ijk}$ [$ln(mobile\_share)_{ijk}$] | [The natural logarithm of] the ratio of the annual number of employees from LoS $k$ in home country $i$ working in host country $j$ to the number of employees from LoS $k$ in home country $i$, as average over the years 2005–2008. We multiply $mobile\_share_{ijk}$ by 100 for representation in tables and figures. |
| $acc\_sim_{ij}$ | Accounting similarity following Bae et al. [2008]. |
| $tax\_sim_{ij}$ | Tax similarity based on data from the tax attractiveness index [Schanz, Dinkel, Fritz, Grosselfinger and Keller, 2017; Keller and Schanz, 2013]. |
| $gdp\_high_{ij}$ | Indicator variable equal to one if the per capita GDP of host country $j$ is larger than the per capita GDP of home country $i$. GDP data are from the World Bank's World Development Indicators. We employ the average over the years 2005–2008 and set missing values to zero. |
| $taxhaven\_host_{ij}$ | Indicator variable equal to one if host country $j$ is among the five closest tax havens for home country $i$. We identify tax havens following Dyreng and Lindsey [2009]. |
| $crosslisting_{ij}$ | Indicator variable equal to one if firms from home country $i$ are cross-listed in host country $j$. We use cross-listings from the dataset described in Sarkissian and Schill [2004, 2009, 2016]. For country pairs that are not part of the dataset, we set cross-listings to zero. |
| $tradeflow_{ij}$ | Indicator variable equal to one if exports from home country $i$ to host country $j$ are above the median for home country $i$. Data are from CEPII [Fouquin and Hugot, 2016]. For country pairs that are not part of the dataset, we set exports to zero. |
| $comlang_{ij}$ | Indicator variable equal to one if home country $i$ and host country $j$ share the same language. |
| $colony_{ij}$ | Indicator variable equal to one if home country $i$ and host country $j$ have a common colonial history. |

*Continued on next page*

*Continued from previous page*

| **Variable** | **Description** |
|---|---|
| $distance\_high_{ij}$ | Indicator variable equal to one if the distance between home country $i$ and host country $j$ is larger than the sample median. Data on language, colonial links and distance are from CEPII [Mayer and Zignago, 2011]. We set missing values to zero. |

*Notes*: $i$, $j$ and $k$ denote home country, host country and line of service (LoS), respectively.

## References


Bae, K.-H., Tan, H. and Welker, M. [2008], 'International GAAP differences: The impact on foreign analysts', *The Accounting Review* **83**(3), 593–628.

Beaverstock, J. V. [1991], 'Skilled international migration: An analysis of the geography of international secondments within large accountancy firms', *Environment and Planning A* **23**(8), 1133–1146.

Beaverstock, J. V. [1996], 'Subcontracting the accountant! Professional labour markets, migration, and organisational networks in the global accountancy industry', *Environment and Planning A* **28**(2), 303–326.

Beaverstock, J. V. [2017], 'Migration: Skilled international labor', *The International Encyclopedia of Geography*.

Beine, M., Docquier, F. and Özden, Ç. [2011], 'Diasporas', *Journal of Development Economics* **95**(1), 30–41.

Beuselinck, C., Deloof, M. and Vanstraelen, A. [2015], 'Cross-jurisdictional income shifting and tax enforcement: Evidence from public versus private multinationals', *Review of Accounting Studies* **20**(2), 710–746.

Bloomfield, M. J., Brüggemann, U., Christensen, H. B. and Leuz, C. [2017], 'The effect of regulatory harmonization on cross-border labor migration: Evidence from the accounting profession', *Journal of Accounting Research* **55**(1), 35–78.

Borjas, G. J. [1987], 'Self-selection and the earnings of immigrants', *The American Economic Review* **77**(4), 531–553.

Cascino, S. and Gassen, J. [2015], 'What drives the comparability effect of mandatory IFRS adoption?', *Review of Accounting Studies* **20**(1), 242–282.

Dustmann, C. and Görlach, J.-S. [2016], 'The economics of temporary migrations', *Journal of Economic Literature* **54**(1), 98–136.

Dyreng, S. D. and Lindsey, B. P. [2009], 'Using financial accounting data to examine the effect of foreign operations located in tax havens and other countries on U. S. multinational firms' tax rates', *Journal of Accounting Research* **47**(5), 1283–1316.

Edström, A. and Galbraith, J. R. [1977], 'Transfer of managers as a coordination and control strategy in multinational organizations', *Administrative Science Quarterly* **22**(2), 248–263.

Fouquin, M. and Hugot, J. [2016], 'Two centuries of bilateral trade and gravity data: 1827–2014', *CEPII Working Paper No. 2016–14*. Available at http://www.cepii.fr/CEPII/-fr/publications/wp/abstract.asp?NoDoc=9134.

Grogger, J. and Hanson, G. H. [2011], 'Income maximization and the selection and sorting of international migrants', *Journal of Development Economics* **95**(1), 42–57.

Keller, S. and Schanz, D. [2013], 'Measuring tax attractiveness across countries', *arqus Working Paper No. 143*. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2368766.

Mayda, A. M. [2010], 'International migration: A panel data analysis of the determinants of bilateral flows', *Journal of Population Economics* **23**(4), 1249–1274.

Mayer, T. and Zignago, S. [2011], 'Notes on CEPII's distance measures: The GeoDist database', *CEPII Working Paper No. 2011–25*. Available at http://www.cepii.fr/CEPII/-en/publications/wp/abstract.asp?NoDoc=3877.

Millar, J. and Salt, J. [2008], 'Portfolios of mobility: The movement of expertise in transnational corporations in two sectors – aerospace and extractive industries', *Global Networks* **8**(1), 25–50.

Peixoto, J. [2001], 'The international mobility of highly skilled workers in transnational corporations: The macro and micro factors of the organizational migration of cadres', *International Migration Review* **35**(4), 1030–1053.

Roy, A. D. [1951], 'Some thoughts on the distribution of earnings', *Oxford Economic Papers* **3**(2), 135–146.

Sarkissian, S. and Schill, M. J. [2004], 'The overseas listing decision: New evidence of proximity preference', *The Review of Financial Studies* **17**(3), 769–809.

Sarkissian, S. and Schill, M. J. [2009], 'Are there permanent valuation gains to overseas listing?', *The Review of Financial Studies* **22**(1), 371–412.

Sarkissian, S. and Schill, M. J. [2016], 'Cross-listing waves', *Journal of Financial and Quantitative Analysis* **51**(1), 259–306.

Schanz, D., Dinkel, A., Fritz, J., Grosselfinger, C. and Keller, S. [2017], 'The tax attractiveness index: Methodology', *Working Paper*. Available at https://papers.ssrn.com/sol3/papers.cfm?-abstract_id=2368766.

Schanz, D., Dinkel, A. and Keller, S. [2017], 'Tax attractiveness and the location of German-controlled subsidiaries', *Review of Managerial Science* **11**(1), 251–297.

Sjaastad, L. A. [1962], 'The costs and returns of human migration', *Journal of Political Economy* **70**(5), 80–93.

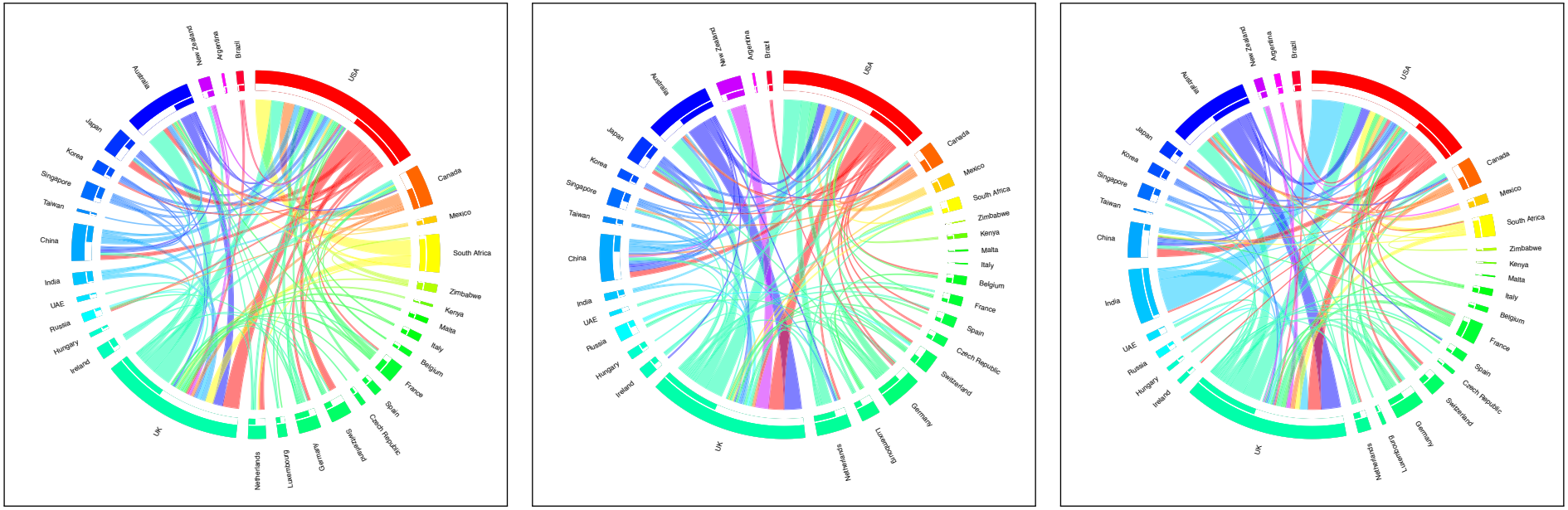
Street, D. L. [2002], 'GAAP 2001 – Benchmarking national accounting standards against IAS: Summary of results', *Journal of International Accounting, Auditing and Taxation* **11**(1), 77–90.

Tzeng, R. [1995], 'International labor migration through multinational enterprises', *The International Migration Review* **29**(1), 139–154.

# Figures

Figure 1: **Distribution of Mobile Employees**

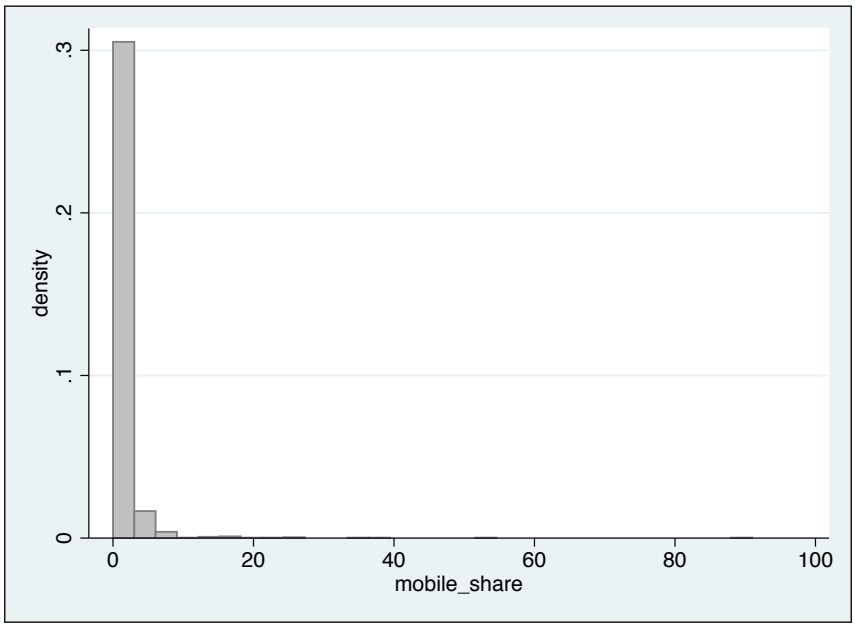
(a) ***auditing*** (b) ***taxation*** (c) ***consulting***



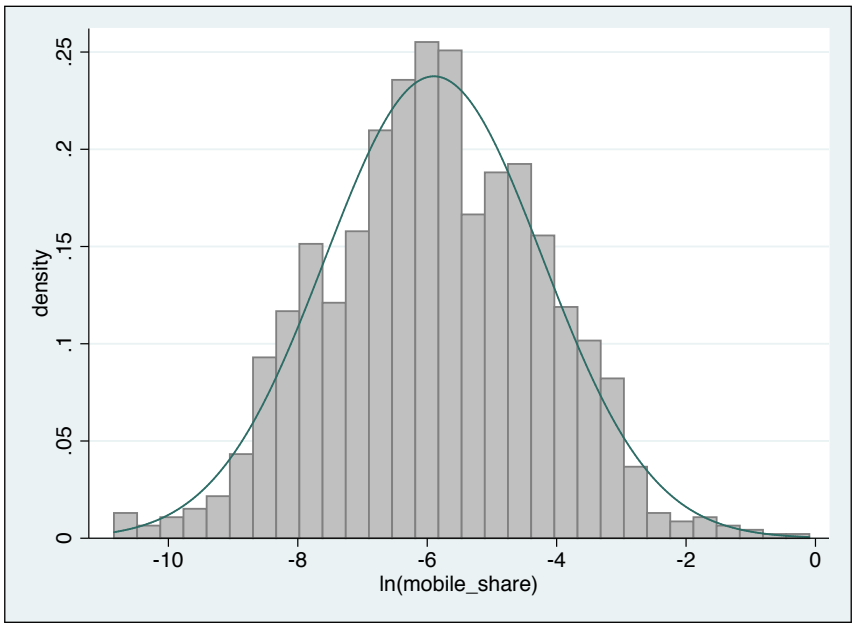

*Notes:* The figure displays shares of mobile employees in country pairs, separately for the three lines of services (LoS): *auditing* in Figure 1a, *taxation* in Figure 1b and *consulting* in Figure 1c. For each country, it depicts flows to the host countries to which mobile employees are sent (in the color of the country) and flows from the home countries from which it receives mobile employees (in the colors of the respective home countries). Flows are expressed relative to the overall number of mobile employees in each LoS. Country pairs are limited to the observations used in Section 3.2. We further drop country pairs whose share is below a threshold.

Figure 2: **Distribution of the Mobility Share**

(a) ***Raw***




(b) ***Logarithmic Transformation***




*Notes:* The figure displays the distribution of the mobility share for all lines of services (LoS) $k$, i.e., *auditing*, *taxation* and *consulting*. Figure 2a shows the distribution of $mobile\_share_{ijk}$, i.e., the ratio of the annual number of employees from LoS $k$ in home country $i$ working in host country $j$ to the number of employees from LoS $k$ in home country $i$, as average over the years 2005–2008. We multiply $mobile\_share_{ijk}$ by 100 for the graphical representation. Figure 2b shows the distribution of the natural logarithm of $mobile\_share_{ijk}$ and the normal density with the same mean and standard deviation. The number of observations is equal to 1,291.

# Tables

Table 1: **Summary Statistics**

| Variable | N | Mean | SD | Min | Median | Max |
|---|---|---|---|---|---|---|
| $mobile_{ijk}$ | 50,060 | 0.034 | | | | |
| *auditing* | 18,360 | 0.045 | | | | |
| *taxation* | 18,360 | 0.023 | | | | |
| *consulting* | 13,340 | 0.034 | | | | |
| $mobile\_share_{ijk}$ | 1,291 | 1.103 | 3.842 | 0.002 | 0.269 | 91.000 |
| *auditing* | 598 | 1.295 | 4.503 | 0.003 | 0.250 | 91.000 |
| *taxation* | 339 | 0.708 | 2.052 | 0.002 | 0.272 | 34.357 |
| *consulting* | 354 | 1.156 | 3.924 | 0.002 | 0.295 | 52.778 |
| $ln(mobile\_share)_{ijk}$ | 1,291 | −5.896 | 1.680 | −10.841 | −5.919 | −0.094 |
| $acc\_sim_{ij}$ | 1,291 | −0.295 | 0.312 | −1.000 | −0.250 | 0.000 |
| $tax\_sim_{ij}$ | 1,291 | −0.504 | 0.167 | −1.000 | −0.500 | −0.100 |
| $gdp\_high_{ij}$ | 50,060 | 0.490 | | | | |
| $taxhaven\_host_{ij}$ | 50,060 | 0.038 | | | | |
| $crosslisting_{ij}$ | 50,060 | 0.020 | | | | |
| $tradeflow_{ij}$ | 50,060 | 0.484 | | | | |
| $comlang_{ij}$ | 50,060 | 0.133 | | | | |
| $colony_{ij}$ | 50,060 | 0.015 | | | | |
| $distance\_high_{ij}$ | 50,060 | 0.485 | | | | |

*Notes:* The table displays summary statistics. $i$, $j$ and $k$ denote home country, host country and line of service (LoS), respectively. LoS is *auditing*, *taxation* or *consulting*. $mobile_{ijk}$ is an indicator variable equal to one if at least one individual from LoS $k$ in home country $i$ works in host country $j$ in the years 2005–2008. $mobile\_share_{ijk}$ [$ln(mobile\_share)_{ijk}$] is [the natural logarithm of] the ratio of the annual number of employees from LoS $k$ in home country $i$ working in host country $j$ to the number of employees from LoS $k$ in home country $i$, as average over the years 2005–2008. We multiply $mobile\_share_{ijk}$ by 100 for representation in the table. $acc\_sim_{ij}$ is accounting similarity following Bae et al. [2008]. $tax\_sim_{ij}$ is tax similarity based on data from the tax attractiveness index. $gdp\_high_{ij}$ is an indicator variable equal to one if the per capita GDP of host country $j$ is larger than the per capita GDP of home country $i$. $taxhaven\_host_{ij}$ is an indicator variable equal to one if host country $j$ is among the five closest tax havens for home country $i$. $crosslisting_{ij}$ is an indicator variable equal to one if firms from home country $i$ are cross-listed in host country $j$. $tradeflow_{ij}$ is an indicator variable equal to one if exports from home country $i$ to host country $j$ are above the median for home country $i$. $comlang_{ij}$ is an indicator variable equal to one if home country $i$ and host country $j$ share the same language. $colony_{ij}$ is an indicator variable equal to one if home country $i$ and host country $j$ have a common colonial history. $distance\_high_{ij}$ is an indicator variable equal to one if the distance between home country $i$ and host country $j$ is larger than the sample median.

Table 2: **Correlations**

| | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** | **(7)** | **(8)** | **(9)** | **(10)** | **(11)** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **(1)** $mobile_{ijk}$ | 1.000 | | | | $0.048^{***}$ | $0.040^{***}$ | $0.402^{***}$ | $0.164^{***}$ | $0.085^{***}$ | $0.174^{***}$ | $-0.038^{***}$ |
| **(2)** $ln(mobile\_share)_{ijk}$ | | 1.000 | 0.032 | 0.025 | $0.295^{***}$ | $-0.063^{**}$ | $0.188^{***}$ | $0.110^{***}$ | $0.146^{***}$ | $0.166^{***}$ | $-0.048^{*}$ |
| **(3)** $acc\_sim_{ij}$ | | 0.038 | 1.000 | $0.229^{***}$ | $-0.066^{**}$ | $-0.030$ | 0.042 | 0.005 | 0.009 | 0.041 | $-0.178^{***}$ |
| **(4)** $tax\_sim_{ij}$ | | 0.033 | $0.210^{***}$ | 1.000 | $-0.050^{*}$ | $-0.078^{***}$ | $0.067^{**}$ | $0.155^{***}$ | $-0.010$ | $-0.001$ | $-0.064^{**}$ |
| **(5)** $gdp\_high_{ij}$ | $0.048^{***}$ | $0.295^{***}$ | $-0.057^{**}$ | $-0.043$ | 1.000 | $0.058^{***}$ | $0.044^{***}$ | $0.205^{***}$ | $-0.003$ | 0.002 | 0.002 |
| **(6)** $taxhaven\_host_{ij}$ | $0.040^{***}$ | $-0.060^{**}$ | $-0.050^{*}$ | $-0.071^{**}$ | $0.058^{***}$ | 1.000 | $0.037^{***}$ | $0.030^{***}$ | $0.101^{***}$ | $-0.009^{**}$ | $-0.170^{***}$ |
| **(7)** $crosslisting_{ij}$ | $0.402^{***}$ | $0.182^{***}$ | $0.062^{**}$ | $0.068^{**}$ | $0.044^{***}$ | $0.037^{***}$ | 1.000 | $0.131^{***}$ | $0.061^{***}$ | $0.132^{***}$ | $-0.050^{***}$ |
| **(8)** $tradeflow_{ij}$ | $0.164^{***}$ | $0.111^{***}$ | $-0.079^{***}$ | $0.094^{***}$ | $0.205^{***}$ | $0.030^{***}$ | $0.131^{***}$ | 1.000 | $0.091^{***}$ | $0.081^{***}$ | $-0.161^{***}$ |
| **(9)** $comlang_{ij}$ | $0.085^{***}$ | $0.152^{***}$ | 0.045 | $-0.006$ | $-0.003$ | $0.101^{***}$ | $0.061^{***}$ | $0.091^{***}$ | 1.000 | $0.158^{***}$ | $-0.070^{***}$ |
| **(10)** $colony_{ij}$ | $0.174^{***}$ | $0.172^{***}$ | $0.048^{*}$ | $-0.006$ | 0.002 | $-0.009^{**}$ | $0.132^{***}$ | $0.081^{***}$ | $0.158^{***}$ | 1.000 | $-0.032^{***}$ |
| **(11)** $distance\_high_{ij}$ | $-0.038^{***}$ | $-0.041$ | $-0.133^{***}$ | $-0.068^{**}$ | 0.002 | $-0.170^{***}$ | $-0.050^{***}$ | $-0.161^{***}$ | $-0.070^{***}$ | $-0.032^{***}$ | 1.000 |

*Notes:* The table displays Pearson correlations in the lower and Spearman correlations in the upper triangle. $i$, $j$ and $k$ denote home country, host country and line of service (LoS), respectively. LoS is *auditing*, *taxation* or *consulting*. $mobile_{ijk}$ is an indicator variable equal to one if at least one individual from LoS $k$ in home country $i$ works in host country $j$ in the years 2005–2008. $ln(mobile\_share)_{ijk}$ is the natural logarithm of the ratio of the annual number of employees from LoS $k$ in home country $i$ working in host country $j$ to the number of employees from LoS $k$ in home country $i$, as average over the years 2005–2008. $acc\_sim_{ij}$ is accounting similarity following Bae et al. [2008]. $tax\_sim_{ij}$ is tax similarity based on data from the tax attractiveness index. $gdp\_high_{ij}$ is an indicator variable equal to one if the per capita GDP of host country $j$ is larger than the per capita GDP of home country $i$. $taxhaven\_host_{ij}$ is an indicator variable equal to one if host country $j$ is among the five closest tax havens for home country $i$. $crosslisting_{ij}$ is an indicator variable equal to one if firms from home country $i$ are cross-listed in host country $j$. $tradeflow_{ij}$ is an indicator variable equal to one if exports from home country $i$ to host country $j$ are above the median for home country $i$. $comlang_{ij}$ is an indicator variable equal to one if home country $i$ and host country $j$ share the same language. $colony_{ij}$ is an indicator variable equal to one if home country $i$ and host country $j$ have a common colonial history. $distance\_high_{ij}$ is an indicator variable equal to one if the distance between home country $i$ and host country $j$ is larger than the sample median. The number of observations is equal to 1,291 for all correlations with $ln(mobile\_share)_{ijk}$, $acc\_sim_{ij}$ and $tax\_sim_{ij}$; and is equal to 50,600 else. $^{*}$, $^{**}$ and $^{***}$ indicate significance at the 0.1, 0.05 and 0.01 level, respectively.

Table 3: **Existence of Temporary Migration**

| | $mobile_{ijk}$ | | |
|---|---|---|---|
| | *auditing* | *taxation* | *consulting* |
| $gdp\_high_{ij}$ | 0.013 | 0.006 | 0.006 |
| | (0.009) | (0.007) | (0.006) |
| $taxhaven\_host_{ij}$ | 0.033* | 0.036* | 0.016 |
| | (0.017) | (0.020) | (0.013) |
| $crosslisting_{ij}$ | 0.354*** | 0.327*** | 0.318*** |
| | (0.043) | (0.060) | (0.050) |
| $tradeflow_{ij}$ | 0.007* | 0.004* | 0.007** |
| | (0.004) | (0.002) | (0.003) |
| $comlang_{ij}$ | 0.026*** | 0.013** | 0.013 |
| | (0.008) | (0.005) | (0.008) |
| $colony_{ij}$ | 0.144*** | 0.074*** | 0.080*** |
| | (0.035) | (0.022) | (0.024) |
| $distance\_high_{ij}$ | −0.012** | −0.009*** | −0.016*** |
| | (0.005) | (0.003) | (0.005) |
| Home and host country fixed effects | yes | yes | yes |
| adj. $R^2$ | 0.341 | 0.260 | 0.292 |
| adj. $R^2$ within | 0.077 | 0.097 | 0.087 |
| $N$ | 18,360 | 18,360 | 13,340 |
| # home countries | 136 | 136 | 116 |
| # host countries | 136 | 136 | 116 |
| Largest fixed effects | | | |
| Home country | US, GB, DE | US, GB, NL | GB, US, DE |
| Host country | US, GB, CA | US, GB, AU | US, GB, AU |

*Notes:* The table displays the analysis of mobility determinants. $i$, $j$ and $k$ denote home country, host country and line of service (LoS), respectively. LoS is *auditing*, *taxation* or *consulting*. The sample encompasses all possible combinations of home and host countries, conditional on observing employees in LoS $k$. $mobile_{ijk}$ is an indicator variable equal to one if at least one individual from LoS $k$ in home country $i$ works in host country $j$ in the years 2005–2008. $gdp\_high_{ij}$ is an indicator variable equal to one if the per capita GDP of host country $j$ is larger than the per capita GDP of home country $i$. $taxhaven\_host_{ij}$ is an indicator variable equal to one if host country $j$ is among the five closest tax havens for home country $i$. $crosslisting_{ij}$ is an indicator variable equal to one if firms from home country $i$ are cross-listed in host country $j$. $tradeflow_{ij}$ is an indicator variable equal to one if exports from home country $i$ to host country $j$ are above the median for home country $i$. $comlang_{ij}$ is an indicator variable equal to one if home country $i$ and host country $j$ share the same language. $colony_{ij}$ is an indicator variable equal to one if home country $i$ and host country $j$ have a common colonial history. $distance\_high_{ij}$ is an indicator variable equal to one if the distance between home country $i$ and host country $j$ is larger than the sample median. Standard errors are two-way clustered by home country $i$ and host country $j$ and denoted below the coefficients in parentheses. *, ** and *** indicate significance at the 0.1, 0.05 and 0.01 level, respectively.

Table 4: **Extent of Temporary Migration**

| | $ln(mobile\_share)_{ijk}$ | | | | | |
|---|---|---|---|---|---|---|
| | *auditing* | | *taxation* | | *consulting* | |
| $acc\_sim_{ij}$ | 0.757*** | 0.800*** | 0.299 | 0.401* | −0.090 | −0.199 |
| | (0.239) | (0.246) | (0.251) | (0.231) | (0.503) | (0.510) |
| $tax\_sim_{ij}$ | 0.630 | 0.370 | 0.998*** | 0.495 | 0.953* | 1.042** |
| | (0.465) | (0.404) | (0.328) | (0.394) | (0.536) | (0.508) |
| $gdp\_high_{ij}$ | 0.060 | −0.004 | 0.217 | 0.147 | 0.181* | 0.081 |
| | (0.168) | (0.157) | (0.156) | (0.142) | (0.097) | (0.108) |
| $taxhaven\_host_{ij}$ | | −0.134 | | 0.052 | | −0.944*** |
| | | (0.197) | | (0.295) | | (0.246) |
| $crosslisting_{ij}$ | | 0.247* | | 0.111 | | 0.023 |
| | | (0.135) | | (0.201) | | (0.172) |
| $tradeflow_{ij}$ | | −0.254 | | −0.105 | | −0.274 |
| | | (0.266) | | (0.135) | | (1.136) |
| $comlang_{ij}$ | | 0.553*** | | 0.642*** | | 0.481** |
| | | (0.182) | | (0.223) | | (0.197) |
| $colony_{ij}$ | | 0.870*** | | 0.482** | | 0.308 |
| | | (0.162) | | (0.188) | | (0.245) |
| $distance\_high_{ij}$ | | −0.254 | | −0.284 | | 0.052 |
| | | (0.160) | | (0.169) | | (0.150) |
| Home and host country fixed effects | yes | yes | yes | yes | yes | yes |
| | $acc\_sim_{ij}$ | | $tax\_sim_{ij}$ | | $gdp\_high_{ij}$ | |
| Δ *coefficients* | 0.592** | | −0.151 | | 0.060 | |
| | (0.277) | | (0.406) | | (0.155) | |
| adj. $R^2$ | 0.614 | 0.659 | 0.602 | 0.643 | 0.652 | 0.670 |
| adj. $R^2$ within | 0.019 | 0.144 | 0.036 | 0.157 | 0.022 | 0.092 |
| $N$ | 598 | | 339 | | 354 | |
| # home countries | 62 | | 47 | | 51 | |
| # host countries | 44 | | 35 | | 37 | |
| Largest fixed effects (incl. controls) | | | | | | |
| Home country | MK, ZW, HR | | SI, KE, NZ | | ZW, TH, LT | |
| Host country | US, GB, CN | | US, GB, CN | | GB, US, CN | |

*Notes:* The table displays the analysis of mobility determinants. $i$, $j$ and $k$ denote home country, host country and line of service (LoS), respectively. LoS is *auditing*, *taxation* or *consulting*. The sample encompasses all combinations of home and host countries for which we observe mobility in the years 2005–2008 with available data for the similarity scores. $ln(mobile\_share)_{ijk}$ is the natural logarithm of the ratio of the annual number of employees from LoS $k$ in home country $i$ working in host country $j$ to the number of employees from LoS $k$ in home country $i$, as average over the years 2005–2008. $acc\_sim_{ij}$ is accounting similarity following Bae et al. [2008]. $tax\_sim_{ij}$ is tax similarity based on data from the tax attractiveness index. $gdp\_high_{ij}$ is an indicator variable equal to one if the per capita GDP of host country $j$ is larger than the per capita GDP of home country $i$. $taxhaven\_host_{ij}$ is an indicator variable equal to one if host country $j$ is among the five closest tax havens for home country $i$. $crosslisting_{ij}$ is an indicator variable equal to one if firms from home country $i$ are cross-listed in host country $j$. $tradeflow_{ij}$ is an indicator variable equal to one if exports from home country $i$ to host country $j$ are above the median for home country $i$. $comlang_{ij}$ is an indicator variable equal to one if home country $i$ and host country $j$ share the same language. $colony_{ij}$ is an indicator variable equal to one if home country $i$ and host country $j$ have a common colonial history. $distance\_high_{ij}$ is an indicator variable equal to one if the distance between home country $i$ and host country $j$ is larger than the sample median. Standard errors are two-way clustered by home country $i$ and host country $j$ and denoted below the coefficients in parentheses. *, ** and *** indicate significance at the 0.1, 0.05 and 0.01 level, respectively.