Essays in Financial and Applied Economics

by

Yuriy Podvysotskiy

B.A., International Economics, Khmelnytskiy State University, 2004

M.A., Economics, Kyiv School of Economics, 2006

M.A., Economics, University of Pittsburgh, 2017

Submitted to the Graduate Faculty of

the Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2020

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ART AND SCIENCES

This dissertation was presented

by

Yuriy Podvysotskiy

It was defended on

November 23, 2020

and approved by

Daniele Coen-Pirani, Department of Economics, University of Pittsburgh

Osea Giuntella, Department of Economics, University of Pittsburgh

Daniel Berkowitz, Department of Economics, University of Pittsburgh

Marla Ripoll, Department of Economics, University of Pittsburgh

David J. Denis, Katz Graduate School of Business, University of Pittsburgh

Essays in Financial and Applied Economics

Yuriy Podvysotskiy, PhD

University of Pittsburgh, 2020

This dissertation consists of three independent essays in financial and applied economics. Essay One analyzes the impact of common ownership concentration on corporate payouts, investments, and markups. The evaluated hypothesis is that a set of investors who own significant equity stakes in two or more firms within the same industry (common ownership) are able to decrease competition, raise markups, and increase payouts. The results confirm a statistically significant increase in corporate payouts following an increase in industry-level common ownership. I further show that the effect of common ownership concentration on payouts is larger in industries that face relatively less competition from Chinese imports, and is also larger in industries with low values of their Herfindahl indices.

Essay Two estimates the implications of divorce between individuals aged 50 and above ('gray divorce') for all components of their wealth. The results suggest that gray divorce negatively impacts net worth, especially components such as housing equity and financial assets. There is no evidence of higher decline in net worth for females as compared to males. However, divorcing females experience higher decline of their individual retirement account balances, and, are more likely to re-enter the labor force. Wealth is also shown to be a key (negative) predictor of the probability of gray divorce.

Essay Three tests the so-called 'modernization hypothesis', a positive effect of income on democracy, examining the case of the Central and Eastern Europe (CEE) after communism's collapse. I show that a highly important influence on the relation between income and democracy in some post-socialist country is its initial political disruption: time between collapse of the communist system and emergence of a new post-communist government. In particular, I verify that the positive relation between income and democracy is significantly weakened, or even reversed, for the countries that experienced prolonged periods of initial political disruption.

Table of Contents

List of Tables

List of Figures

Preface

My PhD journey would not be possible without guidance and help that some other people generously provided to me. Therefore, I would like to thank my advisors and committee for their feedback and support on the way of writing this thesis, that each provided in their special way. I am grateful to Daniele Coen-Pirani for supporting me across all the stages of this PhD journey. His support and advice were always very helpful, timely and to the point, so that finally this made the real difference. I am thankful to Osea Giuntella for encouraging to learn new methods and attempt with excitement new tasks however difficult they might seem. I was highly motivated and excited by the interest in the subject and connections between Economics science and real life that Marla Ripoll and Daniel Berkowitz generously shared with me. Also, I am thankful to David Denis for his deep expertise, encouragement to think in a structured manner especially when designing tests and interpreting their results, and for the opportunity to learn financial economics. In addition, my family provided enormous support as well, for which I am grateful to them. My dear wife Toma proved a great discussion board to test different research ideas and besides Toma, as well as my daughters Mary and Olena, showed great patience and understanding regarding my spending much time working and having less time together. They along with our lovely pom Ellie were very effective in keeping my spirit high especially when not everything worked-out well. Special thanks go to my mother Mary and brother Alex who also provided invaluable help and support during these PhD years. I also thank to all the friends, whom I did not mention above, but who helped by being a decent company.

1.0 Introduction

This thesis includes three independent essays, one of which explores the role of common ownership concentration for corporate payouts, investments and competition. The second essay investigates implications of marital instability among those aged 50+ in regard to all the components of their wealth. The third essay analyzes the nature of the relation between income and democracy, using the case of post-socialist European economies. The three essays are unrelated to each another, yet all feature a similar research approach, which is based on application of econometric methods and relevant identification strategies.

Each of the considered topics is important in their own way. In relation to Essay One, diversified institutional investors jointly hold 70-80% of the US stock market, resulting in shareholder structures with high density of common ownership (Azar et al. 2016). For example, the same few institutional investors, Vanguard, BlackRock, Fidelity, and State Street, each hold sizable equity stakes of 2 - 7% in key companies across many industries, including technology, pharmaceuticals, banks and airlines (Azar 2018). To date, researchers and regulators still debate the common ownership potential to monopolize the US economy, and also decline capital and R&D investments (Gutierrez and Philippon 2016). Another vital phenomenon, which motivated Essay Two, is that of 'gray divorce'. Between 1990 and 2015, divorce rates among the population aged 50+ more than doubled. Marital dissolution between seniors is known to result in deterioration of mental and physical health, and, increase social exclusion (Del Bono et al. 2007; Luo et al. 2012). Moreover, the population aged 50+ comprises over 1/3 of the total US population, and, has the highest median savings in comparison to all other population age groups. Therefore, gray divorce could possibly affect negatively wealth among the 'gray population' category. Essay Three addresses the importance of the relationship between income and democracy. During the past several years, multiple developed and developing countries with previously high democratization records marked tendency to move towards right-wing populist policies (FreedomHouse 2017). The study of the case of the Eastern European countries who recently underwent the 'natural experiment' of regime change could uncover further insights about the relation between

income and democracy, and the role of the initial political disruption for this.

The remaining part of my thesis is organized as follows: Chapter 2 analyzes the impact of common ownership concentration on corporate payouts, investments, and markups. To establish causality, I utilize a difference-in-differences method based on a major mutual fund scandal. Also, as a test of the mechanism, I compare the effect of common ownership concentration on payouts and investments between industries with relatively high and relatively low competition from Chinese imports. Chapter 3 explores the impact of gray divorce on wealth. Examined categories include financial wealth, value of vehicles, and housing equity. Semi-parametric difference-in-differences methodology is applied to ensure that individuals from 'treated' and 'control' groups are similar in all key respects, beyond the incidence of divorce. The chapter then uses machine learning to study the main predictors of future divorce among the population aged 50+. Chapter 4 analyses the case of post-socialist European economies and the relation between income and democracy for these countries. following the approach in Acemoglu et al. (2008), I use trade-weighted world income as an instrumental variable to deal with reverse causality between income and democracy. The chapter suggests a method to measure initial political disruption, and includes it as a moderation variable in the regression of democracy on income. Finally, Chapter 5 summarizes the main findings.

2.0 Common Ownership Concentration and Corporate Payouts

This essay analyzes the impact of common ownership concentration on corporate payouts, investments, and markups. I hypothesize that when a set of investors owns significant equity stakes in two or more firms within the same industry (common ownership), this concentration of common ownership decreases competition, raises markups, and increases payouts. Consistent with my hypothesis, I find a statistically significant increase in corporate payouts following an increase in industry-level common ownership. To establish causation, I rely on the use of difference-in-differences method based on a major mutual fund scandal. I further show that the effect of common ownership concentration on payouts is larger in industries that face relatively less competition from Chinese imports, as well as in industries with low Herfindahl indices.

2.1 Introduction

Common ownership is a situation when a set of investors jointly own a significant stake in a number of firms within the same industry. My study explores whether concentration of common ownership influences corporate payouts. Therefore, this analysis connects the two hotly disputed areas within the literature, namely that of common ownership concentration and its negative effect on competition among firms (e.g. Azar 2018; Dennis et al. 2018; Koch et al. 2019), and that about concerns regarding growing corporate payouts raising concerns about excessive withdrawal of cash from the business that potentially limits its available resources to invest and innovate (e.g. Fried and Wang 2018; Lazonick 2014).

The analysis in this study is performed in connection to the hypothesized negative effect of common ownership on competition. Firms that operate in industries with high concentration of common ownership are argued to face reduced incentives to compete (Farrell 1985; Rotemberg 1984). There are two points of view in the ongoing debate about the anti-competitive influences of concentrated common ownership. Some studies favored the

anti-competitive effect, including Azar (2018) for the case of airline industry, Azar et al. (2016) for banking sector and Gutierrez and Philippon (2018) for a multi-industry set-up. These researchers observed positive implications of common ownership for markups and prices, and a negative effect on capital investments. Other studies did not confirm the aforementioned negative impact of common ownership on competition, including Dennis et al. (2018) and Koch et al. (2019). Therefore, the dispute regarding the implications of common ownership for competition among firms is far from settled.

The concept of common ownership should be clearly distinguished from the related concept of institutional ownership. Common ownership relates to a situation when two or more firms are partly or fully owned by the same investor (group of investors) (Gilje et al. 2018). Recent decades evidenced an increase in common ownership concentration due to the increase in size and influence of institutional investors that are pursuing portfolio diversification equity investment strategies. Higher shareholdings by institutional investors with portfolio diversification strategies resulted in a situation with a number of large institutional investors commonly owning significant equity stakes in key market players within same industries (Azar 2018) [1].

In the current study, I build on the view that common ownership reduces competition. The link from increased common ownership concentration to lower competition and to increased payouts goes as follows: when common ownership concentration increases in a given industry, this lowers incentives to compete among the industry players because common owners do not have incentives to promote competition (Azar 2018). This lack of competition usually results in higher profits and cash flows due to benefits of higher market power. Sound governance of common owners (institutional investors) would lead to withdrawal of increased cash as payouts (Chung and Zhang 2011). But the latter step is not a mechanic identity, as the increased cash resources available to the firms with concentrated common ownership can also be used for alternative purposes, such as financing acquisitions, making debt repay-

[1] Investment Companies Act of 1940 sets restrictions for some types of investment companies. Specifically, Sec 5(a)(1) defines diversified companies (most mutual funds) to face a 5% upper limit for 75 percent of their assets, and besides to be restricted from acquiring more than 10% of voting stocks in a single company. In addition, under Sec 12(d)(1)(A) of the Investment Companies Act, such companies cannot acquire more than 3% of equity in an investing company, cannot spend more than 5% of their assets on stocks of another investing company, and cannot spend more than 10% of their assets on stocks of several investing companies.

ment, or building-up cash piles. In fact, Matvos and Ostrovsky (2008) evidenced the higher propensity to vote for acquisition by firms with greater common ownership. A number of previous studies showed that payouts can significantly differ across firms, depending on their liquidity constraints, degree of financial leverage, as well as life-cycle stage, among other relevant factors (Denis and Osobov 2008) [2].

Positive association between common ownership concentration and corporate payouts may be due to other mechanisms, instead of the aforementioned anti-competitive effect of the common ownership. For example, positive association between common ownership concentration and corporate payouts could be related to the positive association between institutional ownership and corporate payouts. Crane et al. (2016) evidenced that an increase in institutional ownership usually resulted in higher payouts. Also Grinstein and Michaely (2005) showed that institutional investors are attracted to and thus tend to invest more in the firms that provide generous payouts. Such possible alternative mechanisms are taken into consideration, and are controlled for, in the empirical design of the current study.

Obtaining an answer to the mentioned research question would contribute towards evaluation of the nature of common ownership concentration and its broader economic implications. this can provide further insights into the driving forces behind the corporate payout policy and thus contribute to the ongoing debate about the nature of corporate buybacks[3].

The body of literature in the realm of anti-competitive influences of concentrated common ownership is a rapidly growing one. The possibility of anti-competitive effect of concentrated common ownership was explored and confirmed by Azar (2018) in the case of airline industry[4], and also by Azar et al. (2016) in the case of deposit banking industry. Gutiérrez and Philippon (2017) included common ownership in a synthetic measure of total

[2]SEC Rule 10b-18 sets limit for the open market share repurchases at 25% of the average daily volume. The rule is not mandatory and serves as a liability protection for companies ('safe harbor rule'). If repurchases are implemented in excess of the mentioned rule, company management is not protected for prosecution for fraudulent stock market activities

[3]In the mentioned heated debate about corporate payouts, some have argued that increase in share buybacks by the US corporations can lead towards declining investments and weaker growth (Lazonick (2014)). While the opponents, such as Fried and Wang (2018) countered this view and argued that observed increased payouts are relevant only to large mature firms, mostly constituents of S&P index, which are only a small fraction of all the US firms. Whereas investments for all other firms were on the rise and their cash balances were increasing.

[4]For a detailed overview of recent developments in the common ownership literature refer to Schmalz (2018).

concentration (Modified HHI) to argue about declining domestic competition and explore its effect on investments. They concluded that overall market concentration generally led to higher mark-ups and lower capital investments economy-wide. A somewhat different view was expressed by He and Huang (2017), who explored implications of common ownership for the same industry firms and concluded that it facilitates collaboration and coordination via forming strategic alliances and joint ventures. In addition, there are studies that deny the mechanism that common ownership decreases competition. For example, Dennis et al. (2018) replicated the study by Azar (2018), but did not confirm the mentioned findings for airline industry. Also Koch et al. (2019) investigated the causal effect of common ownership on investments and product market competition on the industry-by-industry basis, and for the majority of industries they could not obtain such evidence. Again, it can be inferred that the link between common ownership and declining competition is far from being fully established. Moreover, neither of the mentioned studies explicitly explored the causal impact of common ownership on corporate payouts. According to my knowledge, available literature has lacked causal analysis of common ownership effect on corporate payouts, so this study aims to fill this research gap. Analysis of the effect of common ownership concentration on corporate payouts is of importance as such, but it also provides a rigorous opportunity to test the anti-competitive effect of common ownership. Such opportunity relates to the fact that a number of other alternative mechanisms exist that underlie the positive correlation between common ownership and payouts. Therefore, the current study conducts several tests of the mechanism. Specifically, it explores whether an external source of competition that is beyond the control of the US common owners, such as that of the 'China shock', lowers the extent that common ownership affects payouts and investments. Also, the study investigates whether the effect of common ownership concentration on payouts is stronger in industries with low concentration, where there is more room for the common owners to facilitate decrease in the competition and foster consolidation.

The remaining parts of the study are organized as follows: the common ownership hypothesis in relation to competition and corporate payouts is discussed in Section 2. Then Section 3 lists the sources of data, describes sample selection procedures, explains calculation of the key variables. Section 4 introduces the research methodology and identification

strategy. The main findings of the study are presented in Section 5. The baseline findings are verified by the tests of the mechanisms in Section 6. Finally, Section 7 concludes with the summary of the key findings and some suggestions regarding venues for the future research.

2.2 The Common Ownership Hypothesis

The hypothesis aims to evaluate the effect of common ownership on payouts via its anti-competitive effect. The main hypothesis:

- *Common ownership concentration in the given industry causes higher corporate payouts in that industry.*

This hypothesis views positive effect of common ownership on payouts as a part of the anti-competitive impact of common ownership. In industries with higher common ownership there are lower incentives to compete, and higher efforts to coordinate and cooperate (Azar 2018, He and Huang 2017). This would lead to higher market power of the commonly owned firms. With the higher market power, the firms in the industries with concentrated common ownership can earn higher profits and generate higher cash flows. Then, because of high monitoring capacity of the common owners who are institutional investors, they can facilitate withdrawal of excessive cash in the form of payouts, in order to mitigate the agency costs of free cash flows (Jensen and Meckling 1976). In addition to the explained above mechanism of common ownership affect on payouts, there can also be an impact of common ownership on capital investments that goes as follows: the optimum level of output for a more concentrated industry is lower than for a similar industry with lower level of concentration (Gutiérrez and Philippon 2017). Hence the required production capacity decreases, the level of capital investments can go down in an industry with concentrated common ownership.

The current study allows for other explanations of positive association between common ownership concentration and corporate payouts. Under such alternative hypothesis, common ownership concentration would be correlated with, but not causing, higher payouts. These

other mechanisms rely on positive correlation between common ownership and institutional ownership, and then explain the link between institutional ownership and payouts. The two such alternative channels are the effect of payouts on institutional ownership, and the effect of institutional ownership on payouts.

The first alternative mechanism is based on the evidence that institutional ownership influences payouts. Crane et al. (2016) determined that increase in institutional ownership positively affected dividend payments. This is also related to earlier studies regarding the role of institutional owners, as well as that of debt holders, to reduce the agency costs of free cash flows (Jensen and Meckling 1976). The study by Crane et al. (2016) relied on exogenous shocks to institutional ownership due to transitions between Russell 2000 and Russell 1000 indices. Limitation of this identification strategy is that transitions between the two indices involves only a small fraction of all listed firms (Almeida et al. 2018; Gutierrez and Philippon 2018). Therefore, presumably positive association between common ownership and payouts could be driven by the effect of institutional ownership alone, and not by the anti-competitive effect of common ownership.

The second alternative is based on the fact that institutional investors prefer mature companies that are generous dividend-payers and that also implement regular buybacks. The mentioned causal relation between institutional ownership and corporate payouts was explored by Grinstein and Michaely (2005). The researchers used panel vector auto-regression methodology and concluded that institutions are attracted by dividend paying firms, and moreover, by those who regularly implement share repurchases. But no evidence was obtained by Grinstein and Michaely (2005) in support of the hypothesis that institutional ownership concentration influences payout policies.

To allow for the above mentioned alternative explanations, the regression analysis in this study includes institutional ownership among the covariates. In addition, one of the tests relies on exogenous shock to common ownership in order to establish causality using difference-in-differences method. Several tests of the anti-competitive mechanism of the common ownership are performed also. The first such test explores whether exposure to the 'China shock' lowers the effect of common ownership concentration on payouts and capital investments. The second test of the mechanism investigates whether the effect of common

ownership on payouts and investments differs between industries with different levels of Herfindahl index.

2.3 Data, Sampling and Variables Construction

2.3.1 Data and sampling

The data for the study was obtained from the following databases. Financial statements data was obtained from Capital IQ, institutional ownership information - from Thomson Reuters 13F[5], and the data on the imports and total sales by industry - from Bureau of Economic Analysis database.

The initial sample includes all public companies in the US for the time period between 2000 and 2018. Thus the focus is specifically on the period when the decrease in the domestic competition was taking place, according to Gutiérrez and Philippon (2017). Industries were defined based on 3-digit codes of standard industry classification (SIC3 codes). Following standard approach in the finance research (e.g. Chung and Zhang (2011); Denis and Osobov (2008); Koch et al. (2018)) the database was filtered to exclude financial services companies (SIC codes between 6000 and 6999) and those from utilities (SIC codes between 4900 and 4949). Furthermore, the firms with annual sales less than $0.25 million, and total assets less than $1.0 million were excluded. Also, the listed firms with sales revenue that are lower than their EBIT (operating profit) were omitted. All variables were winsorized at 1% and 99% in order to eliminate possible contaminating effect of outliers, following He and Huang (2017). The resulting sample included data on over 10,350 corporations.

2.3.2 Variables construction

The study considers several measures of common ownership concentration, which are mostly industry-level variables. These industry-level measures include Density of common

[5]Thomson Reuters 13F database is based on the data from mandatory quarterly 13F forms that all US institutional investors, with at least $100 million in assets under management, are required to file to the US Securities and Exchange Commission.

ownership (DCO), Percentage of common funds (PCF), Percentage of common stocks (PCS), as well as delta HHI (ΔHHI) [6]. Besides, there are two company-level measures, namely the dummy for cross-owned firms (CrossDummy) and the number of the same industry rivals that are commonly owned by at least one same institutional owner (NumConnected). Definitions of the firm-specific indicators follow He and Huang (2017).

Ideally, these measures should be applied to individual markets and by considering all investors into the market players. Practical implementation is limited to consider institutional investors only, and also is applied using grouping of firms into industries, using SIC3 industries classification, rather than markets (Schmalz (2018)).

The current study performed simulation of the ways that main industry-level common ownership indicators (DCO, PCF, PCS and ΔHHI) respond to changes in the number of market players, common owners and ownership shares by common owners (*Appendix A.3*). The results of such simulation indicated that while neither of the considered measures of common ownership is fully compelling, there is relatively better performance of DCO measure in comparison to the other considered ones.

Density of common ownership (DCO) is the ratio of connected firm pairs to total number of firm pairs. DCO measures the density of the network of firms within the same industry, and it is an average level of connections within the industry. Connected firms pair has at least one common institutional investors with ownership share of at least 5% in each of the two firms in the pair. This definition follows the use of DCO in Azar (2011) and Koch et al. (2018). In equation below I_{ij} is equal to 1 for each connected pair of firms and to 0 for a not connected ones.

$$DCO = \frac{\Sigma_{i=1}^{n}\Sigma_{j<i}I_{ij}}{n(n-1)/2},\ where\ I_{ij} \in \{0,1\} \tag{1}$$

Percentage of common funds (PCF) is the ratio of the number of institutional investors who own at least two stocks in the given industry to the total number of institutional investors in that industry. The considered ownership share threshold is 5%, following the use of PCF in Anton and Polk (2014) and Koch et al. (2018). The equation below indicates that there

[6] Gilje et al. (2018) implement review of alternative measures of common ownership concentration and their use in the literature. There is no straightforward consensus evidenced as for the choice of some most preferred measure of the common ownership.

are K institutional investors in the given industry. Also, $|\omega_k|$ denotes the size (cardinality) of the set consisting of all the companies in the given industry that are commonly owned by investor k.

$$PCF = \frac{\Sigma_{k=1}^{K} I_{|\omega_k|>1}}{K}, \ where \ |\omega_k| \geq 0 \tag{2}$$

Percentage of common stocks (PCS) is the ratio of the number of stocks that are commonly owned with at least one other stock in the same industry to the total number of stocks in that industry. The considered threshold for ownership share is 5%, following Koch et al. (2018). In the expression below, N is the total number of firms in the industry. Also, $|J|$ is the size (cardinality) of the set of the other firms in the industry that are commonly owned with the firm i.

$$PCS = \frac{\Sigma_{i=1}^{N} I_{|J|\geq 1}}{N}, \ where \ |J| \geq 0 \tag{3}$$

Delta HHI index (ΔHHI) is an indicator that was designed as an extension to the standard HHI (Herfindahl index), aiming to capture concentration among owners (Azar 2018). Only ownership shares for institutions with at least 0.5% are considered in the calculation of ΔHHI, following Koch et al. (2019). The indicator is well-rooted in theory, but it has relatively limited practical applications (Schmalz 2018).

$$\Delta HHI = \Sigma_j \Sigma_{k \neq j} s_j s_j \frac{\Sigma_i \gamma_{ij} \beta_{ik}}{\Sigma_i \gamma_{ij} \beta_{ij}} \tag{4}$$

Where, s_j and s_k are market shares (calculated based on revenue) of firms j and k, respectively. Also, γ_{ij} indicates the percent share of control by shareholder i in firm j, and β_{ik} represents the percent share of ownership by shareholder i in firm k. For practical consideration, for the calculation it is assumed that the two shares (control and ownership) are equal, namely that $\gamma_{ij} = \beta_{ij}$ (Schmalz 2018).

Despite its theoretical appeal and connection to HHI, ΔHHI has some major limitations. One limitation was pointed out by Koch et al. (2019) - a merger between two common owners would lead to a decrease in ΔHHI, rather than to an increase that could be reasonably expected. Simulations that were performed in the current study (Appendix A.3), indicate

that mergers among market payers can also lead to decrease in ΔHHI, which is also a counter-intuitive behavior. Such surprising results are in line with the criticism by Schmalz (2018) who claimed that empirical qualities of ΔHHI are not well studied, unlike those of traditional HHI. While HHI is limited at 10,000 for a perfect monopoly, the upper limit of ΔHHI is not well defined, often can reach values in excess of millions, and this prevents its use for the purpose of antitrust regulation and oversight (Schmalz 2018) [7]. A simple test, namely evaluation of possible change in a measure of common ownership due to simulated merger between institutions, can be used for validating performance of different common ownership concentration indicators. Unlike for ΔHHI, in case of DCO, PCF and PCS, a merger between two block holders is likely to increase value of the measures of common ownership, which is in line with the expectations.

As additional measures, few firm-level indicators are considered. *CrossDummy* is equal to one for the firm in the given year, when there is at least one more firm in the same industry that is commonly owned by one or more institutional investors. The considered ownership threshold is that of 5%. NumConnect shows the number of other firms, which are commonly owned with the given firm i in the same industry and in the given year. The same 5% ownership threshold is applied. 'Commonly owned' denotes that there is at least one institutional investor that has ownership share in the given firm i and in at least one more company in the same industry.

2.3.3 Descriptive statistics

Summary statistics for common ownership concentration, payouts, and control variables are in Tables 2-1 and 2-2. For the total sample, the results indicate that among firms with positive payouts, the mean payouts to total assets ratio is 5.82%, there are 27% of dividend-paying firms-year observations, and 31% of observations with positive net share buybacks. On average, there are 101 firms per industry. Also, 'standard' concentration measured by Herfindahl index has the mean of 1,577, and ranges between 0 and 10,000.

[7] "Horizontal Merger Guidelines" (08/19/2010) by the Department of Justice prescribe that mergers leading to an increase in the market HHI by 100-200 points typically warrant scrutiny.

Table 2.1: Descriptive statistics for the whole sample

	Mean	St Dev	Min	Max	N
DCO	0.14	0.16	0.00	1.00	110,382
CrossDummy	0.098	0.297	0.00	1.00	112,012
NumConnect	2.35	13.07	0.00	186	112,012
PCF	0.049	0.034	0	0.25	110,382
PCS	0.748	0.257	0	1.00	110,382
Payouts[1]	5.82	15.43	0.00	1261.6	112,262
Dividends	4.55	16.01	0.00	1261.2	112,262
Buybacks	4.37	10.50	0.00	461.7	106,972
M&A	7.18	11.26	0.00	162.42	102,528
Cash	19.20	22.71	-0.33	100.0	106,964
Total assets	4,027	20,003	1.0	798,000	106,972
Markup	105.96	486.52	0.13	134,450	106,499
Capex	6.15	16.79	-43.60	3696.76	105,852
Leverage	33.49	790.09	-41900	182000	106,646
EBIT	-5.91	47.73	-2369.55	1326.5	106,480
Tobin's q	1.79	7.56	<0.001	1384.8	89,295
HHI	1576.9	1503.3	0.00	10000	112,012
No. firms (sic3)	101.2	130.2	1.00	564	110,382
			Proportions:		
Payouts > 0			44.3%		
Dividends > 0			26.7%		
Buybacks > 0			31.2%		
M&A > 0			38.6%		
Net Income < 0			40.3%		

Sample covers all public firms from 2000 till 2017. SIC3 industry classification is used. For Payouts, Dividends, Buybacks, and M&A: means and standard deviations are calculated for positive values.
[1] Payouts, Dividends, Buybacks, Capex, Cash and M&A are scaled by Total Assets.

Comparison was performed for sub-group of firms by low and high degree of common ownership in the industry. Where low (high) was defined as, respectively, values of DCO below the 1^{st} quartile and above the 3^{rd} quartile. Comparison of the results indicate that the firms in industries with high DCO provide relatively payouts, higher buybacks, and tend to invest less. Moreover, these firms, on average, have higher cash balances. Among the firms in industries with high ownership concentration, there is 9.1 percentage points higher propensity to do share buybacks.

These results cannot used to infer that common ownership concentration lowers investment and drives up buybacks. This is because the compared mean differences can be jointly determined by other variables. There can be selection bias, as common ownership concentration can be driven by higher payouts in these industries. Therefore, the analysis in the following sections is provided to evaluate the causal impact of common ownership concentration on payouts.

Table 2.2: Descriptive statistics by level of DCO

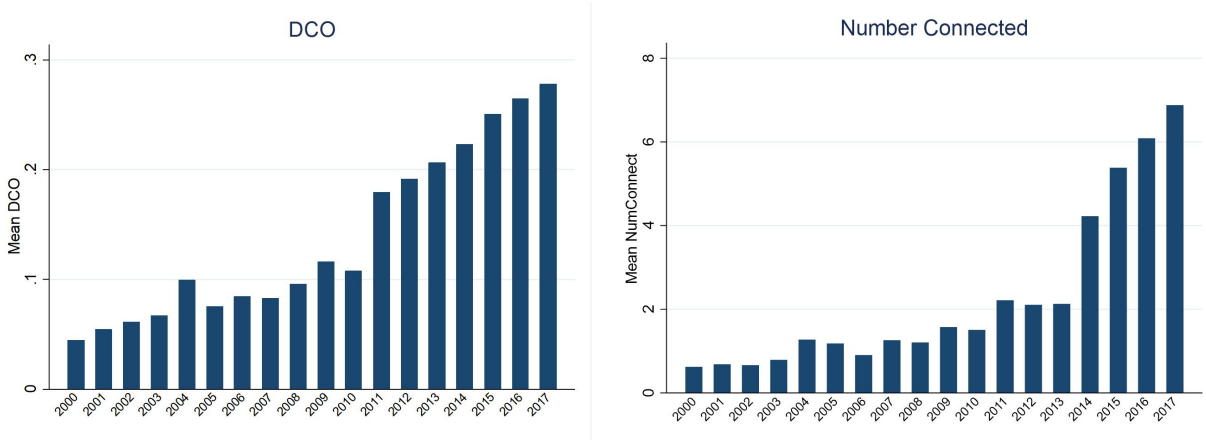
	Low DCO			High DCO			Low DCO – High DCO
	Mean	St Dev	N	Mean	St Dev	N	
DCO	0.018	0.014	27,678	0.347	0.195	27,723	-0.33 ***
CrossDummy	0.089	0.285	27,678	0.114	0.318	27,723	-0.025 ***
NumConnect	1.545	8.691	27,678	3.661	18.427	27,723	-2.116 ***
PCF	0.037	0.041	27,678	0.050	0.029	27,723	-0.012***
PCS	0.496	0.348	27,678	0.898	0.092	27,723	-0.402***
Assets	5,272	24,816	26,395	3,816	16,138	26,737	1,455 ***
Payouts [1]	5.45	14.18	27,678	5.82	18.05	27,723	-0.36 **
Dividends	4.56	12.55	27,678	4.19	20.57	27,723	0.37*
Buybacks	3.86	12.25	26,395	4.33	8.67	26,737	-0.47 ***
M&A	7.23	11.47	24,927	6.87	10.81	25,911	0.36**
Cash	16.49	20.66	26,393	21.19	23.47	26,733	-4.69***
Markup	117.21	937.7	26,171	105.13	266.5	26,677	12.07 **
Capex	7.81	12.28	25,895	4.50	6.16	26,594	3.31 ***
Leverage	34.24	628.03	26,339	32.04	311.46	26,641	2.19
EBIT/ TA	-6.16	49.85	26,165	-5.48	48.12	26,676	-0.68
Tobin's Q	1.59	5.02	21,189	1.86	11.11	22,815	-0.26 ***
HHI (sic3)	1,794	1,684	27,678	1,726	1,665	27,723	67.48 ***
No. firms (sic3)	104.7	163.0	27,678	82.9	111.4	27,723	21.74 ***
	Proportions:						
Payouts > 0		43.4%			49.8%		-6.4%***
Dividends > 0		28.2%			30.8%		-2.6%***
Buybacks > 0		27.9%			37.1%		-9.1% ***
M&A > 0		39.0%			38.4%		0.6% *
Net Income < 0		40.6%			38.4%		2.2% ***

Low DCO and High DCO, respectively, include firms from the industries with DCO below Q_1 and above Q_3. For Payouts, Dividends, Buybacks, M&A, their means and standard deviations are calculated for positive values. In the lower part of the table, proportions of positive values are provided. Significance levels: *** p<0.01, ** p<0.05, * p<0.1.

[1] Payouts, Dividends, Buybacks, Capex and M&A are scaled by Total Assets.

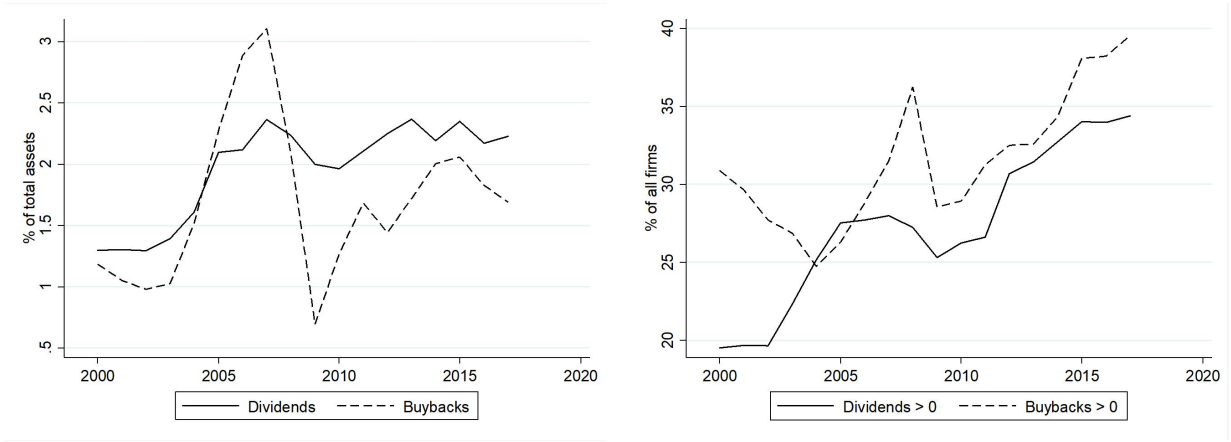
Dynamics of the average common ownership concentration, using two alternative measures (DCO and Number Connect), weighted by market capitalization, are shown in Figure 2.1. These tend to increase over time.

Figure 2.1: Dynamics of common ownership concentration



Note: Considered time period is 2000 - 2017, industries are defined following SIC3 classification. Market value-weighted means are determined for each year

Figure 2.2: Dynamics of dividend and buyback ratios, prevalence of payouts



Note: Left figure includes Dividends/ Assets and Net buybacks/ Assets ratios. Right figure shows percentage of firms with positive dividends and positive net buybacks

Dynamics of corporate payouts is shown in fig. 2.2 and can be compared to that of the common ownership concentration measures in fig. 2.1. During 2000-2017 there was increase in dividends, while buybacks fluctuated wildly (left part). Such behavior is in line with the nature of these two categories of payouts of which dividends tend to be stable as firms are usually reluctant to introduce changes into dividends per share, while buybacks are paid-out in a discretionary manner (DeAngelo et al. 2009). Specifically there was a sharp fall in the buybacks to assets ratio in 2019 during the global financial crisis. The right part of the figure shows proportion of the firms that paid dividends and buybacks among all firms in

the sample. During 2000-2018 there was a steady increase in the proportion of firms with positive payouts.

2.4 Methodology

2.4.1 General set-up

For the analysis of the effect of common ownership concentration on payouts, as well as on capital investments and other outcomes, the following regression is used as the foundation:

$$Outcome_{ijt} = \beta_0 + \beta_1 COC_{jt} + \delta_1 X_{ijt} + \delta_2 Z_{jt} + \nu_j + m_t + u_{ijt} \tag{5}$$

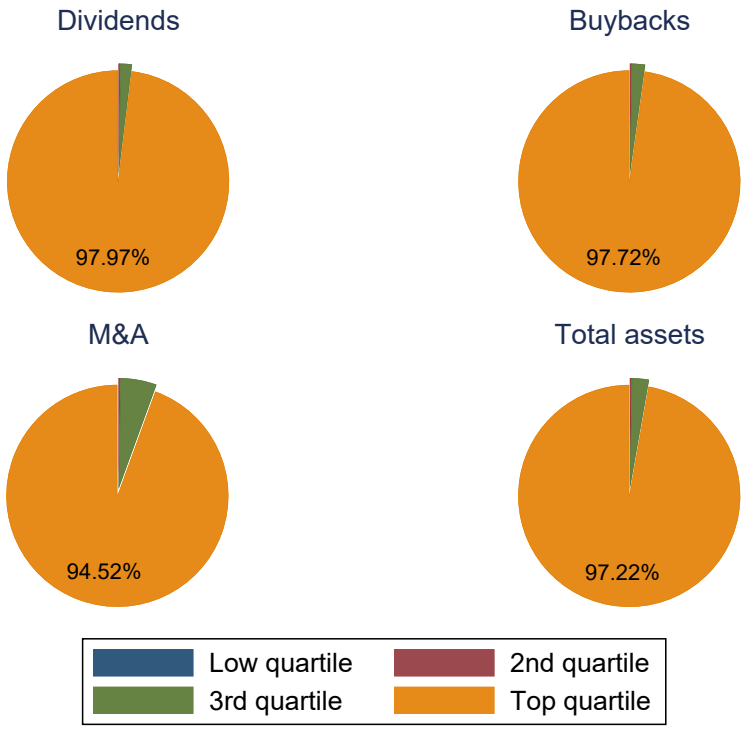
In the equation, outcome variable is one of the following – 'Payouts', 'Dividends', 'Buybacks', 'Markup' and 'Capex', calculated as was detailed in the previous section. Separate equation is estimated for each of the mentioned outcomes. Where, 'Payouts', 'Dividends' and 'Buybacks' are, respectively, the ratio of net total payouts, dividends and net share buybacks, to total assets. CAPEX is the ratio of the sum of net fixed capital expenditures and R&D expenses to total assets. COC is a measure of common ownership concentration, and in most analyses density of common ownership (DCO) is used.

Moreover, Z_{jt} is vector of the industry-level covariates, such as industry HHI, and the number of firms in industry j, while X_{ijt} is vector of the company-specific covariates: log of total assets, leverage, negative net income dummy, retained earnings to total equity ratio, revenue growth rate and cash to total assets ratio. The equation also controls for the set of industry-specific and time fixed effects.

Analysis of the sample structure indicates that large firms account for almost all total dividends, buybacks, acquisitions, and total assets. Firms above the top quartile by total assets accounted for over 97% of all dividends, buybacks and total assets, as well as over 94.5% of acquisitions in 2017 (figure 2.3). Moreover, firms below the median by total assets provided less than 0.5% of all payouts, acquisitions and total assets. These facts dictate that firm size (measured by market capitalization) is used as the weighting factor for all the

regression analyses, providing higher weight for relatively small number of large firms and lower weights for much more numerous small firms.

Figure 2.3: Importance of firms in each size quartile in 2017



Note: Firms are grouped into quartiles by total assets. The figure shows shares in total amount of dividends, buybacks, acquisitions, and total assets that each quartile accounted for in the end of 2017

With equation (5) there is a serious endogeneity problem, as the causal relation between the payouts and common ownership concentration can go both directions. The main considered mechanism is that increased common ownership concentration lowers competition and raises payouts. But there can be other ways for the relation between payouts and common ownership concentration. Such alternative possibility is that investments by institutional investors are directed toward firms with higher payouts, so higher payouts would cause higher institutional investors and thus higher common ownership [8]. To control for the reverse causal link, a number of tests are performed. First, the analysis is performed using matched sample. The use of propensity score matching ensures that the levels of observable covariates, including lagged payouts and institutional ownership, are not systematically different between the

[8]In the obtained sample, correlation between institutional ownership and DCO measure of common ownership is positive and significant at 1% significance level.

treated and control groups. Further, the difference-in-differences approach is applied to ensure identification that relies on exogenous changes in common ownership, which was driven by the well-known mutual fund scandal in 2003 and it can be reliably assumed not to be affected by payouts. Additional testing involves considering for the different exposure of goods industries to international competition captured by the 'China shocks', as well as considering for the different exposure to the market concentration as captured by Herfindahl.

2.4.2 Identification strategy

Relevant identification strategy is required in order to evaluate whether increases in common ownership concentration lead to higher corporate payouts. The analysis based on simple correlations does not reliably work as the positive relation between common ownership and payouts can be due to the fact that institutional ownership is attracted by high payouts. Two approaches are applied in this study. Of these, one is the use of propensity score matching that aims to imitate random assignment of common ownership concentration. While the other approach explores exogeneity of a natural experiment shock due to mutual fund scandal in September 2003.

2.4.2.1 Propensity score matching One of possibilities behind positive association between common ownership concentration and payouts is that institutional investors are attracted to companies with generous payouts policies (Grinstein and Michaely (2005)). As accumulation of institutional ownership in the industry also increases common ownership, this implies that payouts may cause higher common ownership concentration. This possible link was already argued to represent significant identification problem in a test of common ownership concentration effect on payouts. Propensity score matching is therefore aiming to deal with this problem by ensuring a match on lagged observable variables between the treated and control groups. It is implemented by matching observations in the group of firms with high common ownership concentration (above the median) with observations from the group of firms with low common ownership concentration. Propensity score matching aims to achieve the balance in the matched sample so that the situation with the treatment and

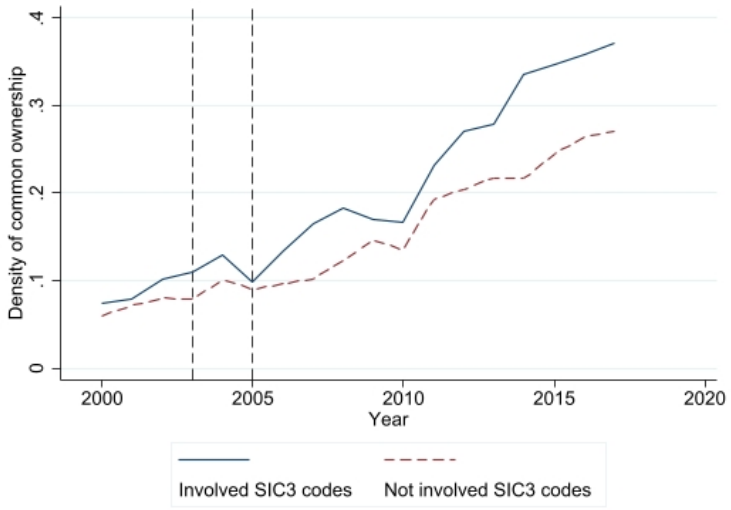
control groups being similar in all other aspects, except for the treatment effect (King and Nielsen (2018)). Matching was performed based on propensity scores, which were estimated using the lagged variables - size, institutional ownership, total payouts, cash, operating profit, Herfindahl index, and CrossDummy. Matching was performed on year-by-year basis.

2.4.2.2 Difference-in-differences This identification approach aims to exploit the exogeneity of a change in common ownership due to mutual fund scandal in September 2003. A large group of mutual funds used malpractices in their trading strategies, which gave rise to the scandal as many investors in these funds were disappointed and embarrassed and withdrew their funds. Subsequently the ownership of these mutual funds decreased significantly [9]. Involved in the scandal 25 families of mutual funds lost over the next two years some 24.3% of their capital, of which 14.1% were lost in the first year following the scandal (Kisin 2011).

Such change can be reliably considered exogenous of the payouts and other corporate decisions of the portfolio firms. Institutional ownership was redistributed among the multiple other institutional investors, and this redistribution influenced the common ownership concentration. In fact, the mentioned scandal event has caused some decline in the average common ownership concentration for the industries with prior heavy ownership of the scandal funds during years 2003-2005 (figure 2.4).

[9] The list of the scandal mutual fund families is obtained from Houge and Wellman (2005).

Figure 2.4: Mean DCO for industries involved in the scandal



Note: Each line shows mean DCO for the respective group of SIC3 industries

Difference-in-differences approach was applied, as follows. The time and treatment dummies, a well as their interaction are included into the regression model, instead of DCO. Of these additional variables, the time dummy is an indicator variable denoting the after-scandal time that starts from year 2003 onward ($Post_t$). The treatment subset indicator ($Treated_j$) indicates the SIC3 industries that were in the top decile by the ratio of the scandal blockholders ownership in the industry (Koch et al. 2018).

$$Outcome_{ijt} = \beta_0 + \beta_1 Treated_j + \beta_2 Post_t + \beta_3 Treated_j \cdot Post_t + \delta_1 X_{ijt} + \delta_2 Z_{jt} + \nu_j + m_t + u_{ijt} \quad (6)$$

While the mentioned scandal resulted in changed institutional ownership of the portfolio companies and not just common ownership concentration, the DID regression directly controls for institutional ownership by including it into the list of covariates.

Additionally, two tests of the mechanisms were implemented aiming to additionally evaluate relevance of the common ownership concentration hypothesis and its impact on corporate payouts. The first test explores the hypothesis that exposure to import competition depresses the anti-competitive effect of common ownership. While the second evaluates the hypothesis

that anti-competitive effect of common ownership, and thus its impact on payouts, is more significant in industries that are less concentrated.

2.5 Findings

The effect of common ownership concentration on payouts and other outcomes is evaluated by controlling for relevant covariates, as well as for the time- and industry-specific fixed effects. The results from fixed effects OLS regressions are considered first (Tables 2.3 and 2.4). Of these, the effect of the industry-level common ownership on payouts and other outcomes is evaluated in Table 2.3. Additionally, interaction between industry-level and firm-level common ownership is explored in Table 2.4. Findings from the analysis using matched samples are provided in tables 2.6 and 2.7. Then the results are presented from the difference-in-differences regression estimation (Table 2.8).

2.5.1 Baseline panel results

The anti-competitive effect of common ownership concentration is expected to increase total payouts, decrease capital expenditures, raise margins and more active industry consolidation via M&A. Basic set of firm-specific covariates is included by following Gutierrez and Philippon (2018) – size (log of total assets), cash flow to assets ratio, book value leverage, annual sales growth, as well as the industry-specific market concentration measured by Herfindahl index (HHI), and log of the number of firms in the industry. Moreover, industry-specific (SIC3) and time-specific fixed effects are included.

Results in Table 2.3 show that DCO exhibits significant positive association with corporate payouts (columns 1 - 3), and significant negative association with capital expenditures (column 5) [10]. The effect of common ownership on markups and acquisition activity is positive but not significant. The analysis in Table 2.3 is prone to the endogeneity issues,

[10]The baseline results were also confirmed using Tobit model, as well as using regression with firm-level fixed effects. All the discussed effects of common ownership were confirmed under these alternative model specifications. These additional results are provided in *Appendix A.4* and *Appendix A.5*, respectively.

due to the possibility of reverse causal linkage between payouts and common ownership. Previously, Grinstein and Michaely (2005) showed positive causal effect of payouts on institutional ownership. Similar reverse linkages are possible between common ownership and other included outcome variables - capital investments, margins, M&A investments. Therefore, the estimated coefficients are prone to the bias. In order to overcome this problem, the difference-in-differences design test was applied and its results are provided further in subsection 2.5.3 below.

Table 2.3: Baseline panel results

	(1) Payouts	(2) Dividends	(3) Buybacks	(4) Markup	(5) Capex	(6) M&A
DCO	0.715***	0.375***	0.357**	3.427	-0.453***	0.218
	(4.11)	(6.14)	(2.43)	(1.13)	(-2.78)	(1.07)
DCO: *Mean* (*SD*)	0.14					
	(0.16)					
Dep var: *Mean* (*SD*)	5.8	4.5	4.4	105.9	6.2	7.2
	(15.4)	(16.0)	(10.5)	(486.5)	(16.8)	(11.3)
Covariates	yes	yes	yes	yes	yes	yes
Industry FE	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes
Observations	78,487	78,487	78,487	78,487	78,171	76,138
Industries (SIC3)	245	245	245	245	245	245
R-squared	0.580	0.706	0.510	0.045	0.414	0.139

Covariates are *Inst own*, *Size*, *Leverage*, *Net Income*, $Cash/TA$, Div_{t-1}/TA_{t-1}, $Net\ BB_{t-1}/TA_{t-1}$, *HHI*, *Size*. t-statistics are in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Payouts, Dividends, Buybacks, Capex, and M&A are scaled by total assets. DCO is density of common ownership. Fixed effects are included for SIC3 industries and years. Regressions are weighted using total market capitalization. Included mean and sd are general sample statistics conditional on positive values for payouts and M&A)

Further the analysis explores the effect of common ownership concentration by considering the interaction of the industry-level measure of common ownership and the firm-level one. Indicator variable *CrossDummy* is included in the regression equation, as well as its interaction term with *DCO*. The interaction term compares the effect of common ownership concentration for the firms that are commonly owned (cross-owned) and other firms in that industry. The findings confirm positive impact of the industry-level measure (DCO) on payouts and its negative effect on capital expenditures, as in the baseline results. In addition, the interaction of *CrossDummy* with DCO is also relevant as the effect of common ownership on payouts and M&A investment differs significantly between cross-owned and not cross-owned firms. There is stronger positive impact of common ownership on dividend

payments of cross-owned firms. Common ownership concentration is evidenced to have a higher effect on M&A investment in the case of commonly owned firms.

Table 2.4: Panel results with cross-owned firms

	(1) Payouts	(2) Div	(3) Net BB	(4) Markup	(5) Capex	(6) M&A
DCO	0.643***	0.336***	0.361**	2.712	-0.492***	-0.005
	(3.59)	(5.33)	(2.39)	(0.87)	(-2.93)	(-0.02)
DCO × *CrossDummy*	0.494	0.339***	-0.142	4.531	0.157	1.748***
	(1.45)	(2.84)	(-0.50)	(0.76)	(0.49)	(4.51)
CrossDummy	-0.187**	-0.027	-0.123*	-2.277	-0.218***	-0.189**
	(-2.25)	(-0.92)	(-1.74)	(-1.57)	(-2.79)	(-1.97)
DCO: *Mean* (*SD*)	0.14					
	(0.16)					
Dep var: *Mean* (*SD*)	5.8	4.5	4.4	105.9	6.2	7.2
	(15.4)	(16.0)	(10.5)	(486.5)	(16.8)	(11.3)
Covariates	yes	yes	yes	yes	yes	yes
Industry FE	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes
Observations	78,487	78,487	78,487	78,487	78,171	76,138
Industries (SIC3)	245	245	245	245	245	245
F-test	9.28***	23.18***	2.85*	0.88	4.31**	10.79***
R-squared	0.580	0.706	0.510	0.045	0.414	0.139

Covariates include *Inst own*, *Size*, *Leverage*, *Net Income*, $Cash/TA$, Div_{t-1}/TA_{t-1}, $Net\ BB_{t-1}/TA_{t-1}$, HHI. Payouts, Dividends, Buybacks, Capex, and M&A are scaled by total assets. DCO is density of common ownership. Fixed effects are included for SIC3 industries and years. CrossDummy equals 1 and 0 to denote cross-owned firms. Regressions are weighted using total market capitalization. Parentheses include t-stats. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. F-test evaluates the null hypothesis that the sum of DCO and DCO × CrossDummy coefficients is zero. Included mean and sd are general sample statistics conditional on positive values for payouts and M&A

Appendix A.6 includes wide format results that are equivalent to the ones above. Wide format results are organized by considering separately the sub-set of firms that are cross-owned and those that are not cross-owned. The wide representation of results can be more straightforward for the purpose of interpretation of the findings.

2.5.2 Propensity score matching analysis

Propensity score matching methodology is used to emulate random assignment of the treatment (high density of common ownership) effect. The treatment group included firms in industries with high level of DCO (fourth quartile and above). Matched sample was constructed as a subset of the total sample, for each 'treated' firm a control one was picked, based on matching variables - lagged size, profitability margin, cash ratio, institutional

ownership, payouts ratio, Herfindahl index and CrossDummy. Matching was performed on the year by year basis, to prevent selection of the same firm into the treated and control groups. Selection of the matched control observations was implemented without replacement. After propensity scores were predicted from the logit model, nearest neighbor algorithm was used to select the best control firm. [11].

Table 2.5: Descriptive statistics by DCO levels for unmatched and matched samples

	Unmatched sample			Matched sample		
	High DCO	Low DCO	Diff	High DCO	Low DCO	Diff
DCO_t	0.311	0.089	-0.221***	0.323	0.086	-0.236***
$Size_{t-1}$	5.964	5.580	-0.383***	5.975	5.994	0.019
Own_{t-1}	0.040	0.028	-0.012***	0.042	0.041	-0.001
$Payouts_{t-1}/TA_{t-1}$	2.395	1.846	-0.549***	2.484	2.458	-0.025
$Cash_{t-1}/TA_{t-1}$	14.97	21.36	6.39***	15.02	14.13	-0.889***
$EBIT_{t-1}/TA_{t-1}$	2.918	-5.521	-8.439***	2.906	3.092	0.186
HHI_{t-1}	2007.2	1416.3	-590.8***	1976.1	1893.5	-82.59***
$CrossDummy_{t-1}$	0.100	0.095	-0.004**	0.102	0.105	0.003
$No.firms$	23,689	72,238		22,506	22,506	

Notes: The two considered sub-samples are 'High DCO' and 'Low DCO'. These are respectively firms from the industries with DCO above the 3^{rd} quartile and that below the 1^{st} quartile. Sig. *** p<0.01, ** p<0.05, * p<0.1.

The table above compares the treatment and control groups before (unmatched sample) and after (matched sample) the matching procedure was performed. The summary statistics comparison indicates that matching helped to improve balance between the two groups by eliminating or at least considerably lowering the differences between the treatment and control subsets. The differences was eliminated for size, institutional ownership, payouts, profitability and $CrossDummy$ (Table 2.6). Between the distribution of treated and control samples significant overlap was confirmed by comparison of their propensity scores distribution (Appendix A.2)

Results obtained from the regression using the matched sample (table 2.6) provide confirmation of the baseline panel data results. Namely, common ownership concentration has positive effect on payouts, including both dividends and buybacks. Also, there is evidenced of negative common ownership concentration effect on capital investments, and positive impact on M&A investments. Moreover, the obtained coefficients from matched samples are higher

[11] Analysis was performed using R package 'MatchIt', following the guidelines regarding application of PSM in Randolph et al. (2014). Critical evaluation of PSM method in King and Nielsen (2018) was taken into consideration, specifically the need to scale the matching variables, and the reservations against using replacement sampling.

in their magnitude and more statistically significant as compared to the ones in baseline panel regression.

Table 2.6: Panel results using matched sample

	(1) Payouts	(2) Div	(3) Net BB	(4) Markup	(5) Capex	(6) M&A
DCO	1.955***	0.432***	1.495***	2.760	-0.739***	0.653**
	(7.42)	(5.04)	(6.62)	(1.54)	(-4.51)	(2.43)
DCO: *Mean* (*SD*)	0.14					
	(0.16)					
Dep Var: *Mean* (*SD*)	5.8	4.5	4.4	105.9	6.2	7.2
	(15.4)	(16.0)	(10.5)	(486.5)	(16.8)	(11.3)
Covariates	yes	yes	yes	yes	yes	yes
Industry FE	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes
Observations	24,327	24,327	24,327	24,327	24,273	23,644
Industries (SIC3)	239	239	239	239	239	239
R-squared	0.614	0.765	0.537	0.169	0.554	0.174

Covariates include *Inst own*, *Size*, *Leverage*, *Net Income*, $Cash/TA$, Div_{t-1}/TA_{t-1}, $Net\ BB_{t-1}/TA_{t-1}$, *HHI*. Payouts, Dividends, Buybacks, Capex, and M&A are scaled by total assets. DCO is density of common ownership. Fixed effects are included for SIC3 industries and years. Regressions are weighted using total market capitalization. Parentheses include t-stats. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Included mean and sd are general sample statistics conditional on positive values for payouts and M&A

The next table below provides the results of regression estimation including the interaction between *DCO* and *CrossDummy*. Such analysis aims to compare the effect of industry-level density of common ownership for the firms that are commonly owned versus the other firms in that industry. The obtained results are generally in line with the equivalent baseline specification. In fact, for commonly owned firms the effect of DCO on dividends is higher, but the effect on net buybacks seems to be lower, than for non-commonly owned ones. In addition, the effect of DCO on M&A investments by commonly owned firms is high in magnitude and significant suggesting that common owners use these firms to facilitate industry consolidation. There could probably be a trade-off between the effect of higher DCO on buybacks and its effect on M&A investments by commonly owned firms. The exact relation would depend on the specific industry, available possibilities for its further consolidation and future growth opportunities.

Table 2.7: Panel results with cross-owned firms using matched sample

	(1) Payouts	(2) Div	(3) Net BB	(4) Markup	(5) Capex	(6) M&A
DCO	2.014***	0.359***	1.642***	2.171	-0.663***	0.292
	(7.37)	(4.04)	(7.01)	(1.16)	(-3.90)	(1.04)
$DCO \times CrossDummy$	-0.448	0.578***	-1.195***	1.798	-0.107	2.003***
	(-0.92)	(3.67)	(-2.88)	(0.54)	(-0.35)	(4.16)
$CrossDummy$	0.081	-0.091*	0.168	-2.137**	0.338***	-0.741***
	(0.51)	(-1.75)	(1.23)	(-1.97)	(3.41)	(-4.58)
DCO: *Mean* (*SD*)	0.14					
	(0.16)					
$Dep\ Var$: *Mean* (*SD*)	5.8	4.5	4.4	105.9	6.2	7.2
	(15.4)	(16.0)	(10.5)	(486.5)	(16.8)	(11.3)
Covariates	yes	yes	yes	yes	yes	yes
Industry FE	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes
Observations	24,327	24,327	24,327	24,327	24,273	23,644
Industries (SIC3)	239	239	239	239	239	239
F-test	27.61***	19.88***	25.24***	1.05	8.56***	11.19***
R-squared	0.614	0.765	0.537	0.169	0.554	0.175

Covariates include $Inst\ own$, $Size$, $Leverage$, $Net\ Income$, $Cash/TA$, Div_{t-1}/TA_{t-1}, $Net\ BB_{t-1}/TA_{t-1}$, HHI. Payouts, Dividends, Buybacks, Capex, and M&A are scaled by total assets. DCO is density of common ownership. CrossDummy equals to 1 and 0 denoting cross-owned firms. Fixed effects are included for SIC3 industries and years. Regressions are weighted using total market capitalization. Parentheses include t-stats. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. F-test evaluates the null hypothesis that the sum of DCO and DCO × CrossDummy coefficients is zero. Included mean and sd are general sample statistics conditional on positive values for payouts and M&A

2.5.3 Difference-in-differences results

The use of difference-in-differences approach is based on the exogeneity of the mutual fund scandal that took place in September 2003 and involved 25 mutual fund families. These funds were reported to be involved in trading malpractices during the after-market hours. After the scandal broke-out, during 2003 - 2005 these funds lost over 25% of their net assets. It was associated with a decrease in the density of common ownership, especially in the industries with the highest ownership concentration of these involved mutual funds. For the purpose of the difference-in-differences analysis, the 'treated' group is defined as all the firms in the industries with the highest ownership concentration of the mutual funds that were involved in the scandal. It is therefore expected that the 'treated' group after the time of treatment exhibits the impact of reduced common ownership concentration, so that

estimated coefficients are opposite to those from the baseline results. Analysis was conducted using matched sample and was limited to years 2002 - 2005, which was the interval around the time of the event.

Table 2.8: DID analysis results

	(1) Payouts	(2) Dividends	(3) Buybacks	(4) Markup	(5) Capex	(6) M&A	(7) DCO
$Treated_j \cdot Post_t$	-13.912***	-3.452***	-10.460***	-0.540	1.046*	-1.149	-0.034***
	(-12.18)	(-7.53)	(-10.49)	(-0.11)	(1.66)	(-0.93)	(-2.64)
$Post_t$	3.139***	0.638***	2.502***	2.152**	0.567***	1.747***	0.028***
	(12.19)	(6.17)	(11.13)	(1.99)	(3.99)	(6.12)	(9.64)
$Treated_j$	19.117***	3.901***	15.216***	12.102	-2.050*	4.726**	-0.120***
	(8.64)	(4.39)	(7.88)	(1.30)	(-1.68)	(2.02)	(-4.78)
Dep Var:	5.8	4.5	4.4	105.9	6.2	7.2	0.14
Mean (SD)	(15.4)	(16.0)	(10.5)	(486.5)	(16.8)	(11.3)	(0.16)
Covariates	yes	yes	yes	yes	yes	yes	yes
Industry FE	yes	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes	yes
Observations	7,707	7,707	7,707	7,707	7,691	7,505	7,707
R^2	0.491	0.649	0.447	0.276	0.597	0.184	0.727

Covariates are $Inst\_own$, $Size$, $Leverage$, $Net\ Income$, $Cash/TA$, Div_{t-1}/TA_{t-1}, $Net\ BB_{t-1}/TA_{t-1}$, HHI. Fixed effects included for SIC3 industries and years. Regressions are weighted using total market capitalization. Sample is limited to years 2002 - 2005. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Included mean and sd are general sample statistics for positive values.

Coefficients of DCO in the panel DID model have signs that are opposite to those of DCO in the baseline models. From table 2.8 column 7 one observes that the mutual fund scandal was associated with a decrease in DCO of about 1/5 standard deviations. Such impact of the 'treatment' by the exogenous variation is negative for payouts, buybacks, mark-ups, acquisitions and DCO; the impact of 'treatment' is positive on capital investments. In other words, lower common ownership concentration is associated with increased competition, lower payouts and M&A investments, and higher capital investments. This 'natural event' was associated with the decreased common ownership and it exhibits impact that is opposite to that of the DCO in the baseline model. Such findings are in line with the suggested framework of the common ownership concentration effect on payouts and investments.

Placebo test was performed to evaluate robustness of this DID analysis. Specifically, false 'scandal' time was used instead of correct one, to explore whether the results are not due to some systematic difference between the treated and control group of firms. In the

obtained placebo test results, none of the difference-in-differences coefficients was significant (*Appendix A.6*).

2.5.4 Robustness checks

Several robustness checks were performed in order to evaluate consistency of the obtained results with respect to different analysis set-ups. Specifically, robustness checks consider relevance of different alternative measures of common ownership concentration (DCO, PCF and PCS), and evaluate results across different time sub-periods (2000-2005, 2006-2011 and 2012-2017). Ideally, the results would remain consistent across all the considered specifications.

Dynamics of the considered three measures of common ownership concentration is presented in the following figures. While DCO and PCS have different levels, their change over time shows similar trends, piques and troughs. Unlike DCO and PCS, PCF has a different dynamics, it shows much higher volatility, especially in 2008-2010 during the global financial crisis.

Figure 2.5: Dynamics of density of common ownership

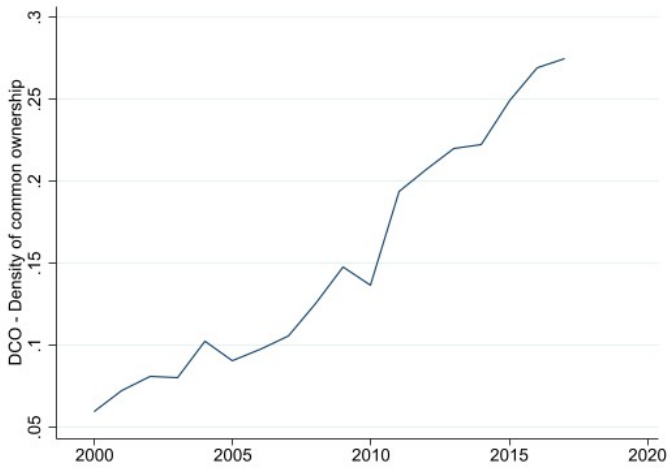
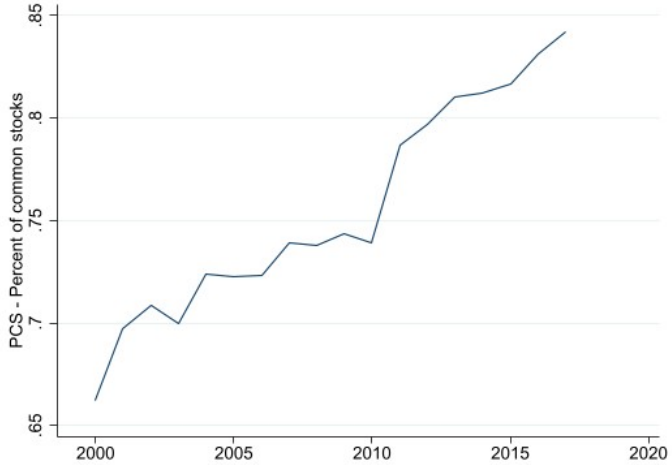


Figure 2.6: Dynamics of percentage of common funds

Figure 2.7: Dynamics of percentage of common stocks



Estimation of the baseline model for payouts with each of the three measures of common ownership concentration is summarized in table 2.9. The effect of all three measures of common ownership concentration is qualitatively similar, as for each measure positive effect of higher common ownership concentration is evidenced. The effect of PCF on buybacks is not statistically significant, presumably due to high volatility of PCF especially during the crisis years.

Table 2.9: Effect of different measures of common ownership on buybacks

	Common ownership measures:		
	DCO	*PCF*	*PCS*
Common ownership conc.	0.582***	0.342	0.582***
	(2.91)	(0.21)	(5.01)
Dep var. *Mean* (*SD*)	0.143	0.049	0.748
	(0.160)	(0.034)	(0.257)
Covariates	yes	yes	yes
Indus. FE	yes	yes	yes
Year FE	yes	yes	yes
No. obs	78,487	78,487	78,487
R2	0.387	0.387	0.387

Dependent variable is Buybacks/ Total assets. Common ownership measures: *DCO* - density of common ownership, *PCF* - percentage of common funds, *PCS* - percentage of common stocks. Fixed effects included for SIC3 industries and years. Regressions are weighted using total market capitalization. t-statistics are in parentheses. Significance levels: ***p<0.01, **p<0.05, *p<0.1

The following test explores whether the results are preserved for different time sub-periods: 2000 - 2005, 2005-2012, and 2013 - 2017. Although these sub-periods are of similar length, they are qualitatively different. Specifically, there was dot-com crisis in 2000-2005 along with the mentioned exogenous shock to the mutual fund industry. During 2006 - 2012 there was the global financial crisis. As common owners strived to increase shareholder value they would reduce payouts during the times of crisis in order to improve liquidity and avoid cash deficits. To allow for the differences in the nature of each sub-period, an interaction term was added between DCO and a dummy for negative net income. After controlling for time-specific and firm-specific situation, DCO has positive sign in its effect on net buybacks in each of the considered three periods.

Table 2.10: Effect of common ownership on buybacks by sub-period

		Sub-periods:		
	2000-17	2000-05	2006-11	2012-17
DCO	0.073*	0.239***	0.042	0.357***
	(1.68)	(3.18)	(0.52)	(4.45)
$DCO \cdot I_{Income<0}$	-0.101	-0.313*	-0.949***	0.567**
	(-0.75)	(-1.66)	(-3.52)	(2.41)
Mean	0.143	0.080	0.144	0.245
(*SD*)	(0.160)	(0.121)	(0.151)	(0.179)
Covariates	yes	yes	yes	yes
Indus. FE	yes	yes	yes	yes
Year FE	yes	yes	yes	yes
No. obs	77,508	26,378	27,107	24,023
R2	0.384	0.427	0.370	0.452

Dependent variable is Buybacks/ Total assets. *DCO* - density of common ownership. Fixed effects included for SIC3 industries and years. Regressions are weighted using total market capitalization. t-statistics are in parentheses. Significance levels: ***p<0.01, **p<0.05, *p<0.1

2.6 Tests of the Mechanism

A number of conceptual tests were performed to further evaluate relevance of the considered mechanism regarding the common ownership concentration effect on competition. In relation to this, external competition that is not under control of the common owners is taken into account by considering 'China shock' (sub-section 2.6.1). It is hypothesized that under external competition that is not under control of, and that cannot be reduced by common owners, the effect on payouts should be substantially lower than in the baseline analysis. Also, the effect of common ownership on payouts and other outcomes is explored in relation to the 'traditional' market concentration using Herfindahl (sub-section 2.6.2). It is hypothesized that in the case of unconcentrated industries, the ability of common owners to reduce competition, increase payouts and facilitate consolidation should be higher that in case of concentrated industries.

2.6.1 Relevance of import competition

The import competition test includes interaction of the density of common ownership with indicator for high China shock. The rationale is that in the industries with high prevalence of imports, competition cannot be lowered by the common owners of the US firms, as the US firms even when commonly owned must still compete with their foreign peers. Among the prior researchers, Acemoglu et al. (2016) showed relevance of import competition for development of the US economy and argued that import competition has seen a surge since 2000 causing job losses to the US of about 2.0-2.4 million. Moreover, Pierce and Schott (2016) also explored the decline in manufacturing employment that took place after 2000.

Current paper follows the approach in Pierce and Schott (2016) for measurement of the exposure to China import shock on the industry-level. This approach relies on the exogenous change in the import tariffs on Chinese goods, due to the US Congress granting Chinese goods with the Permanent Normal Trade Relations (PNTR) regime in 2001, in relation to China accessing the WTO. This marked decline in the US import tariffs on Chinese goods across broad categories of industries. Pierce and Schott (2016) indicate the shock impact was to lower the manufacturing employment by 18% during 2000 - 2007, while it was approximately unchanged at 18 million during 1965 - 2000. Following the considered approach, variable 'NTR gap' is used as exogenous time-invariant industry-level proxy for the China shock. The variable shows the difference between non-NTR and PNTR import tariff rates. Where non-NTR rates are those that would be used if PNTR status was not granted (Pierce and Schott 2016)[12].

$$NTR\ Gap_j = non\ NTR\ rate_j - NTR\ rate_j \tag{7}$$

Considering for the China shock is helpful to test in several ways the anti-competitive hypothesis in relation to the common ownership concentration. One way is to directly test whether higher exposure to non-controlled competition can lower the ability of common owners to increase payouts and reduce investments. Moreover, the 'China shock' helps to

[12]In fact, Pierce and Schott (2016) indicate that since 1980 till 2000, the US applied special procedure towards setting import tariffs for Chinese goods. This required annual re-consideration of whether NTR should or should not be provided. This was still accompanied by significant uncertainty. Removal of this uncertainty was argued to give rise to the 'China shock' since after 2001.

evaluate the link between common ownership and payouts is exclusively due to institutional investors being attracted towards companies with higher payouts.

The major hypothesis of the study is that common ownership concentration causes increase in total payouts, which occurs alongside lower industry competition as common owners aim to mitigating the free cash flow problem. Analysis in this section evaluates the ability of import competition to lower the anti-competitive effect of common ownership. This effect, if confirmed empirically, would provide further evidence in favor of the considered anti-competitive effect of common ownership.

Table 2.11: Panel results by 'China shock' effect

	(1) Payouts	(2) Div	(3) Net BB	(4) Markup	(5) Capex	(6) M&A
DCO	5.897***	0.625***	5.408***	7.446	0.380	-0.588
	(10.87)	(3.78)	(11.61)	(1.52)	(1.61)	(-1.04)
$DCO \cdot High\ CS$	-2.567***	-1.103***	-1.754**	5.061	-0.370	2.545***
	(-2.99)	(-4.22)	(-2.38)	(0.65)	(-0.99)	(2.80)
DCO: *Mean (SD)*	0.143					
	(0.160)					
Dep Var: *Mean (SD)*	2.804	1.127	1.551	114.55	5.367	2.254
	(5.721)	(2.732)	(4.138)	(688.1)	(6.875)	(5.892)
Covariates	yes	yes	yes	yes	yes	yes
Industry FE	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes
No. obs	9,477	9,477	9,477	9,477	9,446	9,173
Industries (SIC3)	80	80	80	80	80	80
R^2	0.646	0.830	0.564	0.111	0.812	0.195

Each regression includes DCO, interaction of DCO and High CS (dummy for high 'China shock'), High CS dummy and continuous CS variable. Dividends, Buybacks, Capex and M&A are scaled by total assets. *DCO* is density of common ownership. Fixed effects included for SIC3 industries and years. Regressions are weighted using total market capitalization. Other covariates: *Size*, *Leverage*, *Net Income*, $Cash/TA$, Div_{t-1}/TA_{t-1}, $Net\ BB_{t-1}/TA_{t-1}$, HHI. t-statistics in parentheses.
Significance levels: *** p<0.01, ** p<0.05, * p<0.1. The analysis is performed using matched sample and industries that are exposed to the China shock

Results from the China shock indicate that positive effect of common ownership on payouts is significantly reduced in the industries with high exposure to the China shock. Also China shock shows high relevance for the effect of common ownership concentration on M&A investment. The latter is in line with findings by Gutiérrez and Philippon (2017) who indicate that firms, which survived under China shock are those highly competitive ones with solid financial performance. Common owners are thus able to facilitate industry consolida-

tion, presumably in order to boost higher competitiveness in the sectors that are exposed to imports. The effect of common ownership concentration on markups and capital investments are insignificant. There is difference in sample composition between these findings for China shock effect and the baseline results. While the baseline results are obtained for the total sample, the 'China shock' findings are obtained for a subset of industries mainly manufacturing, retail and wholesale sectors that involve tradable goods.

The results from analysis of the China shock interaction with DCO is in line with the considered hypothesis about the anti-competitive effect of common ownership. Under higher external competition, common owners prefer to decrease withdrawal of free cash flow from the firms. Instead of higher payouts, common owners tend to facilitate more active M&A investments in the industries that are exposed to the China imports. Industries that have higher consolidation are able to rip economies of scale and are more capable of withstanding international competition. Such impact by common ownership can be associated with higher profitability, yet the effect of $DCO \cdot High\ CS$ on Markup is positive but not statistically significant. Gutiérrez and Philippon (2017) mention that the firms that remain in the industries with high imports competition are more profitable, innovative and more competitive. Similar findings as those in the table above, are provided in wide format. The wide format results, instead of the DCO interaction with the Chinese shock, provide the effect of DCO on the outcome variables by sub-sample of industries with low and high China shock. These results are provided in Appendix A.7.

2.6.2 Relevance of industry concentration

Industries differ from each other by their level of market concentration and competition. While some industries include only several players that control the entire market, others may have hundreds of participants that compete intensively with each other. The effect of common ownership concentration on firms in a given industry could differ conditional on the level of market concentration (Herfindahl index) in that industry. In a highly concentrated industry with only few players, competition may already be at a low level, as firms in this may already lack incentives to compete with each other. Increase in common ownership

concentration might not be able to result in any further decline in competition in such an industry. Conversely, in an industry with high number of players, each of whom has a low market power, increase in common ownership concentration could facilitate a decline in competition much faster and easier. Analysis in this subsection addresses the hypothesis that higher density of common ownership has greater impact on payouts and on other relevant outcomes, such as capital investments, M&A activity, and markups, in industries with low concentration (Herfindahl index). This is the second test of the mechanism, which aims to additionally validate the relevance of common ownership concentration for payouts via its implications for competition.

For testing the above mentioned hypothesis, an indicator variables, $Low\ HHI$ was constructed. Industries (markets) with low concentration have HHI $\leq$ 1,500), while industries (markets) with HHI $\geq$ 2,500 are defined as highly concentrated ones [13].

Table 2.12: Panel results and industry concentration

	(1) Payouts	(2) Div	(3) Net BB	(4) Markup	(5) Capex	(6) M&A
DCO	1.005***	0.413***	0.530***	-0.610	-0.673***	0.567**
	(4.71)	(5.61)	(2.92)	(-0.15)	(-3.18)	(2.19)
$DCO \cdot I_{HHI\leq1500}$	0.124	0.176	0.087	22.892***	-0.826**	1.624***
	(0.32)	(1.30)	(0.26)	(3.04)	(-2.12)	(3.46)
$I_{HHI\leq1500}$	-2.709***	-0.437***	-2.133***	-4.210	-0.943***	1.328***
	(-12.22)	(-5.72)	(-11.31)	(-0.99)	(-4.29)	(4.81)
DCO: Mean (SD)	0.143					
	(0.160)					
Dep var: *Mean (SD)*	2.804	1.127	1.551	114.55	5.367	2.254
	(5.721)	(2.732)	(4.138)	(688.1)	(6.875)	(5.892)
Covariates	yes	yes	yes	yes	yes	yes
Industry FE	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes
No. obs	63,573	63,573	63,573	63,573	63,314	61,670
Industries (SIC3)	244	244	244	244	244	244
R^2	0.575	0.679	0.520	0.040	0.399	0.146

The model includes $I_{HHI\leq1500}$, interaction of DCO and $I_{HHI\leq1500}$. Dependent variables - Dividends, Buybacks, Capex and M&A are scaled by total assets. *DCO* - density of common ownership. Fixed effects included for SIC3 industries and years. Regression weights - Total market value. Covariates: *Inst own*, *Size*, *Leverage*, *Net Income*, *Cash/TA*, Div_{t-1}/TA_{t-1}, $Net\ BB_{t-1}/TA_{t-1}$, *HHI*. t-statistics in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Included mean and sd are general sample statistics for positive values

[13]The US Department of Justice (2018) indicates the two mentioned benchmarks that are used in its activities, as well as used by other relevant agencies. An example of a typical use is in evaluation whether a merger would cause excessive concentration of the market power.

The results in table 2.12 indicate that all the effects of common ownership on payouts can possibly be higher in the industries with lower concentration. Coefficient of the interaction of DCO with $I_{HHI \leq 1500}$ in the effect on payouts is positive but not statistically significant. Additionally, the effect of DCO on markups is significantly higher, and its effect on capital expenditures is significantly lower in industries with low concentration. Also, the impact of DCO on M&A investments is considerably much higher in industries with low concentration. All these effects are in line with the hypothesis that there is more room for common owners to pursue consolidation in the industries with low concentration and high competition, as opposed to the industries with high level of concentration. The obtained findings from analysis of the effect of the interaction between common ownership and market concentration serve to support the effect of common ownership towards lowering competition among firms. The equivalent results in the wide format with the analysis performed by sub-samples are provided in *Appendix A.6*.

2.7 Conclusion

The study evaluated the causal impact of common ownership concentration on corporate payouts based on the mechanism that increased higher common ownership concentration tends to lower competition among the firms. Analysis was based on the sample of over 11,000 listed US firms during 2000 - 2017. It implemented such identification strategies as the use of propensity score matching and difference-in-differences methodology, as well as several tests of the mechanism and feasibility checks. As institutional investors jointly increase their ownership shares in businesses within the same industry, it leads to higher common ownership concentration. Such increase in common ownership concentration is argued to reduce competition among the industry players. This is related to the interest of common owners in pursuing policies, such as not fostering competition among the firms, in order to promote the value of their shareholdings.

The key mechanism that is considered in this study is based on the presumed ability of common ownership to lower competition in commonly owned industries. Decreased competi-

tion is associated with higher market power of the market players, resulting higher markups, lower capital investments, and increase in corporate payouts. The latter is rationalized by the ability of common owners (institutional investors) to deal with the free cash flow agency problem via fostering higher payouts. In addition to the main mechanism, several additional links were considered that explain correlation between common ownership and corporate payouts. For example, institutional investors can be attracted to the firms with generous dividends and buybacks; or institutional investors per se can affect corporate payouts. Besides, the increased cash flow of the firms in the industries with increased common ownership is not necessarily used for making payouts, but it can also be used to finance acquisitions. The current study obtained confirmation of the main considered mechanism and controlled for the mentioned alternative explanations.

The study confirmed positive impact of common ownership concentration on corporate payouts. Based on the results from matched sample, an increase in density of common ownership by one standard deviation results in a 1 percentage point increase of the payouts to assets ratio (1/15 of its standard deviation). Of this increase, 3/4 is due to the raise in share buybacks and 1/4 comes from the dividends growth. This study also re-establishes positive impact of common ownership on mark-ups and acquisitions, and its negative effect on capital investments. These findings provide evidence in favor of the common ownership concentration ability to lower competition to the benefit of common owners. The effect of common ownership concentration was shown to be higher in magnitude for the commonly owned firms, as opposed to the other firms in the same industry.

The study also performed two tests of the mechanism. Specifically, it explored whether the common owners impact on corporate payouts, profitability, and investments changes under increased import competition, as well as whether common ownership concentration has different effect in industries with low concentration as opposed to those with high concentration. When industry is exposed to significant import competition, such source of competition is not under control of the common owners. Thus, their impact on payouts and investments is expected to be altered in such industries, in comparison to the baseline results. The obtained results indicate that high import competition exposure lowers the positive impact of common ownership on payouts and increases its positive impact on concentration among

domestic suppliers via higher M&A investments. It was also confirmed that for industries with low concentration (low Herfindahl) there is greater magnitude of the positive common ownership effect on payouts, markups and acquisitions, as well as more pronounced negative impact on capital investments.

The area of common ownership concentration is an expanding one and there are numerous directions for future research. One aspect is that further analysis would benefit from exploring implications of common ownership concentration for firms within the same between-industries supply chains, opposed to the within-industry setup. Also, the use of structural modeling could help to gain more insight into the implications of common ownership given the complex nature of the interaction between involved stakeholders. Besides, currently there is no a compelling measure of common ownership concentration that could be employed by regulators for purposes such as monitoring mergers and acquisitions. Specifically, based on performed simulation analysis, all considered common ownership indicators under certain conditions showed performance that was not in line with common sense reasoning. Development of a more compelling measure of common ownership is another promising research venue.

3.0 Wealth Implications and Causes of Gray Divorce

I estimate implications of 'gray divorce' for all components of wealth among those aged 50+. The essay uses semi-parametric difference-in-differences methodology, which compares previously married individuals that become divorced to those who remained married. Matching is used to ensure that both groups are similar on key lagged economic and demographic characteristics. My results confirmed negative impact of gray divorce on net worth, especially such components as housing equity and financial assets. There is no evidence of higher decline in wealth among females as compared to males. However, divorcing females are shown to experience greater decline in the value of their IRAs and stocks, and are more likely to re-enter labor force. It was shown that negative wealth implications of gray divorce increased with age among those 50+. The observed deterioration in wealth can be related to the costs of the divorce process, loss of the economies of scale, as well as with deterioration of mental ability due to loneliness. Wealth was shown to also be a key predictor of the probability of gray divorce, as higher wealth lowers chances of gray divorce.

3.1 Introduction

Recent decades evidenced rapid expansion in divorce rates among the elderly, which exceeded those of the younger cohorts. For the purpose of this study, 'gray divorce' is defined as marital dissolution among a couple where one or both partners aged 50 or above. Such definition was also used in Sharma (2015), Crowley (2018), and Brown and Wright (2017) among others. The US population faces steady trend towards increasingly greater share of the elderly in its structure. Those aged 50+ constitute over 1/3 of the total population, according to the US Census. They also have the highest median savings among all other population groups by age (Gilbert 2017). Thus, the possible negative impact of gray divorce for household wealth could have considerable negative implications for the economy as a

whole. Crowley (2018) mentions the so called economic and social penalties of gray divorce. While earlier researchers did recognize the economic losses associated with divorce including its impact on wealth, most prior studies were correlational and lacked a systematic empirical analyses of marital dissolution implications for the wealth and economic well-being of the elderly using an appropriate identification strategy. The current study relies on the use of semi-parametric difference-in-differences methodology in order to isolate and estimate the causal impact of gray divorce on wealth and wealth components of the household, including its effect on total net worth, housing and non-housing wealth, different types of financial assets, and on such related economic and well-being outcomes like cognitive ability, financial and total income, and labor market participation. The study also explores the role of various demographic, economic, and health-related factors for predicting probability of gray divorce among those aged 50 and above.

The tendency towards increasing marital dissolution rates among the elderly is well documented, as is negative implications of gray divorce for health and cognitive ability. Stepler (2017) indicates that during 1990 - 2015 the divorce rates among those aging 50 and above have more than doubled, in comparison to its modest 14% increase for those aged 40-49 and to a 21% decline in divorce rates among the cohort of 25-39 year old. Prior studies have also evidenced that divorces lead to increased loneliness and greater social exclusion among the elderly (Del Bono et al. 2007). In its turn, feeling of loneliness was shown to have adverse health and mental ability implications (DeLiema et al. 2018; Luo et al. 2012). For the case of young cohorts, Gardner and Oswald (2006) showed that divorce increases the level of happiness, especially for individuals who re-partnered soon after the divorce. But the latter may be less relevant for the elderly, and is also evaluated in this analysis. The current study explores the causal link between divorces of the elderly and its implications for their wealth. Because the elderly rely heavily on their wealth to maintain stable consumption level and quality of life, implications of gray divorce could be devastating for the economic well-being of the senior population.

Quite a number of studies considered the effects of marital dissolution on economic and financial well-being (e.g. Smock et al. 1999, Zagorsky 2005, DeLiema et al. 2018, Brown and Wright 2017, Haider et al. 2003), although most of these studies pertained to the younger

cohorts. Of these studies, Zagorsky (2005) explored the impact of divorce on net worth for couples in their 20s, 30s and 40s using correlational analysis. Zagorsky (2005) concluded that per person wealth of married individuals, on average, is 77 percent higher in comparison to that of a single person. Meanwhile for divorced ones, there is an equal magnitude decline in wealth. The result can be driven by econometric selection problem. Theoretical mechanisms that explain faster wealth accumulation in married households include economies of scale and specialization of partners in the marriage. On other hand, raising children, which is relevant for younger couples but usually not for those aged 50+, requires expenses and this has a negative marginal impact on wealth in married couples (Zagorsky 2005, Becker et al. 1977). Many studies (e.g. Smock et al. 1999, Sharma 2015) indicated greater economic 'penalty' from a divorce for women as compared to men in relation to re-entering the job market. These studies argued that, on average, women devoted much time to raising children, they have lower work experience and education levels, which puts then in a disadvantage when re-entering the job market. Haider et al. (2003) reports that among the elderly, single women have the same labor force participation rate as married men. Also, due to the foregone working experience that was instead foregone due to raising children, elderly women that re-enter labor force, are paid less in comparison to elderly men, after controlling for the level of education. Elderly individuals in a marital partnership provide social support and care to one another. When a marriage is dissolved, such care-giving services are foregone, leading to greater loneliness and lower mental ability. Mazzonna et al. (2018) indicates that the lack of awareness about one's own cognitive decline has negative financial implications, due to ineffective disinvestment decisions and higher vulnerability to financial scam. The current study builds on these mentioned mechanisms and tends to explore further divorce implications of various aspects of wealth of the elderly. Despite growing research studying gray divorce and its implications for the elderly, there is lack of a systematic causal evaluation of the gray divorce implications for the wealth and its components. Therefore, the current study aims to address this mentioned gap.

It is also important to understand the relevance of wealth along with relevant demographic and heath factors, to predict gray divorces. An additional component of this study is the extended exploration of the relevance of various economic, demographic and health

factors to predict divorces in the case of the elderly households, using machine learning methodologies. Different factors of marital stability and divorces, in relation to a general population rather than the 'gray' cohort, were broadly covered in prior theoretical and empirical literature. Becker et al. (1977) indicates relevance of marriage-specific assets to lower the probability of divorce. Such assets include giving birth to children and devoting time to non-market working. In relation to the marriage-specific assets mechanism, Stepler (2017) reported higher divorce rates among individuals in their second and subsequent marriages, which is due to a much lower level of marriage-specific assets in the second, third and so on marriages. Prior empirical results (e.g. Crowley 2018; South and Spitze 1986) indicated relevance of such factors as husband's employment, wife's labor force participation, race, home ownership, age at marriage, and living in urban area. South and Spitze (1986) mentioned a changing role of wife's education that lowers probability of divorces in the early age, but increases it in later life periods. The current study considers most of the mentioned determinants and includes additional ones aiming to better predict divorce risks among those aged 50 and above. Anecdotal evidence, along with interview-based studies, indicate such reasons for gray divorce may include spouses having grown apart, one of spouses having mental problems or unhealthy addiction, financial issues, physical cheating by a spouse, and domestic abuse (Crowley 2018). Few prior empirical studies captured the effect of the 'growing apart' concept (in realms of religion, occupation change, physical attractiveness etc.) as a divorce reason. Therefore, in the realm of forecasting gray divorces, the current study contributes to the literature as it adds additional variables to capture differences between spouses, and employs powerful models from the machine learning methodology.

In relation to the above discussion, the current study aims to address the following research questions. First, how does gray divorce affect wealth and wealth components of the household? Second, what are other economic and cognitive outcomes of the divorce in a senior age? Third, what factors predict the marital dissolution among the elderly and does wealth play a significant role in predicting gray divorces?

The remaining parts of the study are organized as follows: Chapter 2 details the methodological foundation of the study, including the data source, sample construction, variables and descriptive statistics, as well as causal inference methodology. Chapter 3 presents the

results of the estimated effects of gray divorce on wealth and its components, as well as additional economic and well-being outcomes. Chapter 4 explores the predictability of gray divorce in senior couples, including evaluation of economic, health and demographic determinants. Finally, Chapter 5 concludes the study and lists its key findings.

3.2 Data, Sampling and Methodology

3.2.1 Data and sampling

Data was obtained from RAND Health and Retirement Survey (HRS) version 2016 release 2. The RAND version of HRS includes imputations for multiple income and asset-related variables, ensuring greater comparability and completeness of information in the database (Bugliari et al. 2020). The US-wide panel dataset covers individuals aged 50 and above. For all married and partnered individuals the dataset provides all the relevant information on their spouses. The data includes 13 waves of surveys and covers 7 cohorts of the US population, totaling to 42,051 individuals from 26,598 households. Current study recognizes that couples who get divorced and those who remain married can be systematically different from each other (Appendix B.1 gives comparison of the summary statistics between the two groups). Thus, the first part of the identification strategy implements propensity score matching in order to select a similar household that remains married for each married household that divorces in the next period. For the purpose of this analysis a 'treated' individual is a one that was married during at least 2 recent waves and then becomes divorced for at least 2 subsequent waves. A 'control' individual is a one that was matched to the 'treated' one and also was married for at least 2 waves prior to and 2 waves after the divorce of its match.

3.2.2 Variables

A number of relevant wealth measures are included into the analysis. The key general measures of wealth are total net worth, housing equity, non-housing wealth, and total

financial assets. These measures and their construction are summarized in the table [1]

Table 3.1: The key wealth variables

	Indicator	Components
1	Net worth	All housing and non-housing assets less all mortgages and other debts
2	Housing equity	Sum of primary and secondary residence less related mortgages and home loans
3	Non-housing wealth	Net worth less Housing equity
4	Financial assets	Stocks, bank accounts, CDs, bonds, other financial assets

In addition to the wealth variables, this study explored such economic and well-being outcomes as happiness status, retirement, labor force participation, total earnings, cognitive score, and memory test score. Variable 'happiness status' is used for the purpose of comparing the current study results for the elderly to the findings of Gardner and Oswald (2006) in case of the younger cohorts. The control variables in this study include household size, the respondent's age, gender, race and religion, time and cohort dummies. All wealth measures, such as net worth or financial assets, are available on 'per household' basis from the RAND HRS 2018. To transform those into a 'per person' basis, square root equivalence scale was applied, according to the approach in OECD (2011). The use of equivalence scales is required for comparison of well-being across households of different number of household members, and this relates to the possibility of under-aged individuals in the household, and to account for any economies of scale and cost saving opportunities (Rojas 2014). Under the square root equivalence scale, a 'per person' estimate is obtained by dividing the per household value by square root of the number of people residing in the household. This approach was preferred to other alternatives (e.g. OECD equivalence scale assigns weight of 1.0 to head of household, 0.7 to each additional adult and 0.5 to a child under 18), because the HRS provides information about household size, but does not decompose into the number of under-age children and adults.

[1] Using HRS codes (Bugliari et al. 2020):

- Net worth = ahous +arles +atran +absns +aira +astck +achck +acd +abond +aothr -amort -ahmln -adebt
- Housing equity = hatoth + hanethb
- Non-housing wealth = arles +atran +absns +aira +astck +achck +acd +abond +aothr -adebt
- Financial assets = astck +achck +acd +abond +aothr

3.2.3 Descriptive statistics

Causal analysis is based on the matched sample, wich was obtained by retaining all 'treated' individuals and by selecting a 'control' individual for each 'treated' one. The matching was performed in a way that ensures similarity between 'treated' and 'control' individuals in all relevant respects, including net worth, income, age, number of household members, education, in the periods that precede the gray divorce of the 'treated'. Descriptive statistics for this matched sample is provided in the table below, indicating central tendency and dispersion for the variables, and also evaluating the balance between treated and control categories.

Table 3.2: Matched sample balance evaluation

	Divorced		Married				
	Mean	SD	Mean	SD	Difference	t-stat	p-val
Non-white, %	0.23	0.42	0.22	0.42	0.00	0.07	0.94
Protestant, %	0.63	0.48	0.64	0.48	-0.01	-0.29	0.77
Catholic, %	0.22	0.41	0.19	0.39	0.03	0.93	0.35
Age, years	59.32	8.87	59.25	8.90	0.07	0.11	0.91
$HH\ size_{t-1}$	2.67	1.04	2.77	1.38	-0.11	-1.14	0.25
$Total\ income_{t-1}$	59300.70	48210.68	52450.32	55516.94	6850.38	1.73	0.08
$Capital\ income_{t-1}$	8420.71	20476.63	7817.01	28049.96	603.70	0.32	0.75
$Net\ worth_{t-1}$	269495.97	764840.26	204443.27	395954.02	65052.70	1.40	0.16
$Hous.\ wealth_{t-1}$	161120.37	437961.84	131537.32	332666.36	29583.05	1.00	0.32
$Fin.\ wealth_{t-1}$	70892.86	339371.10	53626.63	218501.41	17266.23	0.79	0.43
$Stocks_{t-1}$	48322.12	316011.38	25726.58	120940.83	22595.53	1.24	0.22
$Bonds_{t-1}$	2136.63	22646.26	3243.73	35144.70	-1107.10	-0.49	0.62
$Bank\ accs_{t-1}$	19975.99	67624.21	14383.29	64862.38	5592.70	1.11	0.27
CDs_{t-1}	4481.40	17214.02	7365.85	35932.51	-2884.45	-1.34	0.18
$Vehicles_{t-1}$	14118.17	16331.78	12913.56	14384.49	1204.61	1.03	0.31
IRA_{t-1}	31593.49	109791.09	32275.55	104291.55	-682.06	-0.08	0.93

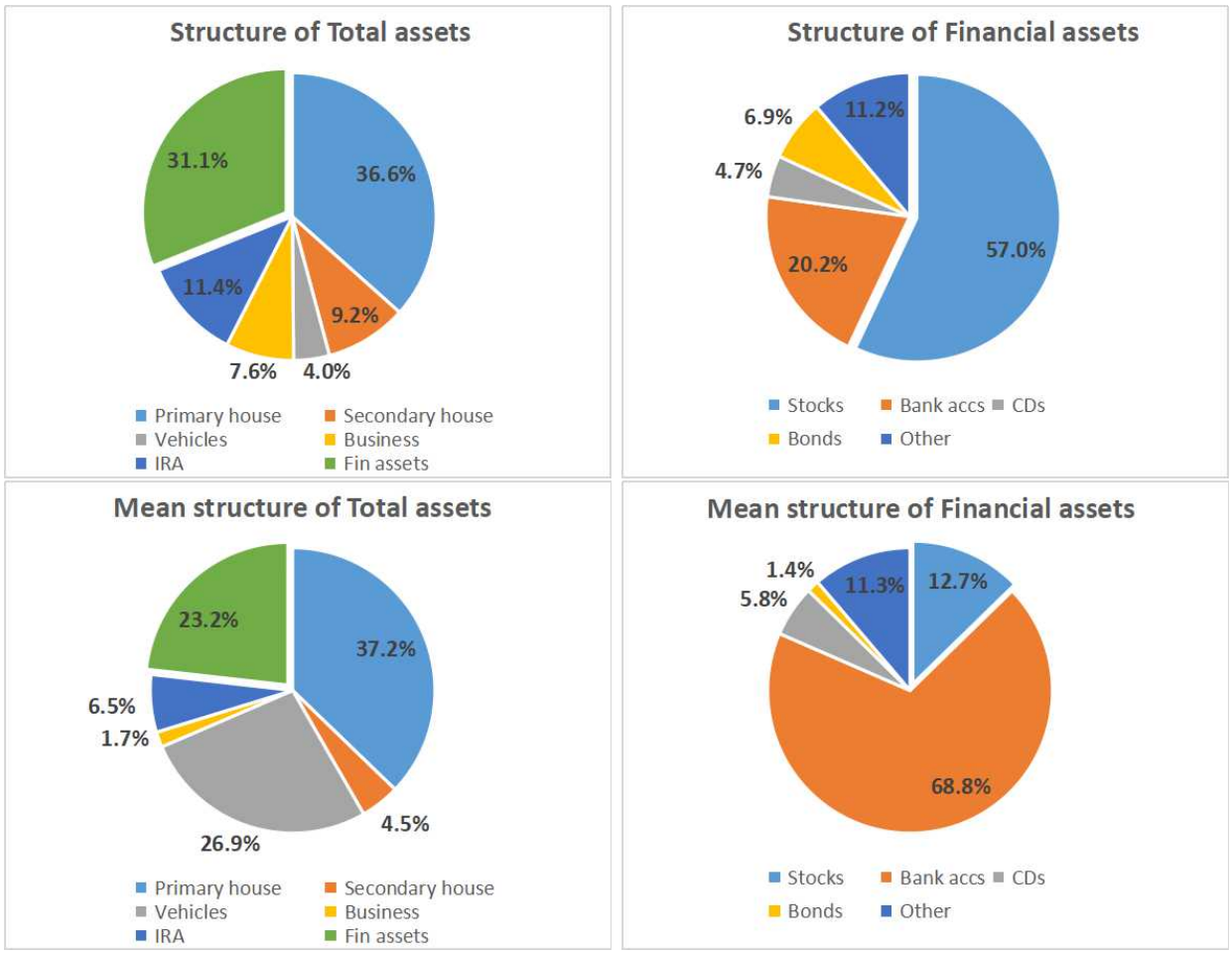
All income and wealth variables are in per equivalent person basis in logarithms. Robust t-statistics in parentheses. Sig: *** p<0.01, ** p<0.05, * p<0.1.

The results indicate that housing wealth comprise about 60 % of total net worth per individual. Among financial assets, stocks is the largest category (68 % of financial wealth) followed by IRA as the second largest (45 %) and bank accounts as the third largest (28 %) [2]. Comparison is performed in the wave that preceded the time when 'treated' individuals appeared divorced for the first time. While the matched sample does not evidence any significant differences between the treated and control subsets in the period before the divorce, the original sample contains significant differences between these groups (Appendix B).

[2]The sum of the weights for stocks, IRAs and bank accounts add up to over 100% because 'financial wealth' is composed of the mentioned components less 'Other debt'

Structure of the household assets for the 'treated' group of households is shown in the figures below. The top row of Figure 3.1 shows the weighted average structure of total assets and financial assets for the 'treated' households. The bottom row of Figure 3.1 indicates the simple average structure of total assets and financial assets, respectively. While the weighted average indicates the categories where the most wealth is stored, the bottom row shows the most typical structure for the considered households. All the results are provided for the time period when 'gray divorce' was observed for the first time for the 'treated' households.

Figure 3.1: Structure of total assets and mean structure of assets



Note: Top row shows weighted average structure of each category for the 'treated'. Bottom shows mean structure of each category for the 'treated' . All data is provided as of the year of 'Gray divorce'

3.2.4 Methodology

3.2.4.1 Difference-in-differences The method of difference-in-differences (DID) is a panel data estimation methodology that aims to isolate pre-treatment differences between the two groups, as well as any common trends in the dependent variable, and thus estimate the treatment effect (Taddy 2019). The following regression version of difference-in-differences captures the effect of gray divorce on wealth from the coefficient of the interaction between treated dummy and the post-treatment time dummy.

$$Y_{it} = \alpha + \gamma D_{div} + \lambda D_{time} + \delta D_{div} D_{time} + \beta X_{it} + \epsilon_{it} \tag{8}$$

Coefficient δ in the model above shows the difference between the mean change in the outcome variable after the divorce for the 'treated' groups and mean change for the 'control' group over the same time (Angrist and Pischke 2009).

The difference-in-differences method assumes that without treatment, both groups would evidence parallel time trend. If this assumption does not hold, the obtained DID estimator could involve a bias. Matching of the sample was performed based on lagged indicators of wealth and income for the past two periods, thus aiming to artificially ensure the parallel trends between both groups. Furthermore, including vector of demographic characteristics X is called to further eliminate differences between the groups and to decrease standard errors of the estimated DID coefficient δ.

3.2.4.2 Event study This method is also referred to as a difference-in-differences lead-lag analysis, or dynamic DID (Gardner and Oswald 2006). The event study DID focuses on changes in the outcome variable in multiple periods before and after the event of gray divorce. The regression approach is based on the following equation, where, τ indicates specific period around the time when treated individual becomes divorced; other variables are used in the same meaning as mentioned previously.

$$Y_{it} = \alpha + \Sigma_{\tau=-n}^{n} \delta_{\tau} D_{div} + \beta X_{it} + \epsilon_{it} \tag{9}$$

Event study DID is useful as it allows exploration of change over multiple periods around

the time when 'treated' individuals become divorced. This is useful in order to explore whether certain changes persist into the long-run, or whether they start with the 'gray' divorce but then vanish over the several years afterwards. For example, as is evidenced in this study, happiness decreases in the years right before the divorce but then it exhibits renewal after the divorce [3].

The lead-lag relation between the divorce and net worth is also explored graphically, by plotting conditional mean levels of net worth (along with the standard errors) for the 'treated' and 'counterfactual' groups over the time around the event time of turning divorced. In this case, 'counterfactual' was obtained as the fitted values from the model by assuming that the 'gray' divorce did not take place.

3.2.4.3 Propensity score matching As was previously mentioned, propensity score matching (PSM) was used for construction of the sample for the subsequent causal analysis. Matching is a procedure to select observations into the control group with the goal of ensuring that they are similar to those in the treated group in all important respects other than the treatment (i.e. 'gray divorce'). Ho et al. (2007a) indicates that propensity score matching strives to summarize all the explanatory variables (matching variables) with a single 'propensity score' variable, which is estimated for each individual as a probability of being selected into the treated group. Thus, implementation of PSM-based sample construction involves two steps. On the first step, a probability of assignment into the treated group is estimated for each individual in the sample. On the second step, the best match is selected for each 'treated' observation using the nearest neighbor method (Angrist and Pischke 2009, Ho et al. 2007b). The main limitation of the PSM methodology is that the list of matching variables may omit certain vital variables that are not be observed. Additionally, there can be narrow support problem, when there is lack of overlap by the propensity scores distribution between the treated and control groups. This was not the problem for the current analysis. The size of the pool for selection of 'control' observations was many times larger as the pool of the 'treated' observations, and there was significant overlap by propensity scores.

[3] In case of Gardner and Oswald (2006) who studies the effect of divorces among younger cohorts, the result was somewhat different. Gardner and Oswald (2006) found that happiness after the divorce came to a higher level as compared to its level before the divorce

3.3 Findings

The obtained causal inference results are organized in several tables and plots. First, the baseline DID results are presented in table 3.3. Then detailed DID results are exhibited with using detailed wealth outcomes that are organized for sub-samples by gender, age group, and income level (tables 3.4 and 3.5 - for intensive and extensive margins). Third, findings from event study DID are provided to evaluate the dynamic of wealth changes around the divorce event (table 3.6, and figures 3.2 - 3.3). Fourth, impact of gray divorce on additional economic and well-being outcomes was explored (table 3.7). Finally, feasibility testing of the results is performed by considering the effect of outliers, as well as by exploring whether the results remain unchanged when the sample is sub-divided by respondent cohort and wave of the survey (table 3.8).

3.3.1 Impact of gray divorce using DID

The basis DID results are provided in the next table, which are the basic results for evaluation of the gray divorce effect on individual wealth and wealth components. The results provide estimates for the lasting effect, as opposed to the temporal effects evidenced with the event study DID. The baseline results involve the following four key measures of wealth: net worth, housing equity, non-housing wealth and total financial assets. If 'gray divorce' has a lasting negative impact on individual wealth, it would be captured by a negative and significant DID coefficient. There are negative and significant coefficients for housing equity and financial assets per person. These indicate that financial assets and housing equity decline and do not subsequently restore following the 'gray' divorce. For the net worth there is negative but insignificant coefficient of -0.022 (similar result that is negative but not significant is observed for the self-reported happiness status). This indicates that total net worth per person does not get significantly affected by gray divorce in the long run. Yet net worth can still experience decline right following the divorce, which is further explored using the event study DID approach.

Table 3.3: Baseline DID results

	(1) Net worth	(2) Housing eq.	(3) Non-hous. eq.	(4) Fin. assets	(5) Happy
$D_{tr}D_{time}$	-0.022	-0.104***	-0.003	-0.925***	-0.021
	(-1.401)	(-4.998)	(-0.183)	(-4.023)	(-0.867)
D_{tr}	0.003	0.034*	-0.002	0.003	-0.061***
	(0.180)	(1.683)	(-0.180)	(0.013)	(-2.859)
D_{time}	0.007	0.048**	-0.000	0.137	0.020
	(0.799)	(2.204)	(-0.036)	(0.660)	(1.001)
Age	0.001	0.001	0.001	0.035**	0.001
	(1.546)	(0.743)	(1.518)	(2.090)	(0.534)
Female	-0.001	0.012	-0.002	0.056	-0.006
	(-0.065)	(0.853)	(-0.213)	(0.256)	(-0.308)
Constant	14.296***	12.604***	14.294***	5.971*	0.694***
	(227.715)	(161.808)	(278.715)	(1.730)	(3.504)
Observations	7,012	7,012	7,012	7,012	6,285
R-squared	0.051	0.062	0.038	0.213	0.019

All dependent variables, except 'Happy' are in per equivalent person basis in logarithms. Robust t-statistics in parentheses. Sig: *** p<0.01, ** p<0.05, * p<0.1.

After the negative implications of gray divorce on housing wealth and financial assets was established from the above set of results, the following two tables provide in-depth exploration of wealth categories (these vary across table rows) for total sample and for subsets obtained from grouping by gender, age and income level (samples vary by columns). Such approach is helpful in determining the cohorts that experience the heaviest decline in the value of certain wealth categories. While table 3.4 provide the results using intensive margin of the dependent variables, the findings for extensive margin are summarized in table 3.5.

Table 3.4: DID results by detailed outcomes: Intensive margin

Dep vars	(1) Total sample	(2) Male	(2) Female	(3) Age $\leq$ 60	(3) Age $>$ 60	(4) Low inc.	(4) High inc.
Happy	-0.021	-0.01	-0.025	-0.041	-0.049	-0.004	-0.029
	(-0.867)	(-0.273)	(-0.849)	(-1.14)	(-1.342)	(-0.122)	(-0.826)
Net worth	-0.022	-0.018	-0.026	-0.045**	-0.005	0.009	-0.047***
	(-1.401)	(-1.099)	(-1.113)	(-2.139)	(-0.229)	(0.395)	(-2.791)
Housing eq.	-0.104***	-0.121***	-0.092***	-0.179***	-0.06***	-0.069**	-0.124***
	(-4.998)	(-3.141)	(-4.328)	(-3.359)	(-2.606)	(-2.4)	(-4.953)
Non-hous. eq.	-0.003	0.002	-0.007	-0.015	0.007	0.018	-0.02
	(-0.183)	(0.146)	(-0.339)	(-0.879)	(0.313)	(0.811)	(-1.517)
Fin. assets	-0.925***	-1.139***	-0.894***	-1.015***	-0.726**	-0.213	-1.487***
	(-4.023)	(-3.25)	(-2.998)	(-2.61)	(-2.022)	(-0.698)	(-4.368)
Vehicles	-1.39***	-1.248***	-1.48***	-1.103***	-2.24***	-1.295***	-1.497***
	(-6.642)	(-4.026)	(-5.345)	(-3.566)	(-7.268)	(-5.033)	(-5.136)
Businesses	-0.141**	-0.008	-0.239***	-0.045	-0.23**	-0.143	-0.157*
	(-2.156)	(-0.095)	(-2.595)	(-0.504)	(-2.323)	(-1.532)	(-1.845)
IRAs	-0.877**	-0.57	-1.094***	-0.353	-1.056**	-1.371***	-0.646
	(-2.528)	(-0.949)	(-2.621)	(-0.607)	(-2.148)	(-2.612)	(-1.463)
Stocks	-0.788***	-0.647	-0.982***	-1.022**	-0.733**	-0.464	-1.154***
	(-2.934)	(-1.543)	(-2.857)	(-2.142)	(-2.014)	(-1.144)	(-3.083)
Bank accs.	-0.84***	-0.996***	-0.83***	-0.589	-0.829**	-0.229	-1.273***
	(-3.982)	(-2.99)	(-3.072)	(-1.627)	(-2.552)	(-0.833)	(-4.164)
CDs	-0.454**	-0.484	-0.521**	-0.443	-0.061	-0.45	-0.334
	(-2.092)	(-1.333)	(-1.994)	(-1.23)	(-0.215)	(-1.459)	(-1.169)
Bonds	-0.078	-0.004	-0.147	-0.061	-0.035	-0.093	-0.059
	(-0.583)	(-0.016)	(-0.945)	(-0.305)	(-0.195)	(-0.448)	(-0.347)
Other assets	-0.107	-0.204	-0.103	-0.209	-0.115	-0.311	-0.222
	(-0.561)	(-0.725)	(-0.406)	(-0.675)	(-0.388)	(-0.946)	(-0.883)
Other debt	-0.094	-0.389	0.122	0.358	-0.121	-0.62	0.179
	(-0.373)	(-0.99)	(0.374)	(0.899)	(-0.303)	(-1.547)	(0.547)

Each number is a Difference-in-differences coefficient from a separate regression (with control variables, time and cohort fixed effects). All dependent variables, except 'Happy' are in per equivalent person basis in logarithms. Columns show considered sub-set, rows indicate outcome variables. T-statistics are provided in parentheses. Sig: *** p<0.01, ** p<0.05, * p<0.1.

'Gray' divorce has lasting negative impact on net worth in case of mid-aged individuals (age $\leq$ 60) and in case of high income individuals [4]. The negative effect on Housing equity is significant in the total sample and each category. The magnitude of the coefficient is the smallest for those in senior years (Age $>$ 60) and those in the low-income category. The effect on vehicles and in businesses indicates significant decline in the case of women and those aged 60+. The gray divorce-induced decline in financial assets is the largest for high-income individuals, for males and those aged below 60. Among all the financial assets, it is bank account and stocks that experience the most pronounced decline. In the case with stocks,

[4]Low and High income categories are those, respectively, below and above the median earnings income for the treated category of $4,900

the decline is higher for females, individuals aged below 60, and those in the high-income group. In case with the bank accounts, the decline is also greater for those in the high-income category. These results can be driven by the fact that low-income individuals most of the time do not have any stocks. Non-housing equity, bonds, other assets and other debts are not significantly impacted by 'gray' divorce for either category. Individual retirement accounts (IRA) exhibit significant decline in the total sample, and the most significant decline is evidenced in the low income subset, as well as senior aged individuals and women.

Substantial proportion of individuals in the sample have zero holdings of certain types of financial and other assets. Moreover, the proportion of those without assets and without positive net worth tends to be higher after gray divorce. Specifically, after the 'gray' divorce 13.7 % of 'treated' individuals lost their positive net worth (8.4% did not have positive net worth prior to the divorce), 28.5 % lost their positive housing equity (23.5% did not have one), 14.8% lost financial assets (6.4% did not have it). Thus, addressing these consideration, extensive margin DID analysis was performed and the results are provided in the following table. The table is organized in a manner that is equivalent to the table with the intensive margin results.All outcome variables were coded as a dummy variable that equals 1 for a higher than zero value of the given wealth category and 0 otherwise. The estimated DID gray' divorce coefficients when using extensive margin are negative and most of then are significant. This indicates that gray divorce is frequently associated with the 'treated' individuals loosing their assets and wealth. The coefficients from the extensive margin analysis tend to have higher statistical significance as compared to their counterparts from the intensive margin analysis. For example, there is significant negative impact on net worth in the total sample as well as in each sub-category (except the low-income group many of whom did not have positive wealth before the divorce). This same reasoning explains the lack of significant result for the low income group in case with stocks and bank accounts. Among various wealth categories, the highest decline is evidenced for housing equity and vehicles. Gray divorce seems to induce individuals to sell their vehicles, especially those aged 60+ and those in high-income category. The strongest negative impact of gray divorce on IRAs is for women, low-income group and those aged 60+. Therefore the elderly (aged 60+) and women are the groups with the highest risk to loose their IRAs due to gray divorce. In case

with bonds, other assets and other debts, there is no evidence of a significant negative effect of gray divorce using extensive margin.

Table 3.5: DID results by detailed outcomes: Extensive margin

	(1)	(2)		(3)		(4)	
Dep vars	Total sample	Male	Female	Age $\leq$ 60	Age $>$ 60	Low inc.	High inc.
Net worth	-0.065***	-0.071***	-0.062**	-0.064***	-0.11***	-0.031	-0.083***
	(-3.591)	(-2.649)	(-2.516)	(-2.968)	(-3.403)	(-1.272)	(-3.267)
Housing eq.	-0.267***	-0.296***	-0.254***	-0.277***	-0.254***	-0.187***	-0.317***
	(-8.456)	(-6.342)	(-5.963)	(-5.332)	(-5.407)	(-4.37)	(-7.067)
Non-hous. eq.	-0.077***	-0.056*	-0.092***	-0.068**	-0.139***	-0.05*	-0.102***
	(-3.695)	(-1.733)	(-3.341)	(-2.431)	(-4.019)	(-1.798)	(-3.346)
Fin. assets	-0.077***	-0.056*	-0.092***	-0.068**	-0.139***	-0.05*	-0.102***
	(-3.695)	(-1.733)	(-3.341)	(-2.431)	(-4.019)	(-1.798)	(-3.346)
Vehicles	-0.124***	-0.121***	-0.122***	-0.121***	-0.198***	-0.102***	-0.147***
	(-5.495)	(-3.646)	(-4.023)	(-3.626)	(-5.892)	(-3.893)	(-4.632)
Businesses	-0.044**	-0.031	-0.057**	-0.024	-0.067**	-0.057*	-0.039
	(-2.07)	(-0.991)	(-1.997)	(-0.782)	(-2.121)	(-1.897)	(-1.371)
IRAs	-0.088***	-0.057	-0.109***	-0.041	-0.113**	-0.134***	-0.07*
	(-2.636)	(-0.997)	(-2.644)	(-0.748)	(-2.369)	(-2.62)	(-1.65)
Stocks	-0.076***	-0.067*	-0.092***	-0.096**	-0.067*	-0.051	-0.106***
	(-2.889)	(-1.653)	(-2.655)	(-2.161)	(-1.86)	(-1.261)	(-2.941)
Bank accs	-0.069***	-0.083**	-0.062**	-0.066*	-0.062*	-0.013	-0.117***
	(-3.139)	(-2.321)	(-2.225)	(-1.84)	(-1.73)	(-0.445)	(-3.569)
CDs	-0.057**	-0.062	-0.063**	-0.067*	-0.014	-0.068*	-0.039
	(-2.341)	(-1.532)	(-2.172)	(-1.689)	(-0.41)	(-1.856)	(-1.274)
Bonds	-0.009	-0.004	-0.014	-0.004	-0.008	-0.01	-0.008
	(-0.64)	(-0.195)	(-0.876)	(-0.225)	(-0.424)	(-0.482)	(-0.465)
Other assets	-0.015	-0.029	-0.013	-0.019	-0.019	-0.035	-0.031
	(-0.77)	(-0.984)	(-0.492)	(-0.572)	(-0.588)	(-0.979)	(-1.168)
Other debt	0.006	-0.037	0.037	0.047	0.002	-0.053	0.034
	(0.192)	(-0.78)	(0.965)	(0.977)	(0.036)	(-1.097)	(0.873)

Each number is a Difference-in-differences coefficient from a separate regression (with control variables, time and cohort fixed effects). All dependent variables are dummies indicating whether the amount is a positive number. Columns show sub-samples, rows indicate different outcome variables. T-statistics are provided in parentheses. Sig: *** p<0.01, ** p<0.05, * p<0.1.

3.3.2 Impact of gray divorce using event study

The above DID analysis indicated that in some cases there was no evidence of significant lasting negative impact of gray divorce on wealth. For example, in total sample gray divorce did not have significant negative impact on the amount of net worth. While the DID method aims to identify a long-term result, event study DID aims to evaluate the effect in different periods around the time of 'gray' divorce. Visual evaluation using event study DID is provided below for selected outcome variables, wile detailed regression-based event study DID findings for each of the considered wealth categories are summarized in the table.

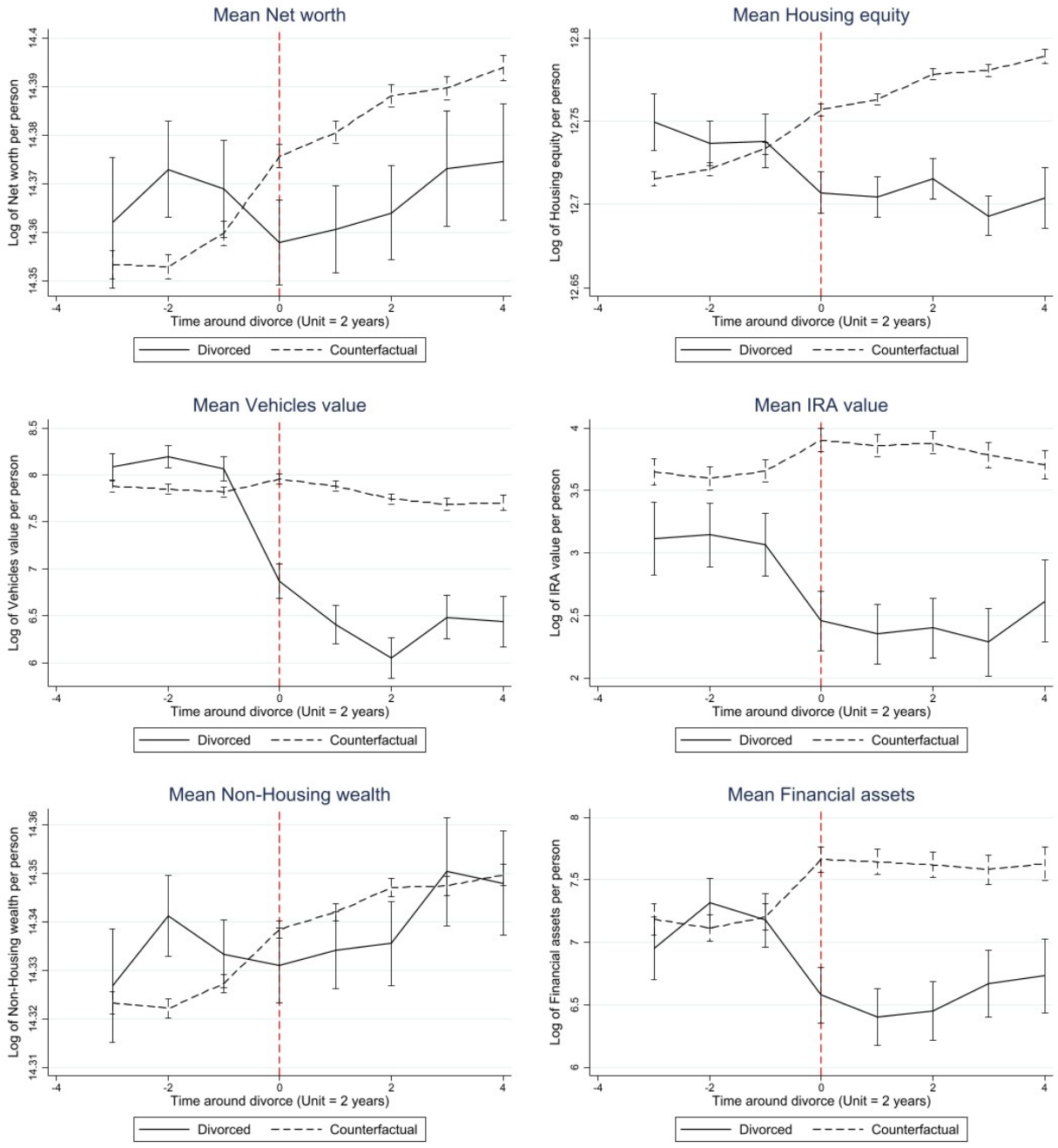
Housing equity and financial assets tend to experience the greatest gap between the divorced and their counterfactual. There is negative net worth outcome observed for the divorced in comparison to the counterfactual, while non-housing wealth does not exhibit as large a difference (figure 3.3).

For the purpose of comparison to the study by Gardner and Oswald (2006), dynamics of happiness for the treated group and its counterfactual were explored as well. In line with the results in Gardner and Oswald (2006), happiness achieved its lowest point in the period right after the divorce, but then it revives in subsequent periods. Unlike the findings in Gardner and Oswald (2006), which showed increase in the level of happiness after the divorce in case of the younger groups, my findings only show that the level of happiness only revives to its level that was observed prior to the 'gray' divorce (figure 3.2).

Figure 3.2: Dynamics of feeling happy around gray divorce

Note: Divorced are 'treated' individuals, while counterfactual are the predicted values from the dynamic DID regression model assuming these are not 'treated' individuals by setting $D_{div} = 0$

Figure 3.3: Dynamics of net worth and asset categories around gray divorce



Note: Divorced are 'treated' individuals, while counterfactual are the predicted values from the dynamic DID regression model assuming these are not 'treated' individuals by setting $D_{div} = 0$

Detailed dynamic DID results for each considered category of wealth are provided in table 3.6 below. For each outcome variable, the table provides a set of DID coefficients in every time period within 4 waves before and 4 waves after the 'gray' divorce event. Significant and positive coefficients in the times prior to the time of 'gray' divorce T combined with the significant negative coefficients in the periods following the divorce suggest negative wealth implications starting prior to the actual divorce [5]. In case of net worth such pre-trends can indicate that decline in wealth started somewhat prior to 'gray' divorce. In case of the happiness status, the pre-trend can be associated with higher levels of stress and decline in happiness as the couple was already on its way to divorce. These pre-trends are evidenced for such outcomes as net worth, housing equity, non-housing equity, vehicles and businesses, as well as happiness status.

The dynamic DID coefficients are helpful for evaluation of timing of the decline in value of various wealth components due to the divorce. Negative impact of gray divorce on net worth is evidenced for the first three periods after the divorce (e.g. from T to $T + 2$). Similar patterns of decline and subsequent revival were evidenced for businesses, stocks and certificates of deposit (CDs). While in case of housing equity, vehicles, and financial assets, and especially bank accounts and IRAs, negative effects are permanent and do not fully recover.

[5] Matching was performed using only the first and second lags, but not higher order lags, of total income and total assets. Thus some higher order lags are significant. Furthermore, as several variables were used for matching, completely perfect match was not achieved in case of net worth and some pre-trend is evidenced in the second lag of net worth. The level of happiness was not included as a matching variable.

Table 3.6: Event study evaluation of gray divorce effect on wealth outcomes

	(1) Net worth	(2) Housing eq.	(3) Nonhous. eq.	(4) Fin. assets	(5) Vehicles	(6) Business	(7) IRA	(8) Stocks	(9) Bank accs.	(10) CDs	(11) Bonds	(12) Happy
$D_{tr}D_{T-4}$	-0.053	0.029	-0.058	-0.153	0.029	0.175**	-0.516	0.255	-0.263	-0.613***	0.063	-0.033
	(-0.665)	(1.364)	(-0.727)	(-0.563)	(0.148)	(2.068)	(-1.474)	(0.770)	(-1.033)	(-2.621)	(0.367)	(-1.008)
$D_{tr}D_{T-3}$	0.011	0.031*	0.006	-0.234	0.191	0.015	-0.455	-0.035	-0.225	-0.419**	-0.094	-0.028
	(0.945)	(1.655)	(0.582)	(-0.954)	(1.179)	(0.273)	(-1.465)	(-0.127)	(-0.989)	(-2.040)	(-0.755)	(-0.982)
$D_{tr}D_{T-2}$	0.023**	0.014	0.021**	0.193	0.362**	0.044	-0.386	0.255	0.075	-0.107	0.013	-0.102***
	(2.130)	(0.934)	(2.322)	(0.904)	(2.549)	(0.855)	(-1.315)	(1.028)	(0.366)	(-0.538)	(0.107)	(-3.844)
$D_{tr}D_{T-1}$	0.011	0.005	0.008	0.027	0.284*	0.041	-0.521*	0.152	0.136	-0.128	-0.111	-0.109***
	(1.087)	(0.306)	(1.007)	(0.123)	(1.843)	(0.795)	(-1.771)	(0.597)	(0.679)	(-0.629)	(-1.009)	(-4.281)
$D_{tr}D_{T}$	-0.018*	-0.050***	-0.007	-0.962***	-1.025***	-0.097**	-1.383***	-0.757***	-0.894***	-0.637***	-0.130	-0.150***
	(-1.859)	(-3.455)	(-0.867)	(-4.171)	(-5.097)	(-2.170)	(-4.941)	(-3.157)	(-4.202)	(-3.342)	(-1.069)	(-5.790)
$D_{tr}D_{T+1}$	-0.018*	-0.057***	-0.006	-1.096***	-1.394***	-0.064	-1.457***	-0.930***	-0.923***	-0.462**	-0.179	-0.088***
	(-1.908)	(-3.867)	(-0.769)	(-4.789)	(-6.550)	(-1.376)	(-5.293)	(-4.009)	(-4.384)	(-2.413)	(-1.566)	(-3.603)
$D_{tr}D_{T+2}$	-0.022**	-0.056***	-0.011	-0.989***	-1.613***	-0.085**	-1.372***	-0.726***	-0.723***	-0.564***	-0.069	-0.046**
	(-2.348)	(-4.166)	(-1.269)	(-4.251)	(-7.254)	(-2.266)	(-4.927)	(-3.047)	(-3.427)	(-3.150)	(-0.559)	(-1.967)
$D_{tr}D_{T+3}$	-0.012	-0.081***	0.006	-0.695***	-1.108***	-0.062	-1.398***	-0.338	-0.693***	-0.249	-0.110	-0.079***
	(-1.048)	(-5.835)	(0.584)	(-2.782)	(-4.621)	(-1.283)	(-4.663)	(-1.277)	(-2.931)	(-1.238)	(-0.919)	(-2.957)
$D_{tr}D_{T+4}$	-0.010	-0.073***	0.004	-0.591**	-1.128***	-0.023	-0.958***	-0.273	-0.516**	-0.268	0.147	-0.018
	(-0.850)	(-3.795)	(0.388)	(-2.153)	(-4.228)	(-0.389)	(-2.841)	(-0.923)	(-2.047)	(-1.217)	(0.856)	(-0.646)
Covariates	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Wave and Cohort FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
R-squared	0.053	0.060	0.040	0.214	0.107	0.016	0.087	0.101	0.168	0.054	0.041	0.024

All dependent variables, except 'Happy' are in per equivalent person basis in logarithms. Robust t-statistics in parentheses.
Sig: *** p<0.01, ** p<0.05, * p<0.1. Number of observations: 6,660.

3.3.3 Related well-being implications of gray divorce

Changes in wealth and its components due to gray divorce can be accompanied by changes in other economic and well-being indicators. This section explores whether 'gray' divorce affected a number of such additional outcome variables, including labor force participation, earnings income, capital income, and indicators of human capital, such as cognitive ability and memory. Obtained dynamic DID estimates indicate that after 'gray' divorce occurred, about 11 - 12 % individuals quit their retirement to undertake some form of employment. For this reason earnings income tends to increase among the 'treated' individuals.

Table 3.7: Event study evaluation of gray divorce effect on selected outcomes

	(1) Retired	(2) In labor force	(3) Earnings	(4) Capital income	(5) Cognitive scores	(6) Memory test
$D_{tr}D_{T-4}$	-0.015	0.002	-0.082	0.289	0.096	-0.046
	(-0.183)	(0.051)	(-0.228)	(0.882)	(0.165)	(-0.154)
$D_{tr}D_{T-3}$	0.059	0.010	0.100	-0.370	-0.596	-0.380
	(0.840)	(0.323)	(0.313)	(-1.330)	(-1.109)	(-1.476)
$D_{tr}D_{T-2}$	-0.025	0.028	0.075	0.243	-0.291	-0.181
	(-0.431)	(1.008)	(0.263)	(0.944)	(-0.625)	(-0.828)
$D_{tr}D_{T-1}$	0.051	0.012	0.089	-0.291	0.515	0.085
	(0.831)	(0.411)	(0.307)	(-1.154)	(0.980)	(0.393)
$D_{tr}D_{T}$	0.005	0.023	0.210	-1.025***	-0.488	-0.256
	(0.090)	(0.791)	(0.718)	(-4.239)	(-1.149)	(-1.268)
$D_{tr}D_{T+1}$	-0.097*	0.020	0.265	-1.139***	-0.755*	-0.368*
	(-1.870)	(0.695)	(0.929)	(-4.890)	(-1.730)	(-1.868)
$D_{tr}D_{T+2}$	-0.098*	0.027	0.463	-1.093***	-1.262***	-0.568***
	(-1.934)	(0.954)	(1.644)	(-4.872)	(-3.097)	(-2.856)
$D_{tr}D_{T+3}$	-0.115**	0.048	0.519*	-0.673***	-0.624	-0.142
	(-2.251)	(1.526)	(1.651)	(-2.695)	(-1.462)	(-0.665)
$D_{tr}D_{T+4}$	-0.075	0.033	0.463	-0.466	-0.258	-0.087
	(-1.350)	(0.943)	(1.328)	(-1.643)	(-0.620)	(-0.358)
Covariates	yes	yes	yes	yes	yes	yes
Wave and Cohort FEs	yes	yes	yes	yes	yes	yes
No. obs	6,758	6,931	7,012	7,012	3,119	5,916
R-squared	0.103	0.212	0.198	0.098	0.192	0.175

All income variables are in logarithms. Robust t-statistics in parentheses. Sig: *** p<0.01, ** p<0.05, * p<0.1.

In addition, there is evidence of a significant decline in capital income, which can be attributed to the previously reported depletion of financial assets among the divorced. Also such decline of capital income can be due to decline in cognitive ability and memory after the 'gray' divorce. In other words, deterioration of mental capability due to 'gray' divorce may be an important source of subsequent wealth deterioration, as individuals tend to make

ineffective investment decisions and are more prone to financial fraud. Such decline in mental ability tends to be temporary and it vanishes three periods following the 'gray' divorce. Hence treated individuals would benefit by limiting their financial decision-making during the first several years following their 'gray' divorces, or preferably by seeking a professional financial advice. The results in this part are in line with findings obtained by Mazzonna et al. (2018) who studied wealth implications of mental ability decline among the elderly.

3.3.4 Feasibility tests

A number of feasibility tests were conducted to verify the findings. These evaluate whether the obtained results are not completely driven by the outliers, and explore whether the findings are similar across waves and cohorts. The feasibility tests results using total financial assets as the outcome variable are shown below.

Table 3.8: Effect of gray divorce on financial assets: feasibility tests summary

	(1) Baseline results	(2) Winsorized at 1%	(3) Winsorized at 5%	(4) Waves 1-7	(5) Waves 8-13	(6) Cohorts 1- 4	(7) Cohorts 5-6
$D_{tr}D_{T-4}$	-0.153	-0.117	-0.125	-0.104		-0.495	1.236**
	(-0.563)	(-0.449)	(-0.484)	(-0.395)		(-1.644)	(2.082)
$D_{tr}D_{T-3}$	-0.234	-0.167	-0.173	-0.153	-0.469	-0.389	0.263
	(-0.954)	(-0.638)	(-0.665)	(-0.582)	(-0.768)	(-1.386)	(0.516)
$D_{tr}D_{T-2}$	0.193	0.437*	0.414*	0.448*	-0.405	0.247	0.018
	(0.904)	(1.838)	(1.769)	(1.875)	(-0.889)	(1.046)	(0.037)
$D_{tr}D_{T-1}$	0.027	-0.150	-0.147	-0.151	0.290	-0.056	0.263
	(0.123)	(-0.531)	(-0.526)	(-0.536)	(0.791)	(-0.224)	(0.567)
$D_{tr}D_{T}$	-0.962***	-0.717**	-0.711**	-0.719**	-1.160***	-0.869***	-1.232***
	(-4.171)	(-2.178)	(-2.176)	(-2.181)	(-3.400)	(-3.290)	(-2.641)
$D_{tr}D_{T+1}$	-1.096***	-1.034***	-1.017***	-1.036***	-1.118***	-1.048***	-1.311***
	(-4.789)	(-2.704)	(-2.678)	(-2.710)	(-3.653)	(-3.992)	(-2.837)
$D_{tr}D_{T+2}$	-0.989***	-0.350	-0.327	-0.352	-1.250***	-1.049***	-0.888*
	(-4.251)	(-0.806)	(-0.754)	(-0.811)	(-4.256)	(-3.903)	(-1.910)
$D_{tr}D_{T+3}$	-0.695***	-0.150	-0.150	-0.150	-0.905***	-0.686**	-0.733
	(-2.782)	(-0.289)	(-0.291)	(-0.288)	(-3.002)	(-2.395)	(-1.452)
$D_{tr}D_{T+4}$	-0.591**	0.407	0.410	0.402	-0.835***	-0.725**	-0.140
	(-2.153)	(0.606)	(0.616)	(0.598)	(-2.814)	(-2.354)	(-0.238)
Covariates	yes	yes	yes	yes	yes	yes	yes
Wave and Cohort FEs	yes	yes	yes	yes	yes	yes	yes
No. Obs.	7,012	3,537	3,537	3,537	3,475	5,391	1,621
R-squared	0.214	0.188	0.186	0.189	0.245	0.191	0.318

Dependent variable is log of financial assets per equivalent person. Robust t-statistics in parentheses. Sig: *** p<0.01, ** p<0.05, * p<0.1.

If the significant negative impact of 'gray divorce' remains in all the considered feasibility tests, this would contribute towards greater confidence in the obtained findings. To explore the effect of outliers, the dependent variable was winsorized at 1% and 5% levels. All the values outside of the mentioned cut-points (e.g. those below 1st and above 99th centiles with 1% winsorizing) were forced to equal the limiting values. Feasibility test results indicate that the negative effect of gray divorce on financial assets is evidenced in each case after the winsorizing. Moreover, the negative effect of 'gray' divorce is evidenced in each category by cohort and wave. Some differences are present though, as in the earlier waves 'gray' divorce impact on financial wealth had somewhat lower magnitude and was shorter-lived in comparison to the later waves.

3.4 Prediction of Marital Dissolution Among Elderly

3.4.1 Prediction analysis methodology

The goal of the section is to explore the role of different economic, demographic and health variables to predict probability of divorce among mid- and senior-aged couples (aged 50+). Thus this study in addition to evaluating the effects of gray divorce on wealth, also explores whether wealth and wealth components are significant predictors of divorce among couples aged 50 and above. This subsection starts with the review of the key elements of the forecasting machine learning methodology. Then the next two subsections provide estimated results that describe the significance of the considered variables, and compare predictive capability of different forecasting models.

3.4.1.1 Sample and features The sample for prediction analysis is constructed from the same HRS panel dataset by retaining only those households that are married in the given period and who either remain married or become divorced in the next period. There were few same gender married individuals (less than 0.5% of the total sample) and these were excluded from the analysis. The resulting sample includes 128,260 observations of which

there are 1,697 those who divorced in the next period. The analysis in this section includes the following predictive features (explanatory variables).

Table 3.9: Summary of features to predict gray divorces

Demographic	**Health**	**Economic**
Length of current marriage*	Health*	In labor force*
Age*	BMI*	Earnings*
Marriages*	Depression*	Pension*
Education*	Smoke*	Net worth
Persons in HH	Drink*	House equity
Children	Lonely*	Financial assets
Race*	Happy*	
Religion*		

For each variable indicated in the table with '*' three values are included - value for the husband, wife, and difference between the two.

The objective of the current section is to forecast gray divorce, the emphasis is on the out-of-sample performance of the considered models. For this purpose the total sample was randomly split into two - training part (80% of the sample) and testing part (20 %). All predictive models that are detailed further, were estimated and their hyperparameters (such as value of Lasso penalty parameter) were estimated using 5-fold cross validation. According to Gareth et al. (2013), k-fold cross validation is a methodology to estimate a predictive model and its hyper-parameters that is aimed to avoid over-fitting.

3.4.1.2 Predictive models This part performs prediction of the next marital status for the currently married couples. This is expressed with the following general model.

$$P(y = 1|\mathbf{X}) = F(\mathbf{X}) = F(X_{DEM}, X_{ECON}, X_{HEALTH}) \tag{10}$$

Predictive analysis in this section involves several models: regularized logistic regression, classification tree, random forest, and artificial neural network. The models differ by functional form and complexity of the predictive function $F(\mathbf{X})$. Each of these models has its advantages and limitations.

Logistic regression with a lasso regularization term is estimated using maximum likelihood method with an objective function that includes the maximum likelihood of a plain logistic function with an additional regularization term, Lasso penalty term: $-lnL + \lambda\Sigma_{j=1}^{P}|\beta_j|$

(Gareth et al. 2013). There are alternatives to lasso regularization term, such as ridge regularization term. Of the two terms, ridge has quadratic penalty form, while lasso has penalty term in absolute value. In addition, third alternative is elastic net regularization that combines both lasso and ridge penalties. Logistic regression is a linear model that can be conveniently presented by indicating its marginal effects for all the explanatory coefficients. Its limitations is lack of consideration for non-linear effects, lack of interaction among the variables and relatively low forecast power as compared to other models.

Classification tree is another straightforward models that can be presented in a convenient form graphically. Classification trees use features (explanatory variables) to sub-divide the feature space into segments that are then classified based on the majority class in that segment (Gareth et al. (2013)). Tree model is constructed by selecting regions that minimize either Gini index or entropy measure (ibid). Classification tree is convenient for the use by a human as it is straightforward to apply in practice. Yet, forecast performance of a single classification tree is limited.

Finally, random forest and artificial neural network models are two highly non-linear models, which have proven high performance in forecasting (Geron 2017). Of these, random forest is a collection of classification tree models that generates majority vote from various trees. Neural network model has a number of hidden layers with elements being forecasts from models in the previous layer of the neural network. The main limitation of these two models is that they technically are 'black boxes' as each involves hundreds and thousands of estimated coefficients that cannot be conveniently presented and interpreted.

3.4.1.3 Performance evaluation With the objective of providing effective forecast, several indicators are used for forecast performance evaluation. These are recall, precision, F_1-*score* and accuracy, all are based on the confusion table. All these measured have domain between zero and one, with high performance models approaching one.

$$Recall = \frac{True\ positive}{True\ positive + False\ negative}$$

$$Precision = \frac{True\ positive}{True\ positive + False\ positive}$$

$$F_1\text{-}score = 2\frac{(Precision \cdot Recall)}{Precision + Recall}$$

$$Accuracy = \frac{True\ positive + True\ negative}{Sample\ size}$$

Hence, Recall shows the percentage of actual divorces that the model is able to correctly classify. This is equivalent to one minus probability of type II error [6]. Precision indicates percent of correctly classified positive cases among all cases that are classified as positive by the model. Therefore, Precision is equivalent to one minus probability of type I error. Also, F_1-*score* is a weighted average measure combining Recall and Precision. F_1-*score* takes higher values when both, Recall and Precision, are high. Accuracy simply indicates the proportion of correctly classified cases by all cases that were classified by the model. When the positive case is a rare in the dataset, relevance of Accuracy for evaluation of forecast performance is low. In case when positive class is a rare one (proportion of newly divorced individuals is below 1.5% in the data), of greater importance is the recall, precision and F_1-*score*, while relevance of accuracy is very low.

3.4.2 Estimated models results

The estimated logistic regression and classification tree models can be presented in a compact form. The current section reports the estimated marginal effects from the Lasso logistic regression, and the decision tree model, as well as feature importance analysis from random forest. Then the following section compares predictive performance of these models to conclude on their predictive power. Logistic regression with lasso penalty term was estimated for the whole sample, as well as for the sub-samples: mid-aged ($\leq$ 60 years) and senior (over 60), low-income and high-income groups. The model included 55 features, the table below presents selected features including all those that pertain to net worth and income, labor force participation and human capital, and some selected demographic ones. The obtained marginal effects indicate that net worth and housing equity are both significant predictors of gray divorces and act to lower the probability of divorce. While net worth is highly relevant for those in Low-income group, Housing equity is more relevant for households in the mid-age

[6] Type I error relates to a false positive classification, while Type II error indicates false negative classification.

category. Unlike the mentioned two wealth categories, financial assets are insignificant in predicting gray divorces. There are other relevant features - female participation in labor force has positive effect on the probability of divorce. Higher education of both spouses has positive impact, while health and BMI have negative impact on the probability of divorce. Higher number of children increases probability of divorce but to a minor extent. Number of previous marriages, as well as feeling lonely, increases probability of the divorce, while feeling happy for male is a predictor of a lower probability of 'gray' divorce. When spouses have the same religion, this decreases the probability of 'gray' divorce, and being of same race does not have any significance in predicting the divorce.

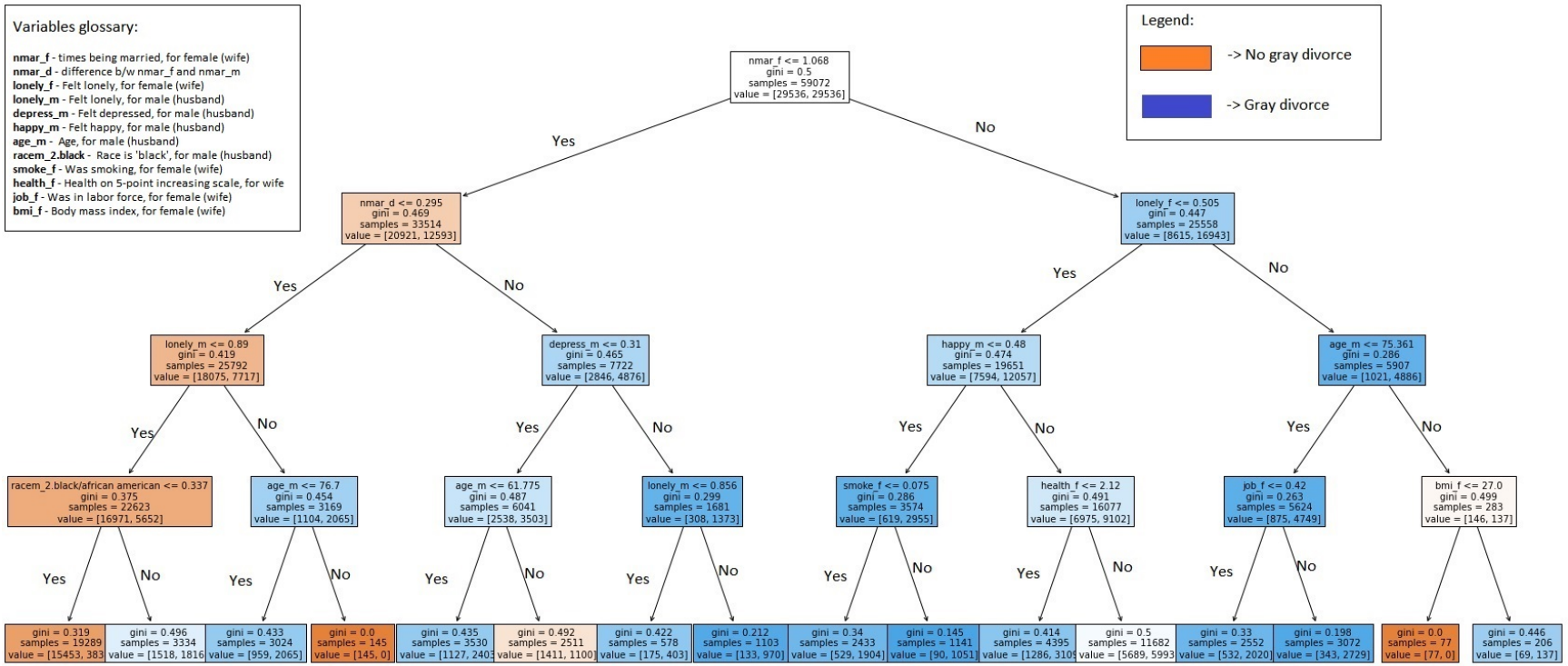
Decision tree results are provided on the next page. The tree is constructed in the way that a reader can walk through the decision nodes from top to bottom, in order to determine her predicted marital state in the next period. The tree was purposefully constrained to have maximum length of four nodes in order to fit the page, and it is not the optimal design for this model. Because of the mentioned constraint it is not the optimal classification tree (optimal design that maximizes the forecast performance has the depth of 12 nodes). The tree model shows high relevance for predicting probability of divorce of such variables as the number of previous marriages for both spouses, feeling lonely, depressed and happy, health and lifestyle related features (figure 3.4).

Table 3.10: Marginal effects in predicting gray divorce: Logit model

Selected features	(1) Total sample	(2) Age ≤ 60	(3) Age > 60	(4) Low income	(5) High income
Net worth	-0.0056*	-0.0010	-0.0027	-0.0084*	0.0013
	(-1.6728)	(-0.1466)	(-0.8316)	(-1.8536)	(0.2115)
Hous. eq.	-0.0033*	-0.0189**	-0.0011	-0.0023	-0.0083
	(-1.8405)	(-2.0293)	(-0.5830)	(-1.3267)	(-1.0451)
Fin. assets	0.0043	-0.0009	0.0040	0.0028	-0.0022
	(1.1414)	(-0.1288)	(0.7894)	(0.3447)	(-0.3331)
Lab. force M	0.0004	-0.0002	0.0012	0.0027**	-0.0009
	(0.4628)	(-0.1326)	(1.2408)	(2.5043)	(-0.5354)
Lab. force F	0.0029***	0.0037**	0.0015	0.0011	0.0047***
	(3.2109)	(2.3874)	(1.5651)	(1.0146)	(2.8594)
Educ M	0.0003**	0.0004	0.0003**	0.0004**	0.0002
	(2.4099)	(1.5047)	(1.9906)	(2.5480)	(0.7711)
Educ F	0.0005***	0.0008***	0.0002	0.0003*	0.0006**
	(2.9789)	(2.9036)	(1.1984)	(1.7937)	(2.3266)
Health M	-0.0010***	-0.0019***	-0.0002	-0.0011**	-0.0009
	(-2.6487)	(-2.9598)	(-0.5758)	(-2.5072)	(-1.3916)
Health F	-0.0003	-0.0009	0.0002	-0.0015***	0.0011*
	(-0.7951)	(-1.3602)	(0.5345)	(-3.3347)	(1.7295)
BMI M	-0.0001	-0.0002	-0.0000	0.0000	-0.0002*
	(-1.3658)	(-1.3665)	(-0.3252)	(0.3803)	(-1.8194)
BMI F	-0.0001*	-0.0002**	0.0000	-0.0001	-0.0001
	(-1.7085)	(-2.3155)	(0.6084)	(-1.3790)	(-1.0803)
Children	0.0007***	0.0011***	0.0003*	0.0007***	0.0007**
	(4.1213)	(3.4841)	(1.8233)	(3.9406)	(2.2391)
Same relig	-0.0032***	-0.0032**	-0.0027**	-0.0034***	-0.0029**
	(-3.5304)	(-2.1568)	(-2.5386)	(-3.0351)	(-2.0591)
Same race	0.0024	0.0017	0.0066**	0.0020	0.0035
	(1.4884)	(0.6798)	(2.2316)	(0.9726)	(1.3759)
Age M	-0.0003***	-0.0010***	0.0001	-0.0002**	-0.0005***
	(-5.3364)	(-7.4602)	(0.8941)	(-2.2267)	(-4.7258)
Age F	-0.0002***	-0.0002	-0.0002***	-0.0001*	-0.0002**
	(-3.0502)	(-1.4561)	(-3.3600)	(-1.8760)	(-2.0323)
N. married M	0.0030***	0.0042***	0.0015***	0.0023***	0.0039***
	(6.5778)	(5.0679)	(3.4952)	(4.5521)	(4.9771)
N. married M	0.0040***	0.0033***	0.0040***	0.0034***	0.0043***
	(8.4597)	(4.0358)	(8.5853)	(6.7086)	(5.4072)
Lonely M	0.0076***	0.0111***	0.0040***	0.0021	0.0136***
	(6.8577)	(6.0137)	(3.3155)	(1.5799)	(7.6766)
Lonely F	0.0065***	0.0099***	0.0026**	0.0015	0.0117***
	(6.0099)	(5.4721)	(2.2929)	(1.2527)	(6.5796)
Happy M	-0.0027**	-0.0028	-0.0031**	0.0003	-0.0063***
	(-2.4277)	(-1.4864)	(-2.5257)	(0.2583)	(-3.5148)
Happy F	-0.0017	-0.0016	-0.0019	-0.0016	-0.0023
	(-1.5397)	(-0.8514)	(-1.5889)	(-1.2600)	(-1.2277)
No. obs	84,388	42,948	40,948	42,218	42,170
Pseudo R^2	0.1263	0.0981	0.1660	0.1740	0.1031

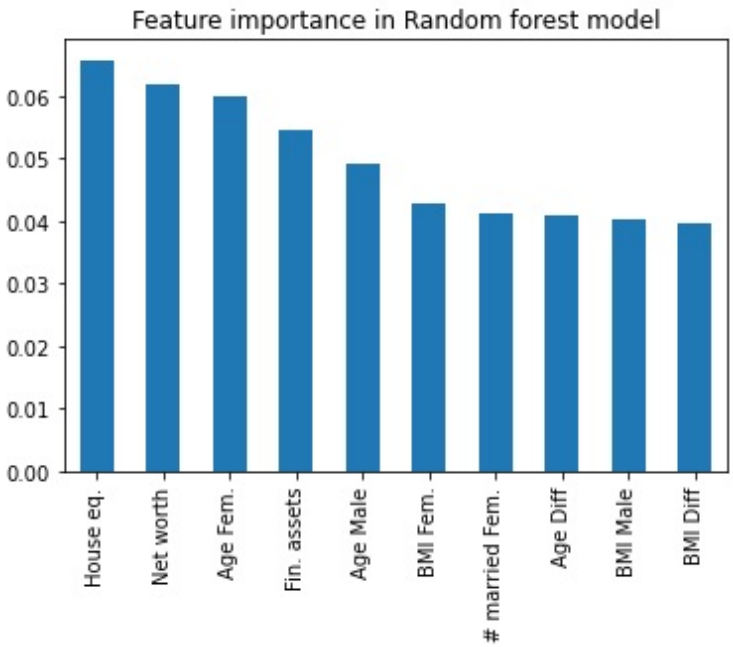
T-statistics in parentheses. Sig: *** p<0.01, ** p<0.05, * p<0.1. Wealth and income are normalized to start from 1 and are in logs (Net worth, Hous eq., Fin. assets).

Figure 3.4: Prediction of the next period marital stability for a couple



Random forest is a highly non-linear model. The model is used to explore feature importance by plotting decrease in the objective function (gini coefficient) that is achieved by separate including each variable. Two wealth indicators, net worth and housing equity, have the highest predictive power (figure 3.5). This suggests that marital stability and wealth involve the two-way relation: 'gray' divorce negatively impacts wealth of the individuals in mid and senior age, while wealth itself is a significant (negative) predictor for the probability of 'gray' divorce.

Figure 3.5: Most important predictors in Random forest model



Note: Y-axis shows decrease in impurity that is achieved by including each predictor

3.4.3 Evaluation of predictive models performance

This section compares predictive performance of the four models and concludes whether these models are able to reliably predict probability of 'gray divorce' for a household. The results are summarized in table 3.11, including the optimal model specifications, and forecast performance be each of the four considered indicators. Summary of the optimal specification for each model is provided in the 'Details' column, this contains the estimated optimal values of the hyperparameters. For example, optimal fandom forest model includes 88 trees, which

have maximum depth of 37 nodes and minimum sample leaf size of 5 observations.

Out of the four models, random forest has by far the highest performance based on F1 score and Precision. Its leading performance is followed by the artificial neural network model, then by regularized logistic regression. Classification tree shows the weakest performance among all the models. In terms of Recall (percent of true positive cases that are identified by the model) performance of logistic regression and classification tree is also high. For these models there is poor performance in terms of Precision as they have very high false positive rate and 'falsely' identify many cases as 'gray divorce'.

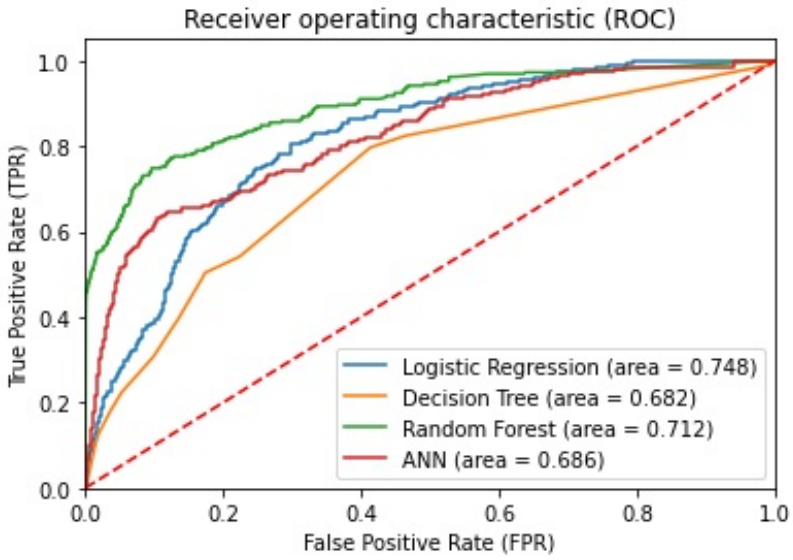
Table 3.11: Predictive model performance summary

Algorithm	**Details**	F_1**-score**	**Precision**	**Recall**	**Accuracy**
Logit + Lasso*	Penalty type: Ridge ('l2') 'l2'=551.3	0.065	0.034	0.741	0.754
Classification tree	Max depth = 4 Min samples leaf = 33	0.043	0.022	0.797	0.589
Random forest*	Trees = 88, Max depth = 37, Min samples leaf = 5	0.566	0.848	0.424	0.993
Neural network*	Hidden layer: 50, 'l2'=0.482	0.185	0.120	0.407	0.959

* Features scaled using standard scaling: $(x_i - x_{min})/(x_{max} - x_{min})$

Additional approach to evaluate performance of a model is by using receiver operating characteristic (ROC) plot (figure 3.6). It shows proportion of true positive rate and false positive rate while the classification threshold of probability moves from 0.0 to 1.0. In case of a model that can provide perfect forecasts, a ROC curve would be a vertical line along y-axis and after it achieves 1.0 (all true positive cases are correctly predicted) it then becomes a vertical line that connects points (0, 1) and (1, 1) on the plot. Area under ROC of the perfect model equals 1.0. On the other hand, a pure random model would be a 45-degree line with area under ROC curve of 0.5. All four models are shown to perform better than a purely random guess, but their performance is inferior in comparison to a hypothetical perfect model. Random forest has the strongest performance among all the considered models.

Figure 3.6: ROC curve of the predictive models



Note: Dashed line represents 'random guess' model. Higher area under the ROC curve points to higher predictive performance

3.5 Conclusion

The study explores implications of gray divorce on wealth and wealth components for those aged 50+. My analysis was motivated by the high and growing significance of this population category in accumulation of wealth in the economy. While there is lack of prior studies employing relevant identification strategies to investigate the effect of gray divorce on wealth, the current study addresses this gap. The study uses semi-parametric difference-in-differences methodology, combining the DID method with propensity score matching. For the purpose of the analysis the 'treated' group was defined to include individuals that previously were married and then changed into the divorced category. The control group included married individuals that did not change their status. For each treated individual a control individual was selected by a matching based on the lagged economic and demographic characteristics.

The results indicate that gray divorce has negative impact on net worth, mostly on such wealth components as housing equity and financial assets, primarily, bank accounts and

IRAs. The negative impact was confirmed with both intensive and extensive margin of the outcome variables. In case of the total matched sample, gray divorce increased the probability of losing positive net worth by 6.5%, and increased the probability of losing positive home equity by 26.7%. No evidence was obtained that females experience higher decline in net worth due to gray divorce compared to males. Yet, the results suggest that females are more likely to experience higher declines in their IRAs and stocks. It is also confirmed that gray divorce has a stronger impact on older individuals (aged 60+) in comparison to those in their 50s. The noted deterioration of wealth in response to gray divorce can be related to the extra costs and foregone economies of scale that are associated with divorce, as well as with deterioration of mental ability. Results from feasibility tests provided additional confirmation towards the validity of the baseline findings. Additionally, wealth was shown to be a key predictor of probability of gray divorce. Specifically, net worth and housing equity displayed potential to lower the likelihood of a divorces among the households aged 50+. The model with the best forecast performance in this study is random forest. Among all positive cases that the model identifies, 85% are correctly identified. Additionally, the model is able to identify 42.4% among the true positive cases. Future research can aim to further raise the forecast performance of the predictive model, for example by including additional features. Moreover, there is need for the future research to contrast the relevance of various features to predict divorces among those aged 50+ as compared to those aged below 50.

4.0 The Modernization Hypothesis in the Post-Socialist Economies

Central and Eastern European countries underwent a unique 'natural experiment' as they saw considerable initial democratization after the Soviet Union collapsed and the communist block dissolved. While these nations started from similar conditions in terms of post-socialist institutional legacy, their subsequent trajectories of economic and political development significantly diverged. The study aims to explore whether positive income-democracy relation, known as the 'modernization hypothesis', holds among the CEE nations, and to evaluate the relevance of initial conditions to explain the divergence of mentioned development trajectories. Initial political uncertainty captures the power vacuum during the period between collapse of the socialist systems and before new democratic governments came to power. The essay implements quantitative measurement of the initial political uncertainty (also known as 'initial political disruption') and evaluates its role to moderate the income-democracy relationship among the CEE economies. This study elaborates on Walder et al. (2015) who conceptualized relevance of the initial political disruption as the major development factor for the CEE nations. My findings confirm a tendency that the initial political disruption reverses the 'modernization hypothesis'. The results remained consistent after considering varying data frequencies, alternative measures of democracy, and after controlling for other relevant initial conditions.

4.1 Introduction

In the late 1980s, there occurred a unique 'natural experiment', as a large group of Eastern European countries experienced rapid democratization due to collapse of the communist block. These countries had similar socialist legacy and were planned economies with a single party rule. After communism collapsed in these countries in 1989-1990, and they underwent initial democratization, the further development paths of these nations diverged. Analysis of the income-democracy relation of Central and Eastern European economies can

provide valuable lessons for other nations that experience regime changes. Walder et al. (2015) conceptualized that initial political disruption was an overlooked critical factor for the subsequent development of CEE economies. The concept of 'initial political disruption' relates to the lag between collapse of socialist and communist governments and emergence of new democratic ones. During this time, disruptive processes occurred due to the vacuum of power and lack of legal institutions. The goal of this study is to explore the relation between economic development and democracy in the post-socialist European countries (Central and Eastern Europe - CEE), and investigate the capacity of initial political uncertainty in shaping the nature of the income-democracy relationship.

The idea of positive relation between income and democracy, known as 'modernization hypothesis' was suggested by Lipset (1959) and Lipset (1960) and it has high political-economic significance. Wucherpfennig and Deutsch (2009) underlines a need for broader understanding of the modernization hypothesis, whereas socioeconomic development is associated with an increased middle class that is subsequently able to achieve democratic transition and greater democratic stability. Since its formulation, the modernization hypothesis was thoroughly tested by numerous researchers, including recent tests by Acemoglu et al. (2008), Acemoglu et al. (2009), Cervellati et al. (2014), as well as Benhabib et al. (2011) and Benhabib et al. (2013). Early studies of the income-democracy relation generally evidenced positive linkage (e.g. Barro 1996). After Acemoglu et al. (2008) for the first time applied fixed effects panel methodology for testing the modernization hypothesis, no evidence of it was found, and this cast doubt on the long believed mechanism. Further insight was contributed by Cervellati et al. (2014), who suggested that the income-democracy relation can behave differently in different country groups. Specifically, the researchers showed negative income-democracy association for the countries with colonial origins, and positive effect for the non-colonial countries.

There is a lack of studies that explored modernization hypothesis in case of the Central and Eastern European nations using relevant empirical identification strategy. The current investigation extends the mainstream literature along the two main dimensions. First, this study evaluates the relation between income and democracy in the case of the Central and Eastern European economies, applying an empirical framework that closely relates to Ace-

moglu et al. (2008) and Cervellati et al. (2014). Second, this study quantitatively evaluates the concept of the initial political uncertainty (referred to as 'initial political disruption' in Walder et al. 2015) as the key factor to shape subsequent relation between income and democracy in these nations. Contributions of this study towards the literature is therefore in undertaking quantitative evaluation of the initial political uncertainty and including it as a major moderation factor in the income-democracy equation.

The considered mechanism that underlies the empirical analysis in this study is rooted in the studies by Acemoglu (2008), Hellman (1998), Passarelli and Tabellini (2017) and Walder et al. (2015). The extended period of political uncertainty, which took place in such countries as Russia, Ukraine, or Kazakhstan, empowered the abuse of power by former political leaders (alongside criminal groups), some of whom managed to transform into oligarchs by capturing control over natural resources, infrastructure, supply chains, and other key economic resources. A major way this process occurred was via non-transparent privatization of the most viable and largest state enterprises at prices that were significantly below their fair values Walder et al. 2015. As the outcome, after oligarchs gained economic power, they became capable of establishing high entry barriers, diverting new players from entering the markets, and postponing any further democratization by leveraging their excessive lobbying capacity (Acemoglu 2008). While economic growth still occurred as large enterprises managed to expand, in part due to their preferential treatment by government regulators, democratization process in such economies was considerably slowed down.

The remaining part of the study is organized as follows: Section 2 reviews the initial socioeconomic situation among the CEE economies and its subsequent development. Section 3 introduces the concept of initial political uncertainty and its measurement. Methodology of the study is detailed in Section 4. The main findings are presented and discussed in Section 5, which is followed by their robustness tests exhibited in section 6. Finally, the concluding Section 7 completes this study.

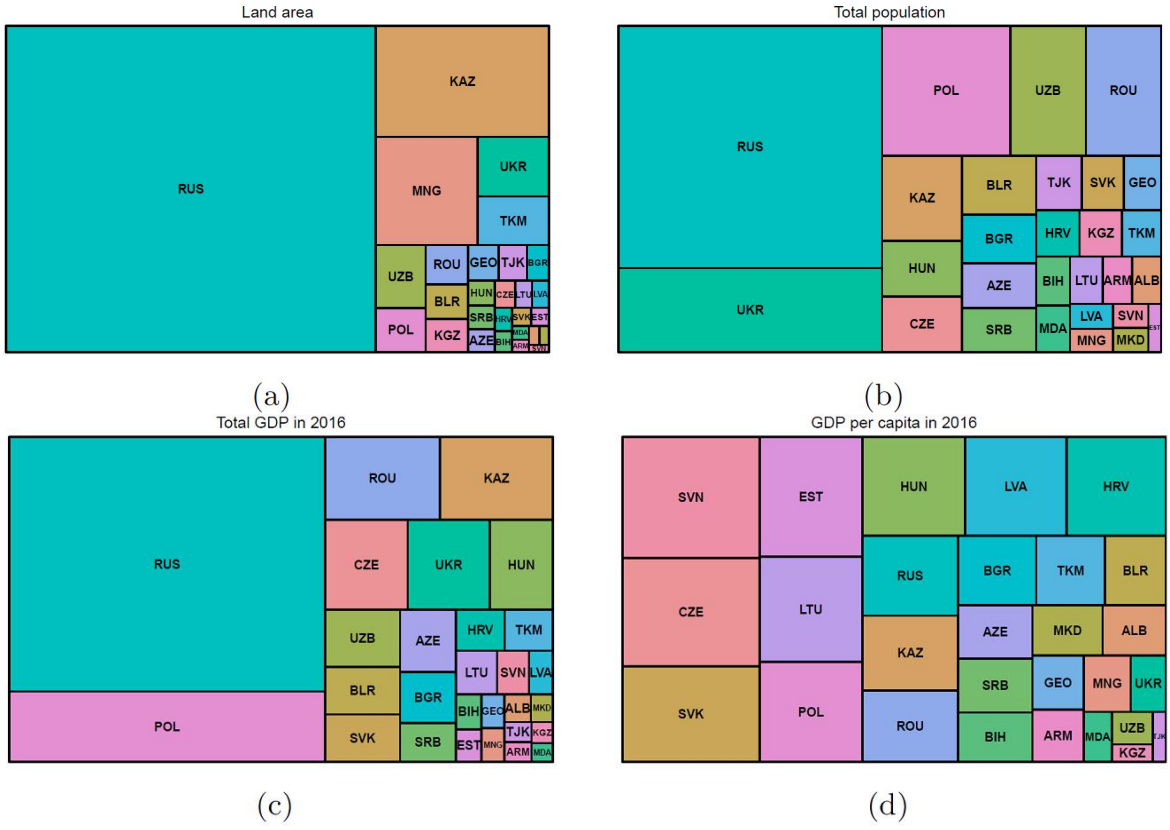
4.2 Democratization and Development in Transition Economies

Before the analysis is presented in the following parts, this section provides an overview of socioeconomic situation in the CEE group of countries. Peculiarity of these countries is their socialist legacy, and the fact that they started from relatively similar institutional conditions. Although the CEE economies started very much alike, their subsequent paths diverged in terms of both, economic development and democratization. The current section explores the initial distribution of these among the CEE countries and their subsequent development dynamics. This section makes a statement that neither initial GDP per capita, nor initial level of democracy could pre-determine the initial political uncertainty of these economies, but that rather the initial political uncertainty was randomly assigned.

4.2.1 Size of considered countries

These countries have joint history as all of them were in the socialist block during the time since the end of World War II till the end of 1980s. As of 2016, Russia was the largest by its territory, total GDP PPP and total population among the CEE countries. But based on the GDP per capita, three Central European economies, Slovenia, Czech Republic, and Slovakia, led the group, whereas Russia ranked tenth.

Figure 4.1: General comparison of the CEE countries



4.2.2 Economic development indicators

In 1990, real GDP per capita was the highest in Czech Republic, followed by Russia, Poland and Kazakhstan. The highest increase was in Bosnia & Herzegovina, Lithuania, Latvia and Poland. In fact, Russia and Kazakhstan saw high initial political disruption despite having some of the highest starting GDP per capita. For three countries, namely Ukraine, Tajikistan, and Kyrgyzstan, real GDP per capita in 2016 was lower than its starting level in 1990.

Figure 4.2: GDP per capita in CEE countries in 1990

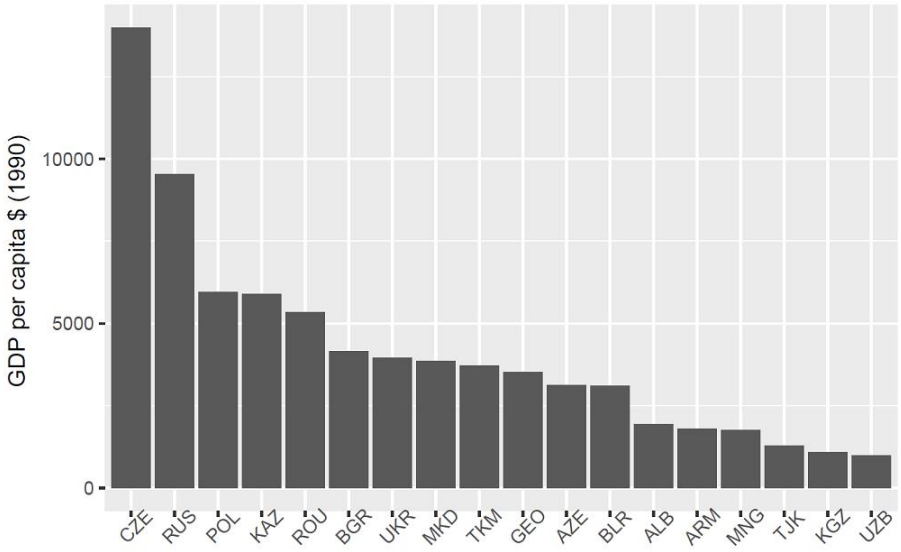
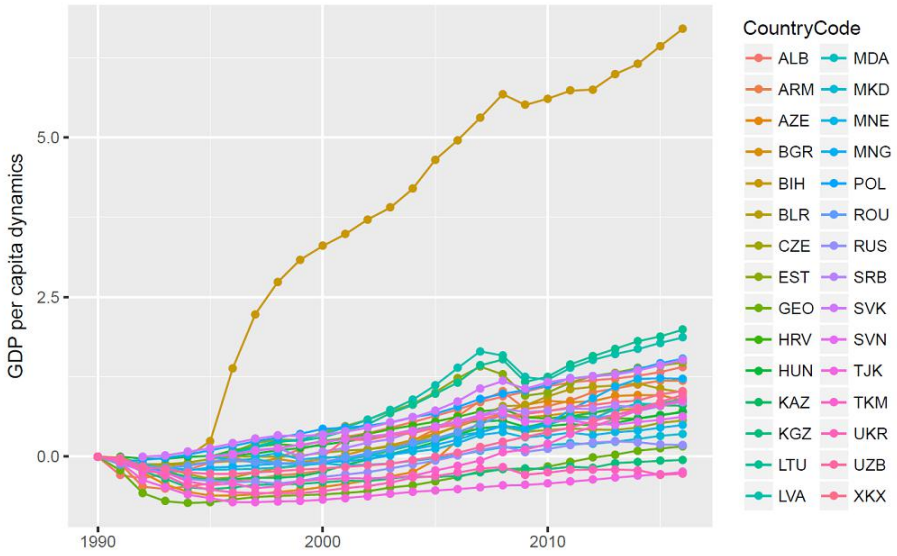
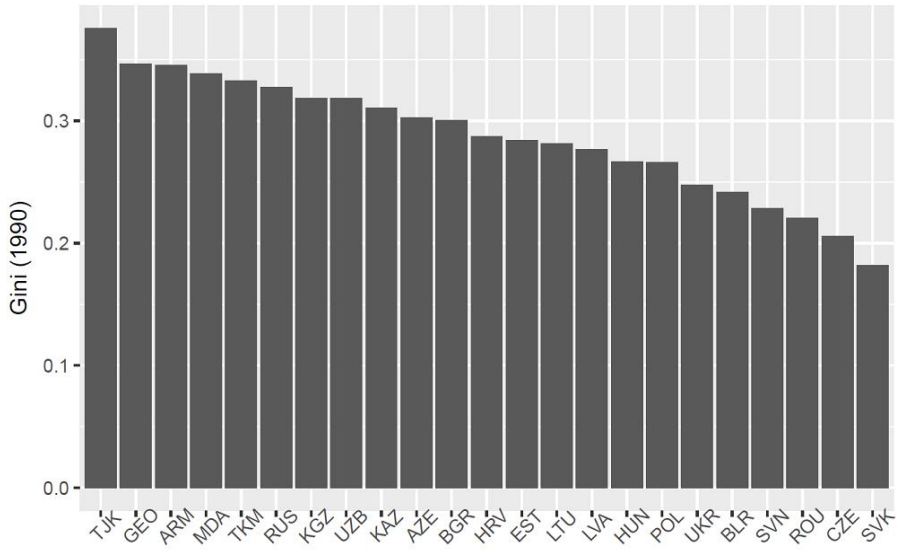


Figure 4.3: GDP per capita changes in CEE countries



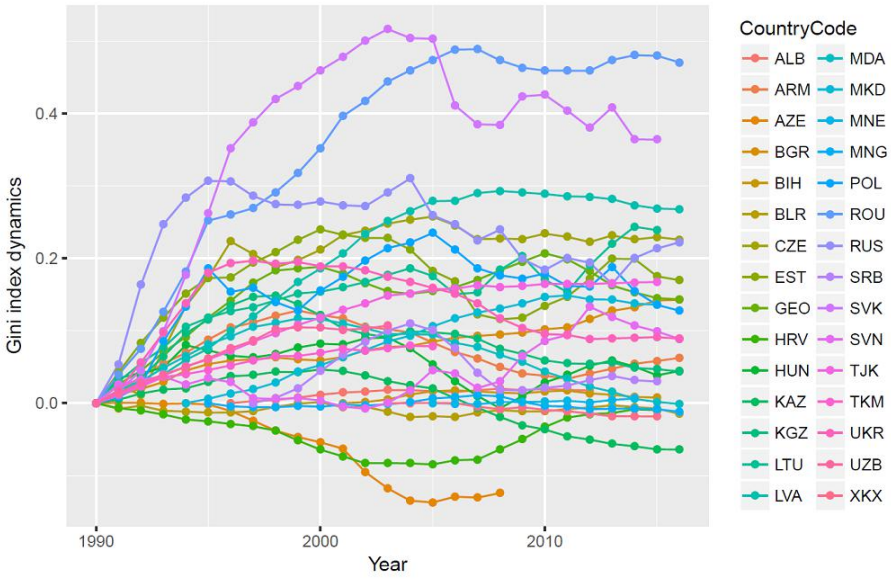
The starting Gini index was the lowest in Slovakia, Czech republic and Romania, while it was the highest in Tajikistan. During 1990-2016, Gini index lowered in just four economies - Moldova, Mongolia, Belarus, and Kazakhstan, indicating increased equality. While for all others it increased showing the growth of inequality among these nations.

Figure 4.4: Gini index in CEE countries in 1990



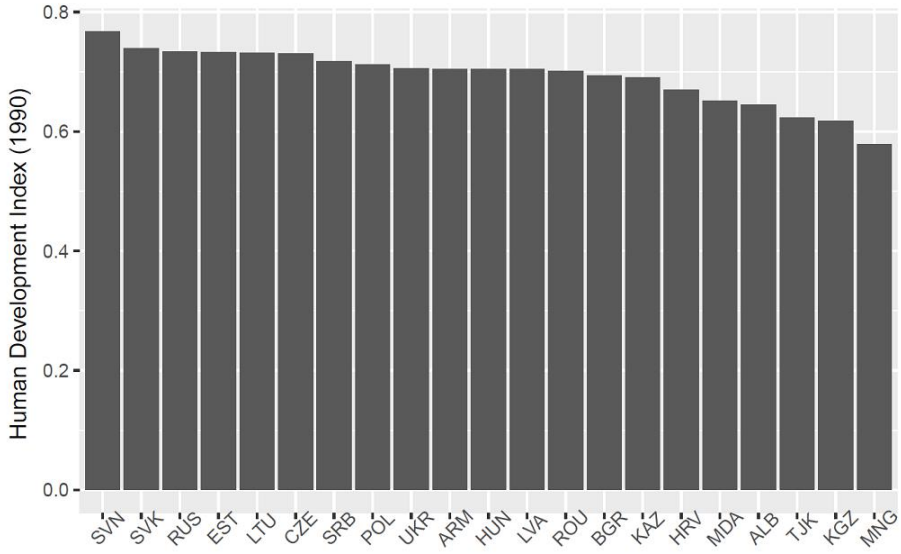
*

Figure 4.5: Gini index changes in CEE countries



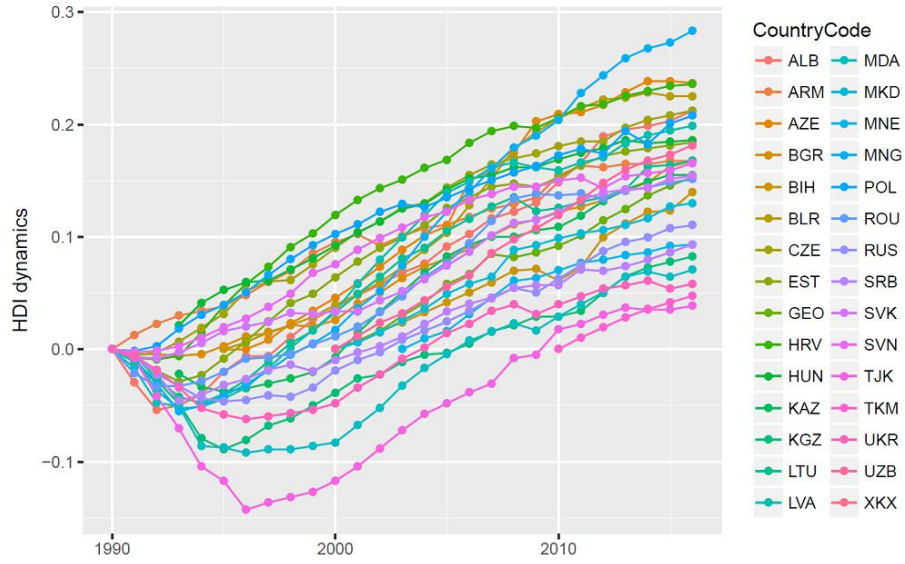
Human development index (HDI) was close among CEE economies. Moreover, over the two decades HDI has grown for all countries. While in some of these nations growth was minimal (Ukraine, Tajikistan, and Turkmenistan), for others it was considerable (Mongolia, Azerbaijan, and Croatia).

Figure 4.6: Human development index in CEE countries in 1990



*

Figure 4.7: Human development index dynamics in CEE countries



4.2.3 Democracy development in CEE nations

Before the collapse of the socialist system, the CEE countries had similar levels of democracy. This is linked to availability of only one ruling party that successfully suppressed any other political forces and deprived people of their political freedoms. Specifically, in 1989 the

level of democracy (based on Polity regime measure) for Czechoslovakia, the former union of Czech Republic and Slovak Republics, for Bulgaria and Mongolia scored zero and it was the same level as for the Peoples Republic of China. Democracy in the Soviet Union (USSR) [1] as well as in Yugoslavia, former union of Serbia, Croatia, Albania, and Kosovo, scored one and it was just a point above the level of China, indicating slightly more democratic regimes. For comparison, democracy level in the US and UK (based on Polity regime measure) scored the maximum of 10 points in 1989. Such figures indicate that initial level of democracy, prior to the collapse of the communist and socialist governments, was similar among all the CEE economies and was not neither able to predetermine the initial political uncertainty in these countries, their their subsequent diverging trajectories.

The main democracy indicator that is used in this study is Political freedom. The indicator is calculated by Freedom House for 195 countries based on 125 lower-level indicators. Two main aggregation parts are blended together, namely political rights and civil liberties (FreedomHouse 2017). The original scaling defines 1 as the most free while 7 is the least free conditions. For the purpose of this analysis, Political freedom variable was re-scaled to range between 0 and 100 with higher values representing freer political conditions.

In 1991 some countries, such as Poland and Czech Republic, already experienced transition towards democratic government, while others were stuck in their period of initial political disruption, as democratic governments in these nations had not formed yet. Thus, the CEE countries already exhibited heterogeneity by their level of democracy in 1991, moreover their subsequent development trajectories diverged further (*Appendix C.6*). The group of countries with relatively high Political freedom included Czech Republic, Poland, Estonia, and Hungary. In the years following the collapse of socialist system, low political freedom was observed in Albania, Georgia, Romania, Serbia, Uzbekistan and Tajikistan. All the countries which initially moved to high levels of democracy managed to subsequently maintain this achieved level. But there was strong divergence among other nations. While Georgia, Albania, Moldova and Serbia managed to considerably improve their political freedom, substantial decline of democracy took place in Russia, Azerbaijan, and Armenia. For

[1] USSR included 15 constituents: Armenia, Azerbaijan, Belarus, Estonia, Georgia, Kazakhstan, Kyrgyzstan, Latvia, Lithuania, Moldova, Russia, Tajikistan, Turkmenistan, Ukraine and Uzbekistan

some countries with low democracy in the beginning of their path there was even further deterioration - Tajikistan, Uzbekistan, and Turkmenistan obtained autocratic rulers who pushed back initial democratization processes.

In the robustness check section of the current study, two other measures of democracy are considered (Polity regime by Polity IV project, and Freedom of the Press by Freedom House). It should be noted that these measures are rather tightly correlated [2]. Dynamics of the democracy based on these measures showed starting distribution and time trends that are similar to that of the Political freedom.

Further technical details about Regime and Freedom of the press are shown in *Appendix C.2*. Also, country-level dynamics of each of the three measures of democracy are shown in Appendices C.6 - C.8.

4.3 Initial Conditions

Evaluation of the income-democracy relation in CEE economies is performed for the total sample, as well as for the sub-samples obtained by splitting the total sample based on the initial political disruption. The concept of initial political disruption, as well as a set of other relevant initial conditions was obtained from political economy and economic development literature (e.g. Barro 1996; De Melo 2001; Walder et al. 2015). Rationale behind the major 'initial conditions' is provided below, while details of their calculation are shown in *Appendix C.1*.

4.3.1 Political disruption

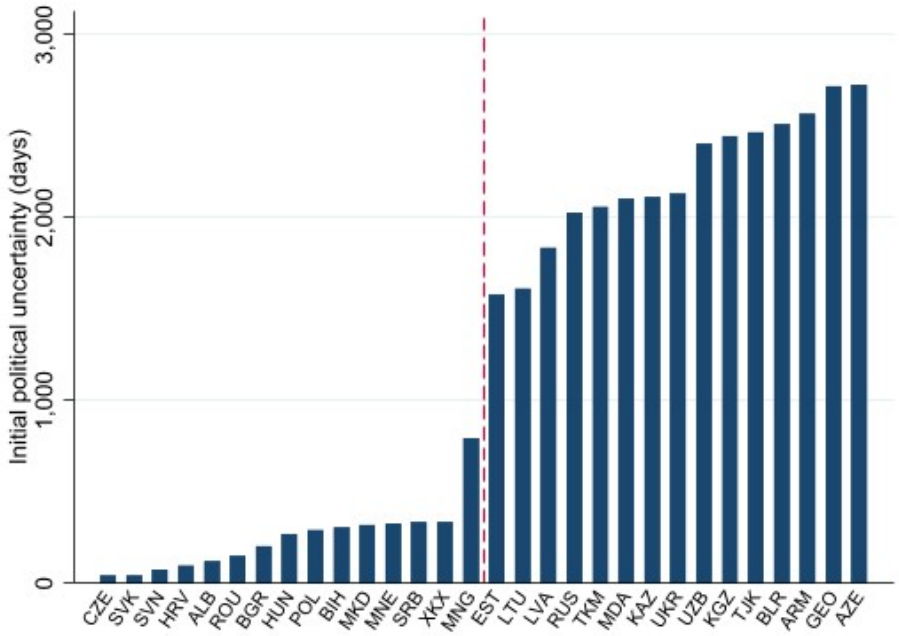
Among the initial conditions, the current study relies on the phenomenon of 'initial political disruption' or 'political origins' of CEE as it was conceptualized by Walder et al. (2015). The 'political origins' of CEE relate to the nature of political transition from social-

[2]Pearson correlation coefficient between Political freedom and Polity Regime is 0.869, between Political freedom and Freedom of the press is 0.943, and that between Polity Regime and Freedom of the press is 0.844. All these correlation coefficients are significant at 1% level.

ism towards democratic institutions. It is suggested that the extent of 'political disruption' significantly affected the nature of diverging trajectories of these countries. The mechanisms (channels) of such effect are explained below. Also, the approach used for construction and measurement of the variable is detailed in *Appendix C.3*.

The idea of 'political disruption' in CEE countries and discussion of its high relevance is due to Walder et al. (2015). The researcher compared group of the post-communist European economies that were not part of the Soviet Union versus the post-communist economies that previously constituted the Soviet Union. Before the end of 1980s, communist party was the sole owner of either all or vast majority of productive assets in each of these economies. At some moment in time of late 1980s it became apparent that the communist party was not able to control productive assets anymore. In several countries, including Poland, Czech Republic, and Mongolia among others, new democratic parliament and government were established rapidly, taking over the communist party. But in other post-communist economies (e.g. Russia, Ukraine, Kazakhstan among others) there was prolonged interim period accompanied by the uncertainty. In these countries it took years after the end of the communism rule and before a new multi-party parliament was freely elected and appointed democratic government, which was then able to foster development of valid democratic institutions. This intermittent period of uncertainty was associated with the lack of property rights protection, weakness of legal institutions as the communist party lacked its previous power while new democratic institutions had not emerged year. The period is referred by Walder et al. (2015) as the period of 'political disruption'.

Figure 4.8: Initial political uncertainty in CEE countries



During the period of uncertainty (or 'political disruption') a number of disruptive processes took place and caused capture of control over productive assets by either former communist leaders, criminals, or newly established dictators. It is thus suggested that the extent to which the control over the economy was captured by a limited groups of 'new winners' further prevented the democratization in these countries, irrespective of whether these countries faced economic growth. The idea of 'new winners' is detailed in Hellman (1998) indicating that on the initial stages of reforms, certain interest groups become winners, but then they can turn into great obstacles of any further reforms. In many post-USSR countries the period of 'political disruption' was associated with rise of oligarchs, killings of many businessmen, unfair privatization of state-owned assets by businesses associated with the ruling politicians. Walder et al. (2015) indicates that during this period there were significant recessions as post-USSR countries experienced hyperinflation and lost about 50-80% of the size of their economy.

4.3.2 Other initial conditions

Therefore, as was previously mentioned, Political disruption was the period of property rights uncertainty. According to Walder et al. (2015), this had immense effect on how countries developed, including their democratization paths. But this was not the only initial condition, and there were other ones. In the robustness check section, the analysis considers relevance of other initial conditions: resource endowment, initial income level, market memory, state independence, and regional tensions[3]. These initial conditions were selected based De Melo (2001), BenYishay and Grosjean (2014), as well as Barro (1996).

'Resource endowment' suggests that higher dependence on resources prevents modernization. The reason is that for resources-rich economies where strong rent-seeking incentive among incumbent elites. The effect was measured as the share of employed in mining and natural resources industry in 1989, available from BenYishay and Grosjean (2014) [4].

Low initial income or threshold GDP level - countries with initially low level of GDP are more likely to revert modernization. The 'starting' real GDP per capita was obtained for all countries as of 1990 from Maddison (2010). The variable was used in Barro (1996) to argue that countries below certain income threshold cannot sustain initial democratization.

'Market memory' indicates the time that an economy was under socialist planning. Following De Melo (2001), longer time under social planning is associated with lower market memory. 'Geographic location' is measured as an indicator variable indicating whether a country borders on a democratic non-communist state. De Melo (2001) uses this variable to evaluate access to Western markets.

State independence and institutions aim to measure the maturity of country institutions. The variable contains three categories - newly established states, countries of federal states, and countries that were independent prior to 1989. Thus, newly created states have to develop more institutions, as compared to countries that were independent over longer time.

Finally 'regional tensions' capture the fact whether a given country was involved in military conflict or now (De Melo 2001). Countries that were in military conflict after the

[3]Technical details of these variables are summarized in *Appendix C.1*

[4]The variable is available in BenYishay and Grosjean (2014) for 26 economies, but it is missing for Kosovo, Macedonia, Mongolia, and Turkmenistan

collapse of communist and socialist parties, incurred disruptive consequences of such conflicts for both, their economic and democratic situation.

4.4 Methodology

4.4.1 Panel data analysis

General set-up of regression analysis follows Cervellati et al. (2014) and Acemoglu et al. (2008). Specifically, Cervellati et al. (2014) used the approach to consider sub-samples based on whether a country was a colony, and they used interaction of GDP and the colony indicator variable for fixed effect estimation. In the current study, the relation between democracy and income per capita is evaluated based on equations (11) and (12) below:

$$d_{i,t} = \beta_0 + \alpha d_{i,t-1} + \gamma y_{i,t-1} + \delta_i + \mu_t + u_{i,t} \tag{11}$$

$$d_{i,t} = \beta_0 + \alpha d_{i,t-1} + \phi_1 (y_{i,t-1} \cdot c_i) + \phi_2 c_i + \gamma y_{i,t-1} + \delta_i + \mu_t + u_{i,t} \tag{12}$$

Where, d - is a measure of democracy, y - log of real GDP per capita, δ and μ are country and year fixed effects. Equation (12) includes c, which is time-invariant dummy that indicates the initial political uncertainty ($Uncer$ or IPU). My analysis aims to investigate the following hypotheses. First, the unconditional modernization hypothesis: H1: $\gamma > 0$. Second, the conditional modernization hypothesis: H2: $\phi_1 < 0$. That is the considered effect of the initial political uncertainty operates to decrease (or to invert) the positive income-democracy relation. In addition to the above, direct effect of the initial political uncertainty on democratization is tested: H3: $\phi_2 < 0$. The above equations represent baseline estimation method, while additional estimation approaches are detailed below.

4.4.2 Two-stage least squares (TSLS)

Estimation of the effect of income on democracy involves a significant endogeneity problem. In fact, while modernization hypothesis argues that income causes democracy (e.g. Acemoglu et al. 2008; Cervellati et al. 2014), there is certain evidence of the reverse casual link indicating democracy having an impact on economic growth (Acemoglu et al. 2019; Benhabib et al. 2011; Rodrik and Wacziarg 2005. Therefore, under endogeneity, estimated coefficients in the baseline model could be biased. To deal with this potential setback, two-stage least square methodology was applied. Lagged trade-weighted world income was used as the instrument for lagged GDP per capita. The approach follows Acemoglu et al. (2008). The F-statistic of robust regression of GDP per capita on the instrument has F-statistic of 27.14, which is above the conventional threshold of 10, indicating solid performance of this instrumental variable. The instrument was constructed as follows. For a given country i from among the CEE economies, for each country 'j' with which it has trade relations, weights w_{ij} are determined based on significance of the trade between i and j relatively to total international trade of country i. Such weights were constructed based on aggregated 4-year periods during the sample period (weights are hold constant within the 4-year periods). Then the trade-weighted income is determined as follows:

$$\hat{Y}_{i,t-1} = \Sigma_{j=1,j\neq i}^{N} \omega_{ijt} Y_{j,t-1}, \quad j \in 1, ..., N \tag{13}$$

4.4.3 System GMM model

Another approach aimed to deal with the fact that inclusion of a lagged dependent variable can incur a bias into the estimated coefficients. The system GMM method obtains first differences of the baseline model, so that the country fixed effects are eliminated from the equation (4).

$$\Delta d_{i,t} = \alpha \Delta d_{i,t-1} + \gamma \Delta y_{i,t-1} + \Delta \mathbf{X}'_{\mathbf{i,t-1}} \beta + \Delta \mu_t + \Delta u_{i,t} \tag{14}$$

Also, the Arellano and Bond (1991) estimator requires a number of moments conditions, specifically no autocorrelation in lags of differenced dependent variable, no serial correlation

in the residual term, as well as no correlation between lagged dependent variable and the residual term.

4.4.4 Robustness tests

For robustness check, a number of additional procedures were implemented. First, alternative data frequencies were considered, using three-year and five-year time intervals. Second, other measures of democracy were evaluated in addition to the Freedom House indicator of Political freedom'. These alternative measures were Polity International measure of political regime, and Freedom House measure of the freedom of the press. Third, a number of other initial conditions variables are considered, the baseline results are shown to remain after controlling for all these initial conditions. Besides, performance of other initial conditions is compared to that of the initial political uncertainty.

4.4.5 Data

The main dataset contains annual frequency data for all 30 post-socialist countries of Central and Eastern Europe (*Appendix C.4*). The sample includes data for 16 years, ranging from 1991 up to 2016. Relevant data was obtained from multiple sources, including the World Development Indicators, Freedom House, Polity International, and UN Comtrade. Relevant details for each variable as well as data sources for each variable are stated in the *Appendix C.1.*

4.5 Results and Analysis

This section evaluates the relation between development and democracy for the Central and Eastern European economies, and explores the moderating role of the initial political uncertainty. The analysis firstly compares mean values of relevant economic and democracy indicators by the sub-set of countries with high and low initial political uncertainty, then regression results are provided.

4.5.1 Univariate analysis

The descriptive statistics for the overall sample, as well as for the two sub-sets of countries by initial political disruption, are shown in table 4.1. For each included variable the difference of means between the two subsets was calculated and tested for its statistical significance.

The comparison between the countries from the subsets organized by level of their initial political uncertainty indicates substantial differences between the two groups. There is lower level of democracy among the countries with high initial political uncertainty, as well as lower GDP per capita, lower HDI and higher inequality based on Gini index. This is despite the fact that in this group of countries, average initial GDP per capita was significantly higher than in the subgroup of countries that did not experience high initial political uncertainty. The results suggest negative impact of the initial political uncertainty on both, income and democracy.

Table 4.1: Descriptive statistics by the level of initial political uncertainty

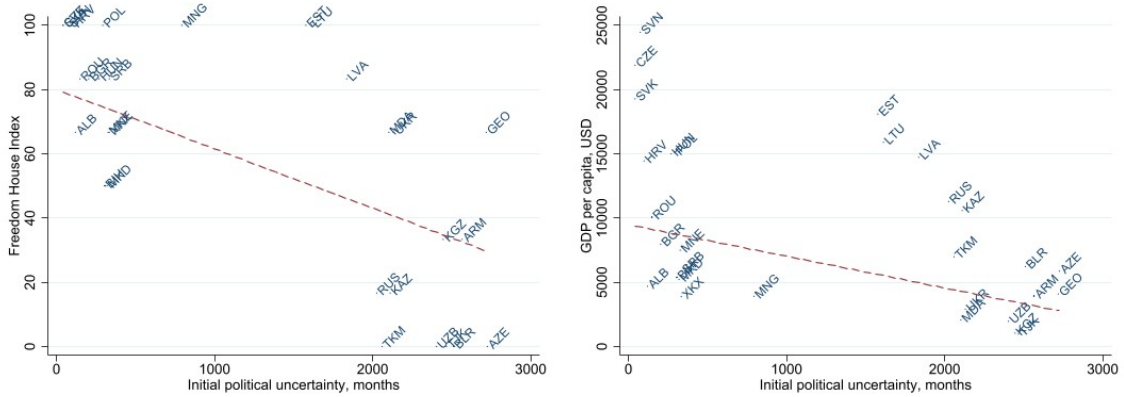
	Total sample			Subsets by level of uncertainty		
	Obs	Mean	Std. Dev.	Low	High	Difference
Uncert	810	1233.530	1026.250	247.260	2219.800	-1972.530 ***
Fhouse	810	57.202	34.121	72.757	41.646	31.111 ***
Polity	810	67.609	33.988	81.715	53.502	28.213 ***
Freedom of Press	690	57.177	27.931	71.880	43.698	28.182 ***
GDP	751	6357.590	5390.520	8213.840	4592.950	3620.880 ***
GDP 90	810	5966.030	2386.360	5474.800	6457.260	-982.460 ***
HDI	698	0.732	0.076	0.753	0.711	0.041 ***
Gini	703	0.322	0.053	0.302	0.339	-0.037 ***

Sig: *** p<0.01, ** p<0.05, * p<0.1.

The initial findings from descriptive statistics are further explored using scatter plots, evaluating the linkages between economic development, democracy and initial political uncertainty. The results in figure 9 point towards negative relationship between the initial political uncertainty and democracy, as well as between the initial political uncertainty and the level of income per capita. High level of democracy and low level of initial political uncertainty was observed in Poland, Slovakia, Czech Republic, as well as Romania and Bulgaria. On the other hand, high initial political uncertainty and low level of democracy occurred in Azerbaijan, Belarus, Uzbekistan, Turkmenistan, Russia and Kazakhstan (Figure 9A). The suggested relation between initial political uncertainty and income is also evident

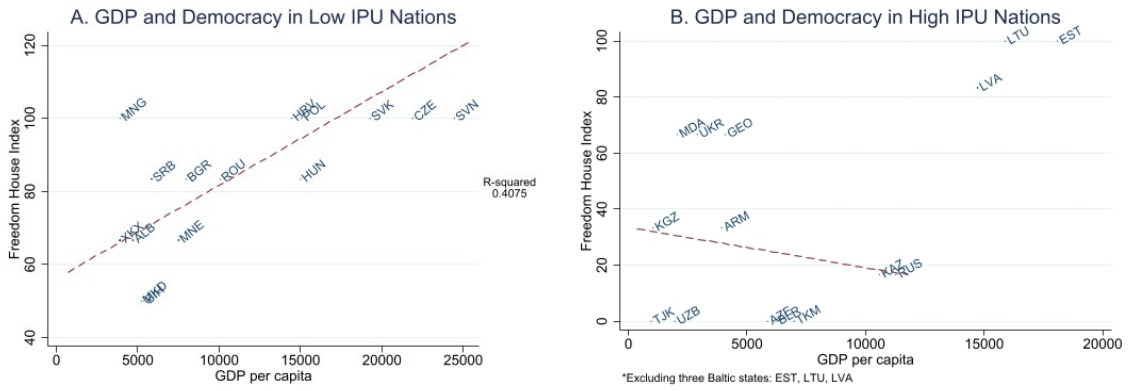
(Figure 9B). Countries that have high income and low initial uncertainty include Slovenia, Czech Republic and Slovakia, while nations with the highest initial uncertainty and lowest income include Kyrgyzstan, Tajikistan, Uzbekistan, as well as Georgia, Armenia, Ukraine, and Moldova.

Figure 4.9: Relevance of initial political uncertainty for income and democracy



The relation between economic development and level of democracy seems to be positive among the countries with low initial political disruption. But the relationship is negative for the economies with high IPU (figure 4.10). In the second group the three Baltic states are considered outliers as they have both high level of democracy and high GDP per capita, despite experiencing significant initial political uncertainty. These three Baltic countries are very small economies. They are in close proximity to European Union, have market memory and also lack natural resources. These additional factors could be responsible for these nations not following the oligarchic development path.

Figure 4.10: Economic development, democracy and initial political uncertainty



4.5.2 Main results

Analysis of the relation between income and democracy and of the moderation role of the initial political disruption starts with the fixed effects regressions in table 4.2. Such approach is close to that in Cervellati et al. (2014) who explored moderation role of colonial origins.

Positive relation between economic development and democracy is observed in column 1, where no fixed effects are included, neither standard errors are clustered at the country level. This represents conventional positive correlation between income and democracy previously evidenced by numerous studies.

Furthermore, after controlling for the past democracy history, as well as for the time and country fixed effects, the effect of economic development is not statistically significant. Therefore, the results in columns (1) and (2) follow the approach by Acemoglu et al. (2008). As was the case with Acemoglu et al. (2008), the findings in columns (1) evidence positive relation between income and democracy, while after taking fixed effects into account the relation becomes insignificant. The results starting from column (3) take 'Uncertainty' into the account. In column (3) the interaction term between income and uncertainty is significant and has negative sign. This suggests that the initial political uncertainty can decrease the income-democracy relation and possibly make it negative. Based on the results in column (3), it can be argued that a CEE country with *Uncertainty* period of 270 days (1st quartile) has overall GDP slope coefficient of 2.16, while a country with *Uncertainty* period of 2,131 days (3rd quartile) has overall GDP slope of -1.57. In case of the latter, a 10% increase in GDP is associated with a subsequent average decline in democracy by 0.16 percentage points.

Table 4.2: Relation between development and democracy using fixed effects

	A. Total sample			B. Level of IPU	
				Low IPU	High IPU
	(1)	(2)	(3)	(4)	(5)
$Fhouse_{t-1}$	0.972***	0.735***	0.710***	0.701***	0.699***
	(94.12)	(15.42)	(14.23)	(8.53)	(11.00)
GDP_{t-1}	0.628*	-2.459	2.696	1.710	-4.748
	(1.70)	(-1.00)	(0.98)	(0.55)	(-1.38)
$GDP_{t-1} \cdot Uncert$			-0.002***		
			(-3.25)		
$Uncert$			-0.021**		
			(-2.29)		
Time FE	no	yes	yes	yes	yes
Country FE	no	yes	yes	yes	yes
Clustered SE	ne	yes	yes	yes	yes
Num. obs	721	721	721	351	370
Num. countries	30	30	30	15	15
Adj R^2	0.952	0.964	0.964	0.918	0.968

Dependent variable is Freedom House measure of democracy normalized to 0-100 scale. *GDP* is natural log of real GDP per capita. *Uncer* measures initial political uncertainty. High IPU and Low IPU subsets are separated by the median. Robust t-statistics in parentheses. Sig. *- 10%, ** - 5%, *** - 1%.

Furthermore in panel B, the same model as in column (2) is estimated for the subsets obtained by splitting the sample based on initial political uncertainty (IPU). The median level of the uncertainty was used as the benchmark to split the two subsets. Generally, the results in columns (4) and (5) are in line with the findings in column (3), as positive slope coefficient on GDP is evidenced for the countries with low IPU, while negative one is observed for the subset with high uncertainty. Yet their statistical significance is marginal, which could be due to the relatively small sample sizes.

4.5.3 TSLS and dynamic panel results

Further analysis of the income-democracy relationship for CEE countries aims to overcome possible limitations of the main results, which were prone to endogeneity issues - the possibility of the two-way causal link between income and democracy. For this purpose the dynamic GMM panel estimator by Arellano and Bond is used in panel A, as well as the two-stage least squares method - in panel B. Findings from these two methods are similar to each other and to the main results. The results in columns (1) and (4) should be compared

to those in column (3) of the main results from table 4.2. While findings in columns (2) and (3) (as well as those in columns 5 and 6) should be compared to those in columns (4) and (5) of the main results.

Dynamic GMM results in column (1) evidence that the initial political uncertainty weakens the income-democracy positive association. The GMM results for the two subsets indicate positive and negative association between income and democracy in countries with, respectively, low and high IPU levels. Similar findings are shown in panel B that includes the second stage results [5]. The TSLS results for the two subsets are more statistically significant than the equivalent coefficients from GMM method, but still lack statistical power to reject the null hypothesis.

Table 4.3: Relation between development and democracy: Dynamic panel and 2SLS

	A. Arellano-Bond estimator			B. TSLS estimator		
	All	Low IPU	High IPU	All	Low IPU	High IPU
	(1)	(2)	(3)	(4)	(5)	(6)
$Fhouse_{t-1}$	0.604***	0.710***	0.592***	0.901***	0.772***	0.945***
	(8.20)	(8.05)	(11.26)	(8.70)	(13.74)	(12.26)
GDP_{t-1}	10.268	2.903	-4.873	1.154	1.578	-6.761
	(1.54)	(0.65)	(-0.93)	(0.76)	(1.28)	(-1.45)
$GDP_{t-1} \cdot Uncert$	-0.004***			-0.003		
	(-2.58)			(-0.95)		
$Uncert$				-0.000	-0.003***	0.000
				(-0.26)	(-7.97)	(0.07)
Time FE	yes	yes	yes	yes	yes	yes
Country FE	yes	yes	yes	yes	yes	yes
Num. obs	691	336	355	669	196	105
Num. countries	30	15	15	30	30	30
Wald χ^2	48982 ***	228.10 ***	131.57 ***			
Adj R^2				0.934	0.837	0.929

Dependent variable is Freedom House measure of democracy normalized to 0-100 scale. *GDP* is natural log of real GDP per capita. *Uncer* measures initial political uncertainty. High IPU and Low IPU subsets are separated by the median. Robust t-stats in parentheses. Sig. *- 10%, ** - 5%, *** - 1%. First stage F-stat from bivariate regression: (4)- F=10.6***, (5)- F=15.05***, (6)- F=2.43.

These findings can be argued to provide additional evidence towards supporting the statement that GDP and democracy have positive association in countries with low initial uncertainty, while in the countries that experienced prolonged uncertainty period this relation can be inverted.

[5] Trade-weighted world income was used as the instrumental variable for the GDP per capita. The IV construction follows Acemoglu et al. (2008), p.824

4.6 Robustness Tests Results

Goal of the current section is to perform robustness checks for the negative effect of the initial political uncertainty on the income-democracy relation. The first test (Table 4.4, panel A) uses alternative data frequencies (annual frequency, and 3-year and 5-year aggregated periods). The second test employs alternative measures of democracy - regime measure and freedom of the press (Table 4.4, panel B). The third test considers for a broader set of relevant initial conditions and compares their effect to that of the initial political uncertainty (Table 4.5).

4.6.1 Alternative data frequency and democracy measures

The results by measure of democracy and alternative frequency are indicative of the same negative moderation effect of Uncertainty on the income-democracy relation. With different measures of democracy, the cross-term between Uncertainty and lagged GDP is negative in all specifications. Yet limitation of this test is that the cross-term has low statistical significance for two auxiliary democracy measures. Additionally, there is negative and significant direct effect of Uncertainty on each of the three measures of democracy. In panel B, the interaction of uncertainty and income is negative and statistically significant across all three time frequencies. Moreover, magnitude of the cross-term increases with the length of time period, so that the slope coefficient of the cross-term under 5-year frequency is approximately three time higher as compared to that under 1-year frequency. Therefore, the results from table 4.4 provide support towards the main results about negative moderation effect of initial political uncertainty.

Table 4.4: Relation between development and democracy under alternative settings

	A. Measures of democracy (1-year frequency)			B. Time frequencies (Political freedom)		
	Political Freedom	Regime	Freedom of Press	1-year	3-year	5-year
	(1)	(2)	(3)	(4)	(5)	(6)
$Democracy_{t-1}$	0.710***	0.713***	0.810***	0.710***	0.446***	0.254*
	(14.23)	(12.72)	(32.36)	(14.23)	(4.57)	(1.83)
GDP_{t-1}	2.696	-2.736	5.004*	2.696	5.574	13.417***
	(0.98)	(-0.71)	(1.97)	(0.98)	(1.26)	(3.11)
$GDP_{t-1} \cdot Uncert$	-0.002***	-0.000	-0.001	-0.002***	-0.005***	-0.006**
	(-3.25)	(-0.23)	(-1.66)	(-3.25)	(-2.77)	(-2.40)
$Uncert$	-0.021**	-0.001	0.021***	-0.021**	-0.034**	-0.052**
	(-2.29)	(-0.09)	(3.29)	(-2.29)	(-2.10)	(-2.44)
Time FE	yes	yes	yes	yes	yes	yes
Country FE	yes	yes	yes	yes	yes	yes
Num. obs.	721	721	645	721	227	118
Num. countries	30	30	30	30	30	30
Adj R^2	0.964	0.963	0.983	0.964	0.958	0.968

Dependent variables are indicated measures of democracy normalized to 0-100 scale. GDP is natural log of real GDP per capita. $Uncer$ measures initial political uncertainty. High IPU and Low IPU subsets are separated by the median. Robust t-stats in parentheses. Sig. *- 10%, ** - 5%, *** - 1%.

4.6.2 Relevance of other initial conditions

This section considers other relevant initial conditions, such as resource endowment, starting per capita income, market memory, and compares their performance to that of the initial political uncertainty. The results in columns (1) correspond to the main results that are augmented by including controls for all five other initial conditions. Then, columns (2) through (6) separately consider the direct and moderating effects for each of the initial conditions.

Table 4.5: Relevance of other initial conditions for democracy using fixed effects

	Measures of initial conditions					
	Uncer	Resource	Income90	Marmem	State	Rten
	(1)	(2)	(3)	(4)	(5)	(6)
$Fhouse_{t-1}$	0.726***	0.737***	0.728***	0.722***	0.731***	0.736***
	(14.78)	(15.35)	(14.66)	(15.12)	(14.95)	(15.15)
GDP_{t-1}	2.920	-3.352	19.910	4.878	-2.632	-1.733
	(1.14)	(-1.12)	(1.39)	(1.11)	(-1.05)	(-0.63)
$GDP_{t-1} \cdot IC$	-0.002***	0.988	-2.662	-0.110*	1.111	-1.348
	(-2.87)	(1.03)	(-1.60)	(-1.81)	(1.13)	(-0.69)
IC	-0.041***	-15.249*	9.845	8.826***	-5.022	3.302
	(-4.88)	(-1.86)	(0.79)	(4.77)	(-0.65)	(0.22)
Time FE	yes	yes	yes	yes	yes	yes
Country FE	yes	yes	yes	yes	yes	yes
Num. obs.	721	721	721	721	721	721
Num. countries	30	30	30	30	30	30
Adj R^2	0.964	0.964	0.964	0.964	0.964	0.964

Model in column (1) controls for the five other initial conditions: Resource, Income90, Marmem, State, and Rten. Freedom House 'Political Freedom' is the dependent variable. IC indicates a measure of initial conditions. Considered ICs: $Uncer$ - initial political uncertainty, $Resource$ - resource endowment, $Income90$ - initial income, $Marmem$ - market memory, $State$ - state independence, $Rten$ - regional tensions. Robust t-statistics in parentheses. Sig. *- 10%, ** - 5%, *** - 1%. GDP is natural log of real GDP per capita.

Among the considered six initial conditions, significant moderation effect is evidenced for just market memory ($Marmem$), in addition to $Uncer$. Also, significant direct effect, in addition to $Uncer$ is evidenced for the following two - resource endowment ($Resource$), and market memory ($Marmem$). The effect of resource endowment is in line with the expectation that higher resource endowment facilitates greater rent-seeking with its adverse impact on democracy (BenYishay and Grosjean 2014). The direct effect on democracy turns out to be positive for market memory, suggesting that countries with greater free market legacy, ceteris paribus, have greater chances to experience improvement in democracy. No effect is observed for the starting income per capita ($Income90$), state institutions ($State$) or regional tensions ($Rten$).

In general, initial political uncertainty was confirmed to be one of the key few underlying factors that significantly moderate the relation between income and democracy, as well as pre-determined trends in democracy among the CEE countries.

4.7 Conclusion

The analysis evaluates the income-democracy nexus for post-socialist countries of the Central and Eastern Europe. Special attention is provided towards the role of initial political uncertainty as the major moderating factor that influenced the nature of the relation between income and democracy of these nations.

The initial political uncertainty is a proxy for the period of economic and political anarchy that occurred in the interval between collapse of the socialist party and enabling of the new democratic governments. A prolonged period of initial political uncertainty is associated with extensive rent-seeking, office abuse by top-ranking authorities, lack of property rights protection, and weakness of other crucial governmental institutions.

The obtained results evidence a negative moderation effect of the initial political uncertainty on the association between income and democracy. It can be claimed that a positive income-democracy relation (known as the 'modernization hypothesis') holds in the CEE countries that had short-lived initial political uncertainty. These CEE economies, which experienced prolonged periods of the initial political uncertainty, exhibit negative relationship between their income and democracy. This latter effect can be characterized as the 'inverted modernization hypothesis'. The moderating effect of initial political uncertainty, despite being statistically significant also has reasonable economic significance. In countries with low initial political uncertainty (of less than 1200 days), a 10% increase in GDP is associated with subsequent increase in democracy by 0.17 percentage points. Meanwhile, a similar increase in GDP for those with high initial political uncertainty (over 1200 days) is associated with subsequent decline in democracy of 0.48 percentage points. The results are shown to be consistent when performing robustness tests using different data frequencies and different measures of democracy.

5.0 Conclusions

My dissertation aimed to address several significant phenomena in the areas of financial and applied economics. Specifically, I explored common ownership concentration, and its impact on corporate payouts and investments, wealth implications of marital instability among the elderly, and the role of political disruption for the subsequent relation between income and democracy of countries in transition. The study builds upon prior literature in the areas of financial economics, health economics, and political economy, and it contributes towards these literature areas in numerous ways. Each of the three essays illustrated the ways that such econometric methods as difference-in-differences, propensity score matching, event study, two-stage least squares can be applied in order to provide causal identification for the considered problems.

Three main conclusions of this thesis are as follows: First, common ownership concentration results in lower competition among firms, and thus leads to higher payouts, more intense industry consolidation and lower capital investments. These effects are lower in magnitude for industries that are exposed to import competition from China. Second, 'gray divorce' has negative implications for financial wealth and home equity of divorcing individuals. Gray divorce increases the probability of losing positive wealth by over 5%. The negative wealth implications of gray divorce increase with age. Also, gray divorces can be predicted with reasonable precision using machine learning modeling, which can be of relevance for social policy-makers. Third, the case of post-socialist European countries evidenced that a long-run development trajectory of a country in transition is highly dependent on its initial political disruption (i.e. the interim period between the collapse of the old regime and emergence of a new government). Countries that experienced high initial political disruption saw negative relation between their income and democracy, while nations with low initial uncertainty exhibited positive income-democracy relation, known as the 'modernization hypothesis'.

6.0 Appendix

6.1 A: Common Ownership Concentration Auxiliary Materials

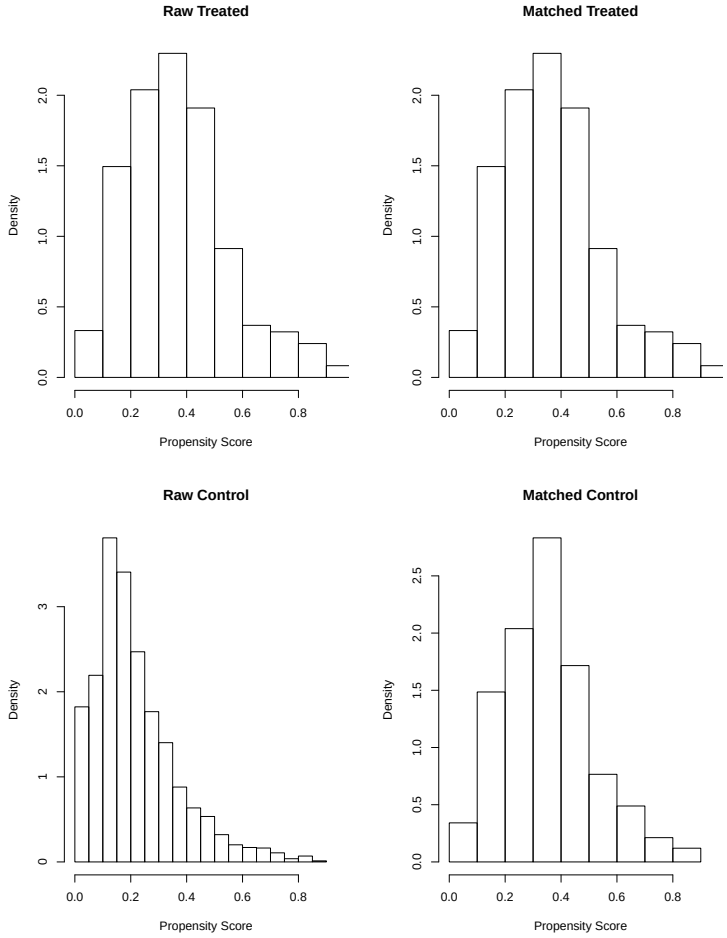
6.1.1 Pairwise correlation

Table 6.1: Pairwise correlation

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1) DCO	1.000													
(2) PCF	0.366***	1.000												
(3) PCS	0.429***	-0.208***	1.000											
(4) Payouts	0.030***	0.011***	0.009***	1.000										
(5) Buybacks	0.031***	0.001	0.019***	0.736***	1.000									
(6) Buybacks > 0	0.085***	0.064***	-0.004	0.309***	0.323***	1.000								
(7) Cash	-0.023***	-0.241***	0.206***	0.030***	0.032***	-0.054***	1.000							
(8) Total assets	-0.018***	0.039***	-0.016***	0.023***	0.012***	0.100***	-0.088***	1.000						
(9) Capex	-0.050***	-0.025***	0.007	-0.012***	-0.011***	-0.031***	-0.077***	-0.003	1.000					
(10) Leverage	0.001	0.005	-0.001	0.001	0.001	-0.003	-0.013***	0.002	-0.000	1.000				
(11) EBIT	0.033***	0.108***	-0.074***	0.100***	0.075***	0.166***	-0.183***	0.059***	-0.017***	0.001	1.000			
(12) Net income < 0	-0.051***	-0.125***	0.077***	-0.096***	-0.111***	-0.215***	0.183***	-0.112***	-0.000	0.003	-0.404***	1.000		
(13) Tobin's Q	-0.002	-0.038***	0.033***	0.016***	0.017***	-0.034***	0.106***	-0.016***	0.012***	-0.002	-0.200***	0.043***	1.000	
(14) HHI	0.236***	0.539***	-0.399***	0.013***	0.009***	0.062***	-0.180***	-0.008***	-0.037***	0.001	0.082***	-0.098***	-0.024***	1.000
(15) No Firms (sic3)	-0.150***	-0.489***	0.408***	0.006	0.020***	-0.062***	0.387***	-0.056***	-0.023***	-0.009***	-0.173***	0.178***	0.065***	-0.448***

6.1.2 Propsensity scores distribution

Figure 6.1: Distribution of treated and control subsamples



Note: The two left plots indicate distribution of treated and control subsamples prior to matching. The two right plots show distribution of treated and control subsamples after matching.

6.1.3 Simulation analysis of common ownership concentration measures

This part explores the nature of the considered measures of common ownership. Three scenarios are considered that start from the same hypothetical starting conditions. It is needed to be mentioned that if all institutional investors ('owners') have a 5% or more in each firm in the same market, further increase in ownership share would not increase common ownership. Therefore scenarios that are considered here include several firms that are not commonly owned.

Table 6.2: Simulation scenarios details

	Scenarios		
State	**(1)** Mergers among firms	**(2)** Mergers among owners	**(3)** Increase in ownership shares
(1) Starting conditions	Five firms each has market share 20% Owner 1 has 5% in Firm 1 Owner 2 has 5% share in Firm 2 Owners 3-4 have 5% share in each Firm 3 - 5		
(2) Change A	Firms 4 and 5 merge	Owner 4 acquires Owner 5	Each owner has share of 7%
(3) Change B	Firms 1 and 2 merge	Owner 1 acquires Owner 2	Each owner has share of 9%

Table 6.3: Simulation results for common ownership concentration measures

	Scenarios											
State	**DCO**			**PCF**			**PCS**			**ΔHHI**		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
(0)	0.30	0.30	0.30	0.60	0.60	0.60	0.60	0.60	0.60	2400	2400	2400
(1)	0.17	0.30	0.30	0.60	0.50	0.60	0.50	0.60	0.60	1600	2400	2400
(2)	0.33	0.30	0.30	0.60	0.50	0.60	0.67	0.60	0.60	1600	2400	2400

6.1.4 Panel results using firm-level fixed effects

While the baseline results in the body of this study include SIC3 industry fixed effects and time fixed effects, for the purpose of feasibility the current part provides results with firm-specific fixed effects and time fixed effects. The results are similar to those with SIC3 industry-specific fixed effects in Table 3. Namely, higher common ownership is associated with greater payouts, higher markups, lower capital investments and more active industry consolidation.

Table 6.4: Panel results using firm-level fixed effects

	(1) Payouts	(2) Dividends	(3) Buybacks	(4) Markup	(5) Capex	(6) M&A
DCO	0.681***	0.437***	0.220	4.819	-0.947***	0.473**
	(3.80)	(6.91)	(2.43)	(0.43)	(-3.26)	(2.30)
DCO: *Mean* (*SD*)	0.14 (0.16)					
DepVar: *Mean* (*SD*)	5.8 (15.4)	4.5 (16.0)	4.4 (10.5)	105.9 (486.5)	6.2 (16.8)	7.2 (11.3)
Covariates	yes	yes	yes	yes	yes	yes
Firms FE	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes
Observations	78,487	78,487	78,487	78,487	78,171	76,138
Firms	10,357	10,357	10,357	10,357	10,330	10,357
R^2 *within*	0.264	0.228	0.249	0.001	0.023	0.072
R^2 *between*	0.273	0.494	0.189	0.011	0.003	0.069
R^2 *overall*	0.177	0.315	0.140	0.002	0.002	0.044

Covariates are $Size$, $Leverage$, $Net\ Income$, $Cash/TA$, Div_{t-1}/TA_{t-1}, $Net\ BB_{t-1}/TA_{t-1}$, HHI, $Size$. t-statistics are in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Payouts, Dividends, Buybacks, Capex, and M&A are scaled by total assets. DCO is density of common ownership. Fixed effects are included for SIC3 industries and years. Regressions are weighted using mean time-invariant market capitalization. Included mean and sd are general sample statistics conditional on positive values for payouts and M&A)

6.1.5 Tobit regressions results

It can indicate that the outcome variable, such as dividends, share buybacks or capital investments are truncated at zero. In this case it can be argued for the need of using Tobit regression that considers latent dependent variable y and its observed non-negative realization $y*$. Marginal effects from Tobit model are qualitatively and quantitatively similar to the coefficients from the baseline results (Table 3). This is in line with expectations, as according to Angrist and Pischke (2008) Tobit marginal effects should be similar in comparison to coefficients of the equivalent linear model.

Table 6.5: Tobit regression results

	(1) Payouts	(2) Dividends	(3) Buybacks	(4) Markup	(5) Capex	(6) M&A
DCO	0.541**	0.235*	0.396	3.463	-0.447***	0.373
	(2.06)	(1.82)	(1.45)	(1.14)	(-2.72)	(0.84)
DCO: *Mean* (*SD*)	0.14 (0.16)					
DepVar: *Mean* (*SD*)	5.8 (15.4)	4.5 (16.0)	4.4 (10.5)	105.9 (486.5)	6.2 (16.8)	7.2 (11.3)
Covariates	yes	yes	yes	yes	yes	yes
Industry FE	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes
Observations	78,487	78,487	78,487	78,487	78,171	76,138
Industries (SIC3)	245	245	245	245	245	245
$PseudoR^2$	0.115	0.226	0.122	0.003	0.060	0.058

Marginal effects are presented (t-statistics in parentheses). Covariates are Own, $Size$, $Leverage$, $Net\ Income$, $Cash/TA$, Div_{t-1}/TA_{t-1}, $Net\ BB_{t-1}/TA_{t-1}$, HHI, $Size$. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Payouts, Dividends, Buybacks, Capex, and M&A are scaled by total assets. DCO is density of common ownership. Fixed effects are included for SIC3 industries and years. Regressions are weighted using total market capitalization. Included mean and sd are general sample statistics conditional on positive values for payouts and M&A)

6.1.6 Placebo test for difference-in-differences results

This part provides results of placebo test for difference-in-difference analysis. For this purpose the event (scandal) time was falsely picked to be 2012, instead of 2004. Such choice is sufficiently far from the actual event time, and also is outside of the timing of financial crisis. In such set-up the difference-in-differences effect is not available, indicating the actual results were not driven by model design or any systematic difference between the two groups. The placebo analysis was performed using matched sample.

Table 6.6: Placebo test for difference-in-differences results

	(1) Payouts	(2) Dividends	(3) Buybacks	(4) Markup	(5) Capex	(6) M&A	(7) DCO
$Treated_j \cdot Post_t$	-0.735	-0.287	-0.448	-0.666	-0.587	0.504	0.003
	(-0.91)	(-0.73)	(-0.65)	(-0.19)	(-1.49)	(0.71)	(0.23)
$Post_t$	-0.532**	0.134	-0.666***	-0.293	1.095***	1.163***	0.189***
	(-2.16)	(1.11)	(-3.19)	(-0.28)	(9.09)	(5.06)	(42.89)
$Treated_j$	-6.336***	-0.443	-5.893***	-14.839**	-1.614*	2.828*	0.205***
	(-3.72)	(-0.53)	(-4.09)	(-2.06)	(-1.94)	(1.91)	(6.76)
Dep var: *Mean* (*SD*)	5.8	4.5	4.4	105.9	6.2	7.2	0.14
	(15.4)	(16.0)	(10.5)	(486.5)	(16.8)	(11.3)	(0.16)
Covariates	yes	yes	yes	yes	yes	yes	yes
Industry FE	yes	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes	yes
Observations	8,983	8,983	8,983	8,983	8,962	8,696	8,983
R^2	0.591	0.580	0.507	0.323	0.641	0.203	0.793

Covariates are $Inst\_own_{t-1}$, $Size$, $Leverage$, $Net\ Income$, $Cash/TA$, Div_{t-1}/TA_{t-1}, $Net\ BB_{t-1}/TA_{t-1}$, HHI. Fixed effects included for SIC3 industries and years. Regressions are weighted using total market capitalization. Sample is limited to years 2010 - 2014. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Included mean and sd are general sample statistics conditional on positive values for payouts and M&A

6.1.7 Panel result by sub-sample in wide format

The current part provides some results in 'wide' format. Specifically, these are the results when effect of common ownership concentration (proxied by DCO - density of common ownership) was tested to be different by sub-samples within the general sample. In the main body of the study these effects is evaluated by including interaction term between DCO and a respective dummy variable. An alternative set-up when a separate model is estimated for each sub-sample is provided in this appendix.

Such analysis was performed in the following cases: (1) to evaluate whether the effect of common ownership is different for the cross-owned firms and not cross-owned firms (Table F1, corresponds to Table 4 in the main body), (2) to evaluate whether the effect of common ownership is different for the firms in industries that face high import competition, aka 'China shock' (Table F2, corresponds to Table 11 in the main body), and (3) to evaluate whether the effect of common ownership is different for the firms in industries with low HHI versus industries with high HHI (Table F3, corresponds to Table 12 in the main body).

Table 6.7: Panel results for cross-owned and not cross-owned firms

	A. Dividends		B. Buybacks		C. Markup		D. Capex		E. M&A	
	CD = 1	CD = 0	CD = 1	CD = 0	CD = 1	CD = 0	CD = 1	CD = 0	CD = 1	CD = 0
DCO	1.347***	0.260***	1.961***	0.383**	6.469	3.650	0.216	-0.273	-0.333	0.251
	(6.11)	(3.99)	(4.15)	(2.43)	(1.51)	(1.08)	(0.44)	(-1.55)	(-0.53)	(1.14)
DCO: *Mean* (*Mean*)	0.14 (0.16)									
DepVar: *Mean* (*Mean*)	4.6 (16.0)		4.4 (10.5)		106.0 (486.5)		6.15 (16.8)		7.2 (11.3)	
Covariates	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Indus. FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
No. obs	7,934	70,553	7,934	70,553	7,934	70,553	7,898	70,273	7,731	68,407
Industr. (SIC3)	226	245	226	245	226	245	226	245	226	245
R^2	0.677	0.711	0.567	0.514	0.195	0.043	0.598	0.404	0.305	0.134

Dependent variable is indicated in each of the four panels. Two subsets are considered - where *CrossDummy* is 1 (CD = 1) and where it is 0 (CD = 0). DCO is density of common ownership. Fixed effects are included for SIC3 industries and years. Regressions are weighted using total market capitalization. Covariates are $Size$, $Leverage$, $Net\ Income$, $Cash/TA$, Div_{t-1}/TA_{t-1}, $Net\ BB_{t-1}/TA_{t-1}$, HHI, $Size$. Parentheses contain t-statistics. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Included mean and sd are general sample statistics conditional on positive values for payouts and M&A

Table 6.8: Panel results by 'China shock' effect

	A. Dividends		B. Buybacks		C. Markup		D. Capex		E. M&A	
	High CS	Low CS	High CS	Low CS	High CS	Low CS	High CS	Low CS	High CS	Low CS
DCO	-0.777**	0.639***	2.641**	4.958***	9.475*	8.303	0.421	0.680	3.132**	-1.251**
	(-1.97)	(3.58)	(2.52)	(9.70)	(1.92)	(1.26)	(0.88)	(1.58)	(2.13)	(-2.17)
DCO: *Mean* (*Mean*)	0.14 (0.16)									
DepVar: *Mean* (*Mean*)	4.6 (16.0)		4.4 (10.5)		106.0 (486.5)		6.15 (16.8)		7.2 (11.3)	
Covariates	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Industry FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
No. obs	3,077	6,400	3,077	6,400	3,077	6,400	3,077	6,400	3,067	6,394
Industries (SIC3)	33	61	33	61	33	61	33	61	33	61
R^2	0.588	0.403	0.627	0.359	0.533	0.331	0.513	0.050	0.317	0.616

Each regression is estimated for two-subsets, which respectively contain companies from the industries with low and high China shock. These are respectively below and above the median value. Dividends, Buybacks, Capex and M&A are scaled by total assets. *DCO* is density of common ownership. Fixed effects are included for SIC3 industries and years. Regressions are weighted using total market capitalization. Covariates are $Size$, $Leverage$, $Net\ Income$, $Cash/TA$, Div_{t-1}/TA_{t-1}, $Net\ BB_{t-1}/TA_{t-1}$, HHI.
t-statistics in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. The analysis is performed on the matched sample. Included mean and sd are general sample statistics conditional on positive values for payouts and M&A

Table 6.9: Panel results by industry concentration

	A. Dividends		B. Buybacks		C. Markup		D. Capex		E. M&A	
	Low HHI	High HHI	Low HHI	High HHI	Low HHI	High HHI	Low HHI	High HHI	Low HHI	High HHI
DCO	0.758***	0.349***	1.314***	-0.036	23.407**	-0.370	-1.878***	-0.332**	2.426***	-0.096
	(4.78)	(4.09)	(3.43)	(-0.16)	(2.42)	(-0.59)	(-3.89)	(-2.15)	(4.39)	(-0.36)
DCO: *Mean* (*Mean*)	0.14 (0.16)									
DepVar: *Mean* (*Mean*)	4.6 (16.0)		4.4 (10.5)		106.0 (486.5)		6.15 (16.8)		7.2 (11.3)	
Covariates	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Industry FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Year FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
No. obs	50,793	12,780	50,793	12,780	50,793	12,780	50,572	12,742	49,248	12,422
Industries (SIC3)	33	61	33	61	33	61	33	61	33	61
R^2	0.665	0.738	0.508	0.598	0.035	0.647	0.386	0.539	0.137	0.229

Data is grouped into two samples - low industry concentration ($HHI \leq 1,500$) and low industry concentration ($HHI \geq 2,500$). Dependent variables - Dividends, Buybacks, Capex and M&A are scaled by total assets. *DCO* - density of common ownership, measures common ownership concentration. Fixed effects included for SIC3 industries and years. Regression weights - Total market value. Covariates are $Size$, $Leverage$, $Net\ Income$, $Cash/TA$, Div_{t-1}/TA_{t-1}, $Net\ BB_{t-1}/TA_{t-1}$, HHI. t-statistics in parentheses. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Included mean and sd are general sample statistics (conditional on positive values for all payouts and M&A)

6.2 B: Gray Divorce Auxiliary Materials

6.2.1 Unmatched sample descriptive statistics

Table 6.10: Unmatched sample descriptive statistics

	Divorced		Married				
	Mean	SD	Mean	SD	Difference	t-stat	p-val
Non-white, %	0.23	0.42	0.22	0.42	0.00	0.25	0.80
Protestant, %	0.63	0.48	0.64	0.48	-0.01	-1.04	0.30
Catholic, %	0.22	0.41	0.19	0.39	0.03***	3.35	0.00
Age, years	62.03	9.87	61.71	9.62	0.33	1.40	0.16
HH size	2.12	1.30	2.58	1.29	-0.45***	-14.60	0.00
Total income	54880.91	226211.26	52700.74	62487.34	2180.17	0.55	0.58
Capital income	15439.63	196707.33	8449.66	31937.61	6989.97***	2.09	0.04
Net worth	224131.72	649289.76	220808.44	368213.39	3323.28	0.26	0.79
Hous. wealth	155461.78	568066.34	140717.53	321081.07	14744.25	1.34	0.18
Fin. wealth	62013.03	358189.81	60072.13	209414.88	1940.91	0.28	0.78
Stocks	39389.87	266011.90	25998.29	107976.36	13391.58***	2.78	0.01
Bonds	4637.49	54386.26	3635.51	38192.86	1001.98	0.89	0.37
Bank accs	14413.54	49663.65	17827.94	76333.86	-3414.40**	-2.21	0.03
CDs	3510.79	19680.61	6716.43	45700.30	-3205.64***	-3.80	0.00
Vehicles	11386.30	23842.79	13285.36	15756.79	-1899.05***	-3.95	0.00
IRA	30477.05	121602.37	36003.96	112110.04	-5526.91***	-1.98	0.05

All wealth variables are in per equivalent person basis. Sig: *** p<0.01, ** p<0.05, * p<0.1. Unmatched sample includes N_1 = 4,472 treated and N_2 = 542,204 non-treated observations.

6.3 C: Modernization Hypothesis Auxiliary Materials

6.3.1 Data sources and details

Table 6.11: Data sources and details

Variable	Measurement details	Source
Uncertainty	Months between collapse of socialist system and first legitimate parliamentary elections	Own calculations, Walder et al. (2015)
Real GDP per capita	US Dollars of 2010	World Bank, WDI
Gini index	Range from 0 (complete equality) to 1 (absolute inequality)	Standardized World Income Inequality Database (SWIID)
Human development index (HDI)	Normalized from 0 (lowest human development) to 1	United Nations Development Program (UNDP)
Real GDP per capita in 1990	1990 international Geary-Khamis Dollars	Maddison Angus database (www.ggdc.net)
Political freedom*	From 0 (lowest democracy) to 100 (highest democracy)	Freedom House
Freedom of press*	From 0 (lowest democracy) to 100 (highest democracy)	Freedom House
Regime score*	From 0 (lowest democracy) to 100 (highest democracy)	Polity International
Resource dependency	US Dollars of 2010	World Bank, WDI
Market memory	US Dollars of 2010	World Bank, WDI
Geographic location	US Dollars of 2010	World Bank, WDI
State institutions and Independence	Initial institutional country characteristics: 0 -newly created states, 1- countries of federal states, 2- were independent before 1989	De Melo (2001)
Regional tensions	Indicator of countries that had military conflict due to border disputes	De Melo (2001)

*Further details of democracy variables are provided in Appendix C.2

6.3.2 Democracy variables details

There is multitude of democracy indexes, among which there is heterogeneity with respect to their input data sources, country and time coverage, and academic impact (Coppedge et al. 2016).

Political freedom Original measurement: defined on the scale between 0 and 100. Freedom House determines countries as 'Free' with values between 0-30, 'Partly free' - 31-60, and 'not free' - 61-100. The variable was transformed to 0 - 100 scale to be increasing in the level of Political freedom.

Political regime Original measurement: Based on Polity2 which is the difference between the polity Democracy index and Autocracy index for the same country. The original Polity2 variable is defined on the range from -9 to +9. The variable was transformed to 0-100 scale to be increasing in the level of democracy.

Freedom of Press Original measurement: defined on the scale between 0 and 100, where 100 represents the least freedom of the press. Variable was transformed to the scale 0 - 100, where indicator is increasing in the freedom of the press.

6.3.3 Construction of political disruption variable

A main contribution of the current study is its development of a measurement of 'Political disruption' in CEE countries. The suggested measurement here closely followed the definition of this phenomenon by Walder et al. (2015). The task requires that the start and the end of 'political disruption' is defined for each of the 28 CEE countries. Therefore 'political disruption' is synonymous to lack of legitimate power, and is measured as duration of the uncertainty is defined as the length, in days, of the respective time interval. The approach taken to code T_1 and T_2 is detailed below.

$$D = T_2 - T_1$$

Start of the uncertainty, T_1, was determined as follows. For the countries that constituted former Soviet Union, the uncertainty about the future ability of communist party emerged as the party leadership implemented new policy of 'perestroika' and 'glasnost'. The leadership of communist party decided it has to change and implement elements of democracy and private ownership into the Soviet economy. For post-USSR countries, the starting event was the adoption of Law on Cooperatives. The law established the possibility of private ownership and enabled private owners to implement international trade and other economic transactions. As the relevant mechanisms were not yet properly defined, this caused processes of cashing out state enterprises and transferring cash into off-shores, as well as de-facto privatization of state enterprises at significant discount to their fair values. For the countries that were not part of the Soviet Union, the beginning of 'political disruption' was either a decisive demonstration, a decisive protest, or other equally significant event, which involved a large number of people, and led to the resign of communist presidents in those economies.

End of uncertainty, T_2, was estimated as the date of the multi-party parliamentary elections in the post-USSR economies. Multi-party parliamentary elections was the way that new forces came to power and implement constitutional reform, as well as adopted other important laws on the path of transition towards market economy. For the non-USSR countries, the end of the 'political disruption' period was the earlier of either multi-party elections, or assignment of a new democratic government. There were no democratic

governments assigned prior to multi-party elections in the post-USSR countries. In one of the post-USSR countries, Turkmenistan, there were no multi-party elections, and the 'end' of political disruption was the date of referendum when the term of the president was extended significantly without the need for re-election. This event ended the uncertainty regarding the control over the country economy and de-facto established a dictatorship.

6.3.4 List of CEE countries

Table 6.12: List of CEE countries

	Country name	Code
1	Albania	ALB
2	Armenia	ARM
3	Azerbaijan	AZE
4	Bulgaria	BGR
5	Bosnia and Herzegovina	BIH
6	Belarus	BLR
7	Czech Republic	CZE
8	Estonia	EST
9	Georgia	GEO
10	Croatia	HRV
11	Hungary	HUN
12	Kazakhstan	KAZ
13	Kyrgyzstan	KGZ
14	Lithuania	LTU
15	Latvia	LVA
16	Moldova	MDA
17	Macedonia	MKD
18	Montenegro	MNE
19	Mongolia	MGN
20	Poland	POL
21	Romania	ROU
22	Russian Federation	RUS
23	Serbia	SRB
24	Slovakia	SVK
25	Slovenia	SVN
26	Tajikistan	TJK
27	Turkmenistan	TKM
28	Ukraine	UKR
29	Uzbekistan	UZB
30	Kosovo	XKX

6.3.5 Pairwise correlation for the annual data

Table 6.13: Pairwise correlation for the annual data

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) *Uncertainty*	1.000											
(2) *Marmem*	0.872*	1.000										
(3) *Locat*	-0.376*	-0.454*	1.000									
(4) *State*	-0.596*	-0.420*	0.324*	1.000								
(5) *RT*	0.182*	0.076*	-0.347*	-0.281*	1.000							
(6) *Fhouse*	-0.556*	-0.629*	0.604*	0.468*	-0.245*	1.000						
(7) *Polity*	-0.487*	-0.539*	0.476*	0.430*	-0.114*	0.869*	1.000					
(8) *Press*	-0.613*	-0.705*	0.614*	0.377*	-0.251*	0.943*	0.844*	1.000				
(9) *GDP*	-0.473*	-0.515*	0.742*	0.232*	-0.322*	0.604*	0.450*	0.622*	1.000			
(10) *GDP*90	0.036	-0.182*	0.594*	-0.218*	-0.148*	0.396*	0.289*	0.442*	0.690*	1.000		
(11) *HDI*	-0.383*	-0.448*	0.515*	0.132*	-0.235*	0.504*	0.446*	0.533*	0.830*	0.584*	1.000	
(12) *Gini*	0.399*	0.351*	-0.291*	-0.153*	0.332*	-0.296*	-0.145*	-0.307*	-0.490*	-0.361*	-0.365*	1.000

Variables: *Uncertainty* - measure of political uncertainty (disruption), *Marmem* - market memory, *Locat* - geographic location, *State* - state institutions and independence, *RT* - regional tensions, *F house* - Freedom House measure of democracy, *P olity* - political regime measure, *Press* - freedom of the press, *GDP* - real GDP per capita, *GDP* 90 - real GDP per capita in 1990, *HDI* - Human development index, *Gini* - Gini index. Significance: *- 10%, ** - 5%, *** - 1%.

6.3.6 Democracy in CEE

Figure 6.2: Political freedom by country

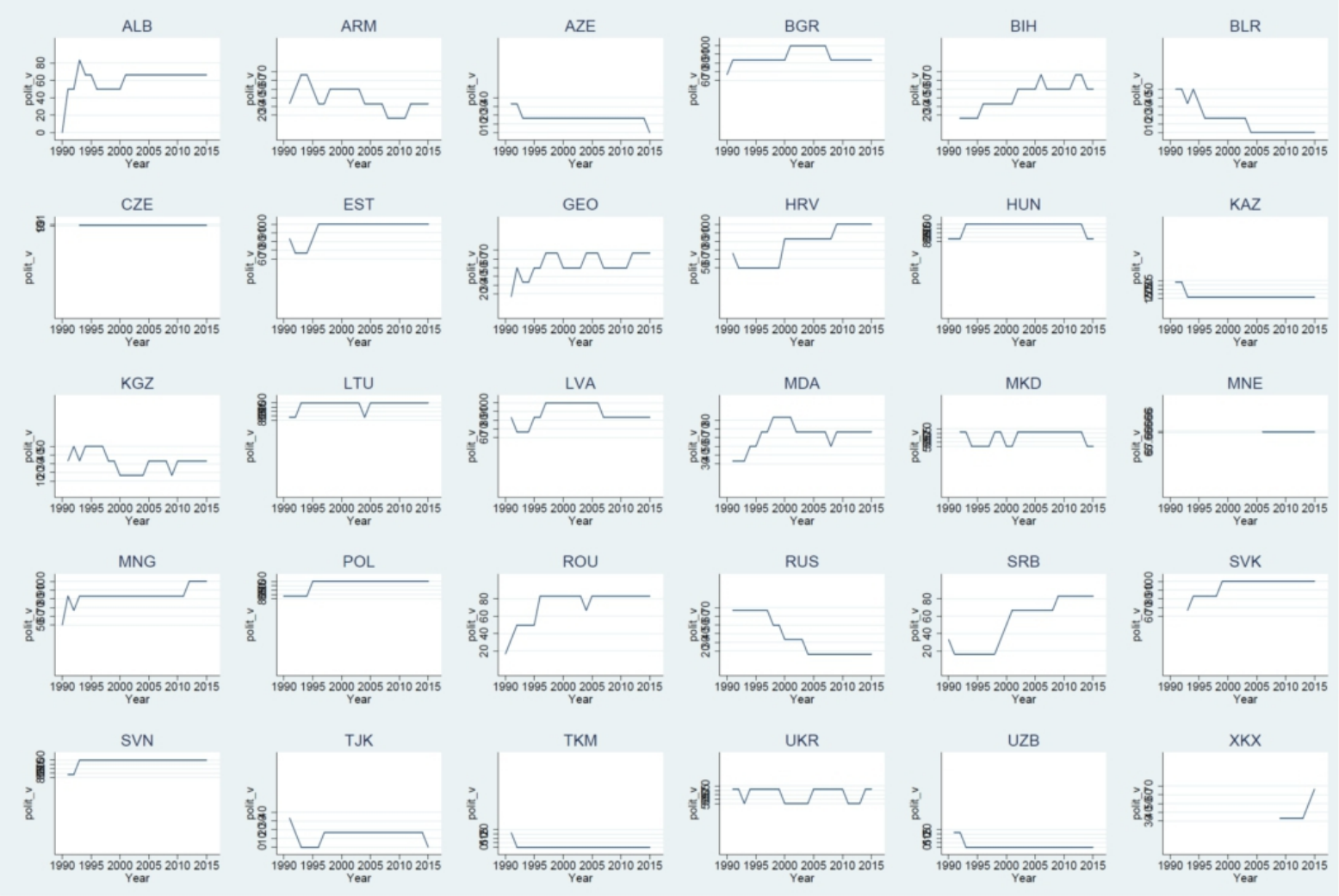


Figure 6.3: Polity regime by country

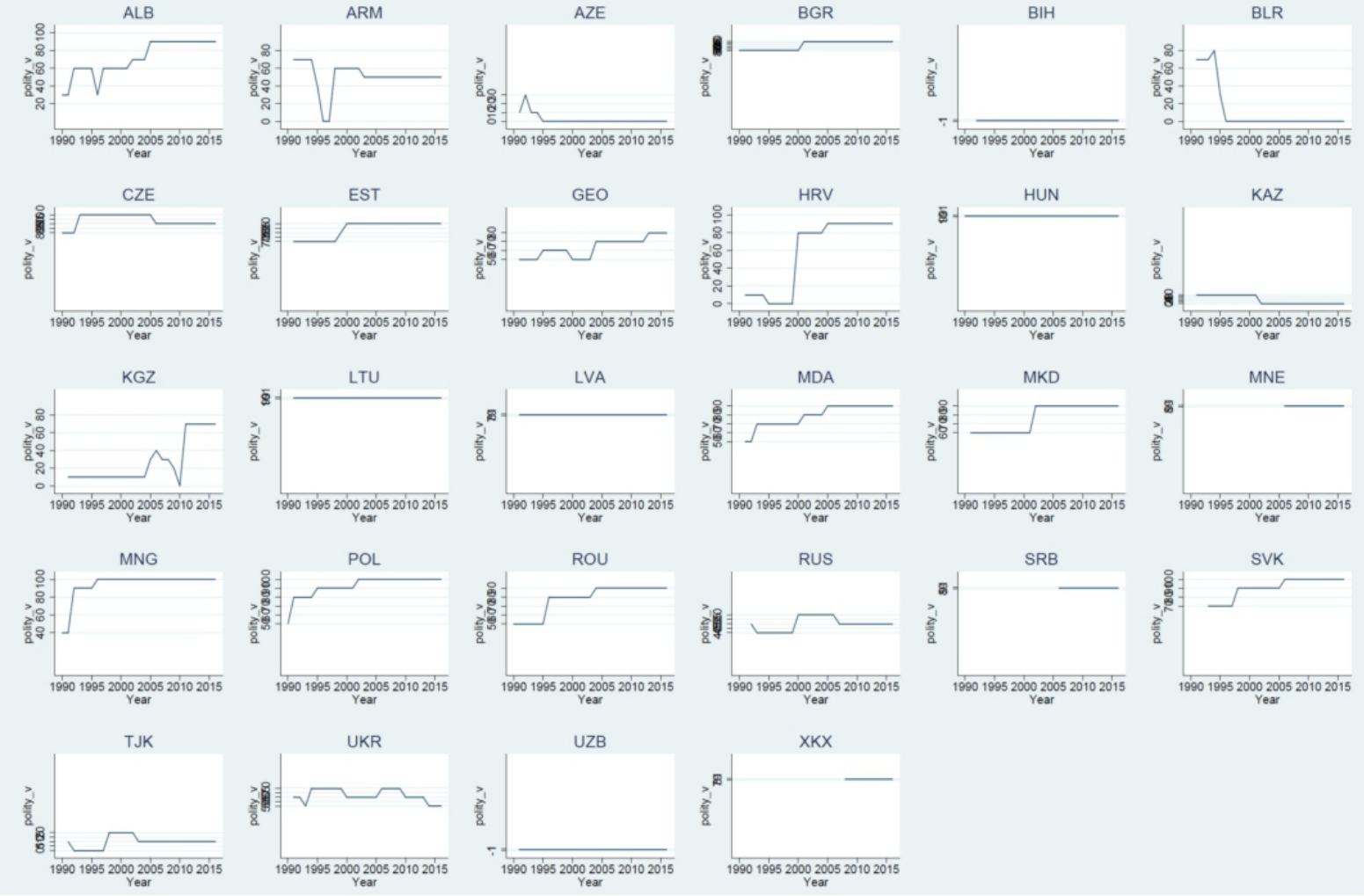
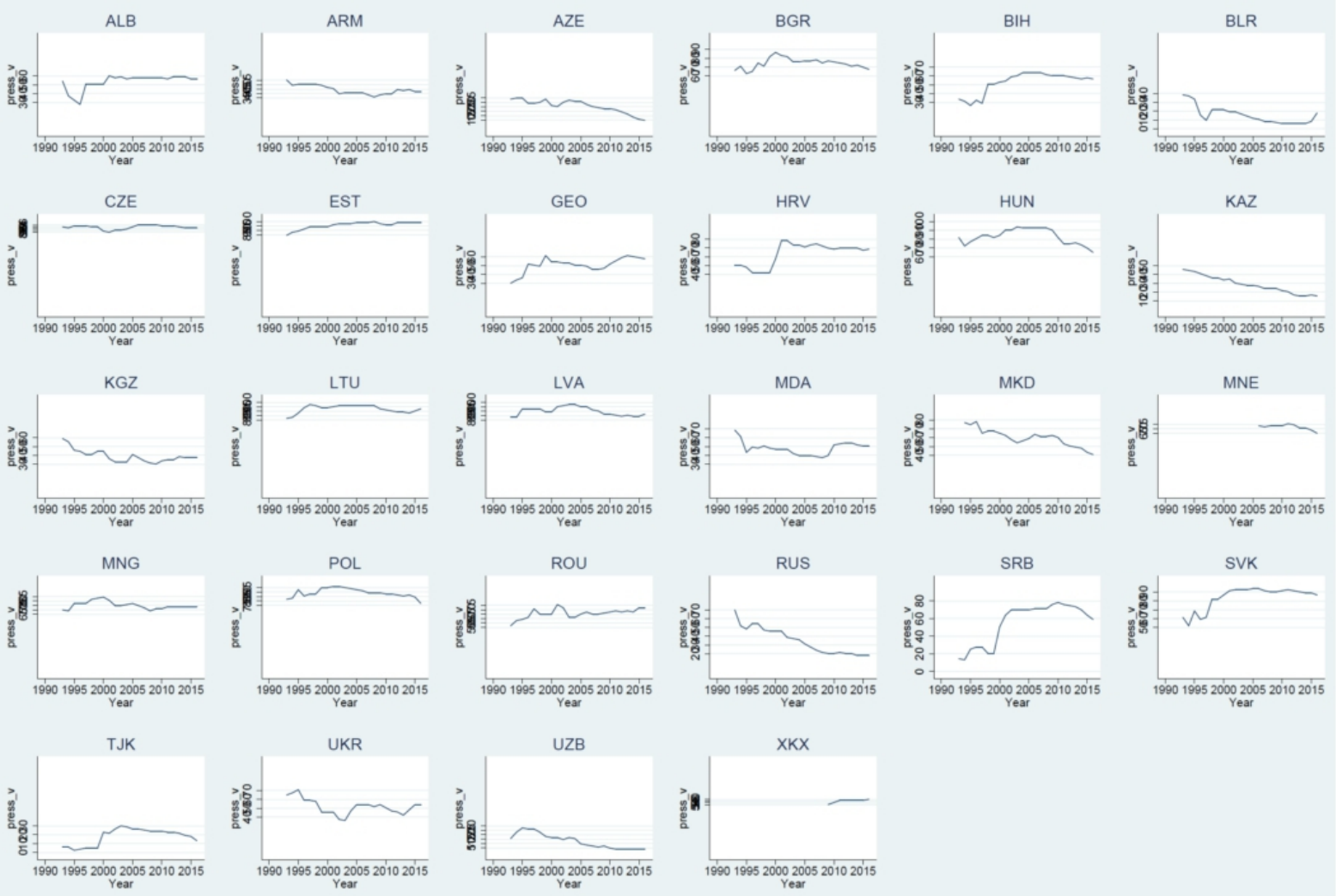


Figure 6.4: Freedom of the press by country



7.0 Bibliography

Acemoglu, D. (2008). Oligarchic versus democratic societies. *Journal of the European Economic Association*, 6(1):1–44.

Acemoglu, D., Autor, D., Dorn, D., Hanson, G., and Price, B. (2016). Import competition and the great us employment sag of the 2000s. *Journal of Labor Economics*, 34(1):141–198.

Acemoglu, D., Johnson, S., Robinson, J. A., and Yared, P. (2008). Income and democracy. *American Economic Review*, 98(3):808–42.

Acemoglu, D., Johnson, S., Robinson, J. A., and Yared, P. (2009). Reevaluating the modernization hypothesis. *Journal of monetary economics*, 56(8):1043–1058.

Acemoglu, D., Naidu, S., Restrepo, P., and Robinson, J. A. (2019). Democracy does cause growth. *Journal of Political Economy*, 127(1):47–100.

Almeida, H., Fos, V., and Kronlund, M. (2018). The real effect of share repurchases. *Journal of Financial Economics*, 119:168–185.

Angrist, J. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Anton, M. and Polk, C. (2014). Connected stocks. *The Journal of Finance*, LXIX:1099–1127.

Arellano, M. and Bond, S. (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The review of economic studies*, 58(2):277–297.

Azar, J. (2011). *A New Look at Oligopoly: Implicit Collusion Through Portfolio Diversification*. SSRN Electronic Journal.

Azar, J. (2018). Anticompetitive effects of common ownership. *The Journal of Finance*, LXXIII(4).

Azar, J., Raina, S., and Schmalz, M. (2016). *Ultimate ownership and bank competition*.

Barro, R. J. (1996). Democracy and growth. *Journal of economic growth*, 1(1):1–27.

Becker, G. S., Landes, E. M., and Michael, R. T. (1977). An economic analysis of marital instability. *Journal of political Economy*, 85(6):1141–1187.

Benhabib, J., Corvalan, A., and Spiegel, M. M. (2011). Reestablishing the income-democracy nexus. Technical report, National Bureau of Economic Research.

Benhabib, J., Corvalan, A., and Spiegel, M. M. (2013). Income and democracy: Evidence from nonlinear estimations. *Economics Letters*, 118(3):489–492.

BenYishay, A. and Grosjean, P. (2014). Initial endowments and economic reform in 27 post-socialist countries. *Journal of Comparative Economics*, 42(4):892–906.

Brown, S. and Wright, M. (2017). Marriage, cohabitation, and divorce in later life. *Innovation in Aging*.

Bugliari, D., Carroll, J., Hayden, O., Hayes, J., Hurd, M., and Karabatakis, A. (2020). Rand hrs longitudinal file 2016 (v2) documentation. Technical report, RAND Center for the Study of Ageing.

Cervellati, M., Jung, F., Sunde, U., and Vischer, T. (2014). Income and democracy: Comment. *American Economic Review*, 104(2):707–19.

Chung, K. and Zhang, H. (2011). Corporate governance and institutional ownership. *Journal of Financial and Quantitative Analysis*, 46(1):247–273.

Coppedge, M., Gerring, J., Lindberg, S. I., Skaaning, S.-E., Teorell, J., Altman, D., Bernhard, M., Fish, M., Glynn, A., Hicken, A., et al. (2016). V-dem codebook v6, varieties of democracy (vdem) project. *Gothenburg, V-Dem Institute*.

Crane, A., Michenaud, S., and Weston, J. (2016). The effect of institutional ownership on payout policy: Evidence from index thresholds. *The Review of Financial Studies*, 29(6):1377–1408.

Crowley, J. E. (2018). *Gray Divorce: What we lose and gain from mid-life splits*. University of California Press.

De Melo, D. (2001). Circumstance and choice: the role of initial conditions and policies in transition economies. *World Bank Economic Review*, 10(1):1–31.

DeAngelo, H., DeAngelo, L., and Skinner, D. (2009). Corporate payout policy. *Foundations and Trends in Finance*, 3:95–287.

Del Bono, E., Sala, E., Hancock, R., Gunnell, C., and Parisi, L. (2007). Gender, older people and social exclusion: A gendered review and secondary analysis of the data. Technical report, ISER Working Paper Series.

DeLiema, M., Deevy, M., Lusardi, A., and Mitchell, O. S. (2018). Financial fraud among older americans: Evidence and implications. Technical report, National Bureau of Economic Research.

Denis, D. and Osobov, I. (2008). Why do firms pay dividends? international evidence on the determinants of dividend policy. *Journal of Financial Economics*, 89:62–82.

Dennis, P. J., Gerardi, K., and Schenone, C. (2018). *Common ownership does not have anti-competitive effects in the airline industry*. Available at SSRN 3063465.

Farrell, J. (1985). Owner-consumers and efficiency. *Economic Letters*, 19:303–306.

FreedomHouse (2017). Freedom in the world 2017. populists and autocrats: The dual threat to global democracy. *Freedom House*.

Fried, J. and Wang, C. (2018). Short-termism and capital flows. *Working Paper 17-062*.

Gardner, J. and Oswald, A. J. (2006). Do divorcing couples become happier by breaking up? *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 169(2):319–336.

Gareth, J., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An Introduction to Statistical Learning: With Applications in R*. Springer.

Geron, A. (2017). *Hands-On Machine Learning with Scikit-Learn and TensorFlow*. O'Reilly Media.

Gilbert, D. L. (2017). *The American class structure in an age of growing inequality*. SAGE publications.

Gilje, E., Gormley, T. A., and Levit, D. (2018). *The rise of common ownership*.

Grinstein, Y. and Michaely, R. (2005). *Institutional Holdings and Payout Policy*. The Journal of Finance, LX(3).

Gutierrez, G. and Philippon, T. (2016). *Investment-less growth: An empirical investigation*. National Bureau of Economic Research.

Gutierrez, G. and Philippon, T. (2018). Ownership, concentration and investment. *AEA Papers and Proceedings, American Economic Association*, 108:432–437.

Gutiérrez, G. and Philippon, T. (2017). *Declining Competition and Investment in the U*.

Haider, S., Jacknowitz, A., and Schoeni, R. (2003). The economic status of elderly divorced women. *Research Paper No. WP2003-046*.

He, J. and Huang, J. (2017). Product market competition in a world of cross-ownership: Evidence from institutional blockholdings. *The Review of Financial Studies*, 30(8):2674–2718.

Hellman, J. S. (1998). Winners take all: the politics of partial reform in postcommunist transitions. *World politics*, 50(2):203–234.

Ho, D., Imai, K., King, G., and Stuart, E. (2007a). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political analysis*, 15(3):199–236.

Ho, D., Imai, K., King, G., and Stuart, E. (2007b). Matchit: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*.

Houge, T. and Wellman, J. (2005). Fallout from the mutual fund tracking scandal. *Journal of Business Ethics*, 62:129–139.

Jensen, M. C. and Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of financial economics*, 3(4):305–360.

King, G. and Nielsen, R. (2018). *Why Propensity Scores Should Not Be Used for Matching*. Massachusetts Institute of Technology.

Kisin, R. (2011). *The impact of shareholders on corporate investment: Evidence from mutual fund holdings*. Washington university.

Koch, A., Panayides, M., and Thomas, S. (2018). *Commonality in institutional ownership and competition in product markets*. Working paper, University of Pittsburgh.

Koch, A., Panayides, M., and Thomas, S. (2019). *Common Ownership and Competition in Product Markets*. Available at: available at: https://ssrn.com/abstract=2965058.

Lazonick, W. (2014). *Profits without prosperity*. Harvard Business Review.

Lipset, S. (1960). Martin, political man: The social bases of politics.

Lipset, S. M. (1959). Some social requisites of democracy: Economic development and political legitimacy. *American political science review*, 53(1):69–105.

Luo, Y., Hawkley, L. C., Waite, L. J., and Cacioppo, J. T. (2012). Loneliness, health, and mortality in old age: a national longitudinal study. *Social science & medicine*, 74(6):907–914.

Maddison, A. (2010). Statistics on world population, gdp and per capita gdp, 1-2008 ad. *Historical Statistics*, pages 1–36.

Matvos, G. and Ostrovsky, M. (2008). Cross-ownership, returns, and voting in mergers. *Journal of Financial Economics*, 89(3):391–403.

Mazzonna, F., Peracchi, F., et al. (2018). Self-assessed cognitive ability and financial wealth: Are people aware of their cognitive decline? Technical report, Einaudi Institute for Economics and Finance (EIEF).

OECD (2011). *Divided We Stand: Why Inequality Keeps Rising*. OECD Publishing.

Passarelli, F. and Tabellini, G. (2017). Emotions and political unrest. *Journal of Political Economy*, 125(3):903–946.

Pierce, J. R. and Schott, P. K. (2016). The surprisingly swift decline of us manufacturing employment. *American Economic Review*, 106(7):1632–62.

Randolph, J., Falbe, K., Kureethara, A., and Balloun, J. (2014). A step-by-step guide to propensity score matching in r. *Practical Assessment, Research and Evaluation*, 19(18):1–6.

Rodrik, D. and Wacziarg, R. (2005). Do democratic transitions produce bad economic outcomes? *American Economic Review*, 95(2):50–55.

Rojas, M. (2014). *Encyclopedia of Quality of Life and Well-Being Research*. Springer Netherlands, Dordrecht.

Rotemberg, J. (1984). *Financial transaction costs and industrial performance*. MIT Press.

Schmalz, M. (2018). Common-ownership concentration and corporate conduct. *Annual Review of Financial Economics*, 10:413–448.

Sharma, A. (2015). Divorce/separation in later-life: A fixed effects analysis of economic well-being by gender. *Journal of Family and Economic Issues*, 36(2):299–306.

Smock, P., Manning, W., and Gupta, S. (1999). The effect of marriage and divorce on women's economic well-being. *American Sociological Review*, 64(6):794–812.

South, S. J. and Spitze, G. (1986). Determinants of divorce over the marital life course. *American sociological review*, pages 583–590.

Stepler, R. (2017). Led by baby boomers, divorce rates climb for america's 50+ population. Technical report, Pew Research.

Taddy, M. (2019). *Business Data Science*. McGraw Hill.

Walder, A. G., Isaacson, A., and Lu, Q. (2015). After state socialism: the political origins of transitional recessions. *American Sociological Review*, 80(2):444–468.

Wucherpfennig, J. and Deutsch, F. (2009). Modernization and democracy: Theories and evidence revisited. *Living Reviews in Democracy*, 1.

Zagorsky, J. L. (2005). Marriage and divorce's impact on wealth. *Journal of Sociology*, 41(4):406–424.