Spring 2022

Essays in Financial Economics

Canyao Liu
Yale University Graduate School of Arts and Sciences, liucanyaoalex@gmail.com

Abstract

Essays in Financial Economics

Canyao Liu
2022

In my dissertation, I dive into three specific areas in financial economics.

Chapter 1 of my Ph.D. dissertation studies how the boom in off-exchange trading at the market close affects the close price discovery. In recent years, investment banks like Goldman Sachs have started a "guaranteed close" business where investors looking to buy or sell shares of a certain stock can get a guarantee from the bank to execute their orders at the close price set on the primary exchange. Using the TAQ data and a quasi-experimental shock from NYSE fee cut, we find that when the fraction of trades through "guaranteed close" increases, the informativeness of close price increases. We develop a model where investors choose which venue to trade in. A bank conducting "guaranteed close" business competes with the exchange on transaction fees, and gains profit from trading strategically utilizing the order flow information. The bank's trading activity concentrates the price-relevant information into the exchange. Consequently, the "guaranteed close" improves price discovery at the market close.

Chapter 2 of my Ph.D. dissertation studies the long-term effects of experiencing high levels of job demands on the aging and mortality of CEOs. The estimation exploits variation in industry crises and takeover protection. First, we apply neural-network based ML techniques to assess visible signs of aging in pictures of CEOs. We estimate that exposure to a distress shock during the Great Recession increases CEOs' apparent age by one year over the next decade. Second, using hand-collected data on the dates of birth and death for 1,605 CEOs of large, publicly-listed U.S. firms, we estimate the resulting changes in mortality. The hazard estimates indicate that CEOs' lifespan decreases by 1.5 years in response to an industry-wide downturn, and increases by two years when insulated from

market discipline via anti-takeover laws. Our findings imply significant health costs of managerial stress, also relative to known health risks.

Chapter 3 of my Ph.D. dissertation provides an economically interpretable and easy-to-calculate approximation to optimal portfolio choice over the life cycle. The standard literature that solves the numerical optimal portfolio policy requires complicated backward induction, making it hard to apply for providing financial advice. Real-world financial advisors, on the other hand, tend to neglect the risky nature of human capital and offer advice that is not truly optimal. We bridge the gap by first using a reduced-form regression to predict discount rates of future incomes over an agent's life. Our prediction method achieves an R^2 of more than 90% over a wide range of simulations. Furthermore, by plugging the discount rates we predict into Merton (1969) formula, we obtain an approximate solution that has an average difference within 2% when compared to optimal solution solved through backward induction.

Essays in Financial Economics

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by
Canyao Liu

Dissertation Director: Nicholas Barberis

May 2022

Acknowledgments

I am deeply indebted to my doctoral advisor Professor Nicholas Barberis and my committee members Professor James Choi, Professor Stefano Giglio, and Professor Kelly Shue. Without their generous and invaluable support, I would never have accomplished the progress so far. Beyond leading me into the world of finance research and offering me guidance on academic research, they also provide me a sense of belonging. To me, coming from a foreign country with different culture sometimes means challenges. But here at Yale, at this department, I have always been feeling at home.

I am also immensely grateful for having my brilliant coauthors: Mark Borgschulte, James Choi, Marius Guenzel, Jingxiong Hu, Ulrike Malmendier, and Jiaheng Yu. Working with them has not only been enjoyable but also been a great learning process for me. I have learned so much from every single one of them by discussing the research ideas and tackling the research problems together. Without their intelligence, dedication, and support, this dissertation would never have come together. The process of producing research is sometimes challenging and tedious, but working with them added so much memorable moments to this journey.

I want to thank my lovely friends at Yale. Some of them are in the same department as me. We chatted about research, about all the finance questions in the world, about each other's stories and dreams over coffee, beers, lunches, and dinners, and recently even breakfasts. I will miss and cherish all the moments and discussions we had, and I believe in one day, some of those will change the world. Some of them come from other departments, and I am so lucky to have known them through this bridge of Yale. They have always opened different worlds for me. We have shared so many joyful moments together and they have added some much color to my life at Yale. I will miss you all very much and I will see you in other corners of the world.

I want to thank my beloved girlfriend. I have always considered meeting her as the luckiest thing that has ever happened in my life. We met on the second day I came to Yale, and from that moment on, I know my life has changed. We have been through moments of laughter and tears together and shared each other's success and failure. Her intelligence, passion for life, and deep care for other people have always touched me and motivated me to become a better person. Standing by her side, I know there's nothing in this world I can't do. I owe every bit of my accomplishments to her.

Last but not least, I want thank my dear parents who have relentlessly supported me throughout my life. They have never said no to any of my needs and have provided me tremendous support for the tough moments I have been through. For that, I owe all of my joy and achievements to them.

This work is dedicated to my beloved parents and girlfriend without whose constant support this thesis paper was not possible.

Contents

List of Figures

List of Tables

Chapter 1

Dual trading and price discovery at market close: Theory and evidence

JingXiong Hu, Canyao Liu, Jiaheng Yu

Abstract

Investment banks like Goldman Sachs have started to offer a "guaranteed close" that executes orders for investors at close prices set on the primary exchanges. About 30% of the close auction volume was executed through this channel in recent years. Using a quasi-experimental shock – an NYSE fee cut in January 2018 – we find that "guaranteed close" improves the informativeness of close prices. This finding cannot be explained by existing theories. We develop a model with two trading venues – the exchange and a bank, executing orders at the same price – where dual trading of the bank improves price discovery. The implications of our model apply more generally to scenarios where brokers trade on order flow information.

1.1 Introduction

Over recent years, the surge of trading volume in the close auctions of primary listing markets has gained great attention – the total trading volume executed in the close auctions has increased by more than 120% from 2012 to 2018, and accounts for nearly 8% of the total trading volume in 2018. This is partly driven by index funds and ETFs which are benchmarked against closing prices for indices and seek to trade at the market close price to minimize tracking error (Bogousslavsky and Muravyev, 2021). In the meantime, given this demand, several investment banks also started to provide liquidity at the market close price since late 2016. Specifically, they offer a "guaranteed close" service, where investors can get a guarantee from a bank to execute their orders at the close price set on the stock's listing exchange.

As shown in Figure 1.1, daily trading volume through "guaranteed close" has risen sharply, reaching almost 30% of the daily volume executed in close auctions in 2018.[1]

"Guaranteed close" is attractive because of its low transaction fee. In fact, since a stock's close auction is held only by its listing exchange, the lack of competition drove up the close auction fees: NYSE's base rate has gone up by 16% and Nasdaq's by 60% from 2012 to 2018.[2] Although the transaction fee in "guaranteed close" is lower, one concern is that the dual trading of the bank may affect the close price. For example, executives at NYSE and Nasdaq suggested that "... if more trading moves to banks, it will make close prices less trustworthy..."(WSJ, 2018). SEC DERA (2017), however, questions the validity of this claim.

Given these concerns and the importance of the close price, we take the first step to formally study the impact of "guaranteed close" on the informativeness of the close price,

[1] WSJ (2018) and SEC report: Securities Exchange Act Release No. 34-80683 (May 6, 2017), document the same observation.

[2] See Securities Exchange Act Release No. 34-80683 (May 6, 2017) (SR-BatsBZX-2017-34). See Tier F under Execution Fees for the Nasdaq Closing Cross in the Nasdaq fee schedule available at [link]; an Liquidity Indicator 7 in the NYSE fee schedule available at [link] .

both empirically and theoretically. [3]

One major contribution of our paper is that we provide quasi-experimental evidence that "guaranteed close" improves price discovery. This cannot be explained by the predictions of the dual-trading literature where, for example, Sarkar (1995) suggests no effect of dual trading on price discovery. Our unique empirical setting is ideal for testing the effect of dual trading: banks trade on their own accounts after viewing order flows of customers, and a quasi-experimental shock reduces the order flow to the banks; designated market makers clear the market in a single auction and set the price, corresponding to the Kyle (1985) framework that the dual-trading literature usually builds on. Different from dark pools which may improve discovery due to execution risk (Zhu, 2014), "guaranteed close" has no execution risk.

To explain the improvement in price discovery, we build a model based on the single-period model in Kyle (1985). Our model provides a novel mechanism through which dual-trading improves price discovery. In our model, traders explicitly choose between two trading venues (i.e., the bank and the close auction) based on transaction costs. As long as the proportions of informed orders relative to uninformed orders are different between two venues, the net total orders to the bank provide the bank with information to trade profitably. The bank trades in the same direction as the net orders when the proportion of informed orders at the bank is higher than in the close auction, and in the opposite direction otherwise. Such trading activity alone amplifies the fraction of informed orders received by the market maker. Unlike previous literature on dual trading (e.g. Sarkar (1995)) that usually assume all informed traders to be large and sensitive to price impact, our model introduces small informed traders with unit demand who do not react to changes in market depth. Hence the amplification effect of the bank on the fraction of informed

[3]The close price is typically regarded as the most important price of the day. It is used in the computation of net asset values (NAV), margin accounts, marking to market at the clearinghouse, settlement of derivative securities and the calculation of risk metrics such as volatility. See Section 1.7 in the internet appendix on how deviations in close prices also have important welfare implications for investors in general.

orders cannot be fully offset by the reaction of informed traders reducing their orders as in Sarkar (1995). While informed traders are usually believed to be large, emerging evidence shows that some of them are in fact small (Farrell et al., 2021; Boehmer et al., 2021; Kelley and Tetlock, 2013, 2017; Kaniel et al., 2012). The mechanism of our model works as long as some informed traders are small. Our model and empirical evidence have general implications for the effect of the assorted venue choices in today's brokerage industry on price discovery. As some brokers like Robinhood charge zero commission fee and sell order flow data to sophisticated investment firms (comparable to the bank's "guaranteed close"), others still charge a commission fee and offer an explicit execution price (comparable to the close auction). The brokerage firm Interactive Brokers even simultaneously offers two options by itself: IBKR Lite and IBKR Pro. "If it is IBKR Lite with zero commissions ,..., we send them off to a market maker,..., and there is payment for order flow that comes back and you may not get as good of an execution ... If it is IBKR Pro, you will get better execution."(Steve Sanders, executive vice president of Interactive Brokers, link).

Our empirical result relies on using the NYSE close auction fee cut in January 2018 as a quasi-experiment. In the fee cut, NYSE reduced the Tier 1 rate and Tier 2 rate (which applies to broker-dealers with high trading volumes in the close auction) significantly, but left the Non-Tier rates (which applies to the others) almost untouched. Given the differential changes, we expect the fee cut to have differential impact on different stocks. The broker-dealers who qualify for Tier 1 or Tier 2 rates may have a client base that has high demand for trading at the market close price – for example, index funds and ETFs. The fee cut for broker-dealers would pass through to them. Stocks with higher ETF/index fund ownership should therefore be more exposed to the fee cut. Similarly, stocks more heavily traded at the bank's venue ex ante would also be more exposed to the fee cut. This is likely because they have higher ETF/index fund ownership, while ETF/index funds are more sensitive to fees and in the meantime received larger NYSE fee cuts.[4]

[4]This is why these stocks are traded more heavily at the bank in the first place.

We use a standard difference-in-differences strategy. To measure each stock's trading volume via "guaranteed close", we use the NYSE Trade-and-Quote (TAQ) millisecond level data, and proxy this volume by the volume of off-exchange trades between 4:00 p.m.–4:10 p.m. EST that are executed at the official close price.[5] We show that ETF/index fund ownership is strongly correlated with the "guaranteed close" volume.[6] We measure the informativeness of the close price by the "closeness" (mean squared error) between the close price and the next day's open price.

Aggregating our data to monthly frequency, we sort the stocks by their pre-shock fraction of close-price volume executed by the bank. The treated group consists of stocks that rank in the top 50%. The control group consists of the remaining stocks. We find that relative to the control group, the fee cut reduced the treated stocks' fraction of close-price volumes executed by the bank, and decreased the informativeness of their close prices. These results are robust to using ETF/index fund ownership to designate treatment group, excluding earnings announcement days from the data, and using a matched control group based on market cap, trading volume and intraday volatility.

In our model, there are two trading venues accepting only market orders: a market maker (representing the close auction) and a bank (representing the "guaranteed close" service). There is a single asset with an uncertain liquidation value v. There is a continuum of liquidity traders and a continuum of informed traders, all with unit demand. Each informed trader receives a signal about the asset value and buys or sells one unit of the asset if the expected profit is higher than an opportunity cost that is heterogeneous among traders. Each liquidity trader has to buy or sell one unit of the asset after receiving a liquidity shock. Trading with the bank incurs a convenience cost, which represents the cost of building connections and contracting with the bank. Convenience costs are heterogeneous

[5]The same measure is used by SEC DERA (2017), and confirmed to be constituted almost surely by off-exchange market-on-close trades.

[6]This echoes anecdotal evidence that "guaranteed close" is used by index-fund managers including Vanguard Group and BlackRock Inc. (WSJ, 2018)

among traders. The distributions of the above features of the traders are publicly known.

A competitive market maker in the close auction sets the close price p. A trader can submit his order either to the market maker or the bank, both executing his order at p, but with different transaction costs. The bank, besides executing orders for the traders, submits orders on its own account to the close auction to maximize profit. The close auction charges a transaction fee μ_c and the bank charges μ_b per unit of asset traded. μ_c is exogenously given and μ_b is chosen by the bank to maximize profit. The fees are public information ex ante.

We solve the model in closed form. In equilibrium, the net total orders from informed traders is linear in the asset's value v. We find a linear Nash equilibrium where the bank's trading strategy is a linear function of the net total orders it receives, and the market maker sets p as a linear function of the net total orders he receives. We show that under mild conditions, the bank participates with a fee $\mu_b < \mu_c$. We then contrast the informativeness of p in this equilibrium to a model with no bank, and find that informativeness of p is higher with dual trading of the bank. As we described before, this is because the bank can trade profitably based on the net total orders received so long as the proportions of informed orders relative to uninformed orders are different between the two venues. The bank's trading activity amplifies the fraction of informed orders received by the market maker, making the close price more informative. Finally, we find that dual trading of the bank unambiguously decreases the expected gain (before fees) of informed traders, and has a mixed effect on the expected loss (before fees) of liquidity traders.

Let us finally caution that our model is stylized, with exogenous noise trading, and no entry of banks, among other assumptions. A dynamic model with multiple exchanges and broker-dealers would be more realistic extrapolating our result to other trading scenarios in general. In addition, our model is only meant to capture one aspect of "guaranteed close", namely its effect on price discovery. The other side of the coin, namely, the liquidity provision feature, especially whether it still provides liquidity when market conditions are

volatile, is not studied. Future research may shed more light on these issues.

Literature Review Our paper links to three strands of literature. First, we add to the recent discussion on the surge of trading volume at the market close. Although market closures by themselves can generate endogenous time-variation in trading activity and price movements (Hong and Wang, 2000), Bogousslavsky and Muravyev (2021) show that the influx of ETF and index fund trades are key determinants of the volume at market close in recent years, and they adversely affect the informativeness of the close prices. To inform the policy attempt to improve this, our results suggest that off-exchange venues for trading at the market close like "guaranteed close" can actually improve price discovery. This is in contrast to conventional wisdom as expressed in the SEC report SEC DERA (2017) and for example, Credit Suisse (2019).[7] Baldauf et al. (2021) studies optimal contracts referencing unrealized prices with client-dealer agency problem due to potential manipulation of the benchmark, and finds that market-on-close contracts are not optimal in general.

Second, our paper contributes to the literature on dual trading, that is, broker-dealers strategically using customers' order flow information to trade on their own accounts. On the theory side, Röell (1990) has a broker observing only the trades of uninformed trader. In Fishman and Longstaff (1992), the broker has private information about whether his customer is informed or not, and allow the customer and the dual-trading broker to trade at different prices. Sarkar (1995) finds that dual trading has no impact on discovery, although it decreases net profits of informed traders and increases the utility of uninformed traders. Our model is different from theirs in many dimensions; most importantly, we explicitly model the venue choices of both informed and uninformed traders.

Yang and Zhu (2020) and Huddart et al. (2001) study the strategies of informed traders when there are back-runners, who partly infer informed traders' information from their order flow and exploit it in subsequent trading. The informed traders counteract back-

[7] They both suggest off-exchange MOC volume has no effect on price discovery.

runners by randomizing their orders (unless back-runners' signals are too imprecise), but back-runners unambiguously improve price discovery. These analyses are different from ours in terms of the economic questions and modeling approaches. In our model, all the trading happens in one period and the price is set only once, and if traders choose to trade in the close auction their order flow is not being traded against.

To our knowledge, we are the first to provide quasi-experimental empirical evidence on the effect of dual trading on price discovery. This is challenging, not only because of the lack of data on trading/order flow of individual traders and brokers, but more importantly, the lack of exogenous shocks to dual trading activities. Chakravarty and Li (2003) use proprietary audit trail transaction data to study dual trading in futures markets, and find that dual traders trade merely to supply liquidity and manage inventory, rather than trading against the customers. Related to our paper, Barbon et al. (2019) find that broker-dealers intermediating large stock portfolio liquidations spread order flow information to their clients.

Finally, our paper connects to the existing models of dark pools and alternative trading venues. Zhu (2014) shows that adding a dark pool along side the exchange can improve price discovery. This is because dark pools do not guarantee execution, and informed traders tend to trade in the same direction and face a higher execution risk in the dark pool, hence have incentive to remain in the exchange.Ernst et al. (2021) find that market participants learn from the publication of general off-exchange transactions, and the off-exchange orders are information-motivated and contributing to price discovery. Other models on trading venue choice include Hendershott and Mendelson (2000), Buti et al. (2017), and Ye (2010). Their models either do not model asymmetric information, or do not allow all the agents to freely select venues, and do not consider the commission fee difference. Our paper is also related to an emerging literature on exchange competition, among which Chen and Duffie (2021) find that market fragmentation and exchange competition could lead to improvement in price discovery when all exchange prices are taken

together.

The remainder of this paper is organized as follows. In Section 2, we present the institutional background of close auctions on the primary exchanges and "guaranteed close" of the banks. Section 3 exhibits quasi-experimental evidence that "guaranteed close" improves close price informativeness. Section 4 introduces our model of dual trading and price discovery at market close. Section 5 concludes.

1.2 Institutional Background

1.2.1 Close auction

In this section, we introduce the mechanism of NYSE's close auction, whose characteristics are to a large extent shared by other exchanges like NYSE Arca and Nasdaq.[8]

Several types of orders can be used in NYSE's close auction, with the most common being market-on-close (MOC) and limit-on-close (LOC) orders. An MOC order is an unpriced order to buy or sell a security at the close price and is guaranteed to receive an execution. An LOC order sets the maximum price an investor is willing to pay, or the minimum price for which an investor is willing to sell. An LOC order priced better than the final close auction price is guaranteed to receive an execution. As shown in Table A.1 in the appendix, 65% of the orders in NYSE close auction are MOC orders and 14% are LOC orders.[9]

From 6:30 am in the morning, MOC and LOC orders can be entered and existing MOC and LOC orders can be canceled until 3:50 pm. At 4:00 pm, the regular session trading ends and the close auction commences. The method for determining the close prices follows two principles: (1) maximize the number of shares that can be executed

[8] See Appendix A of Bogousslavsky and Muravyev (2021) for a detailed summary of Nasdaq's close auction mechanism, as well as NYSE's. Also, see "NYSE Open and Closing Auctions Fact Sheet", 2019 [link], "NYSE Arca Auctions Brochure", 2019 [link], and "Nasdaq Open Close Quick Guide", 2019 [link]

[9] Another 18% are Closing D Orders, a special order accessible only to NYSE floor brokers.

in the close auction; (2) minimize the difference between the close price and a reference price if multiple close prices satisfy principle (1). The auction effectively aggregates the supply and demand curve constituted by the MOC and LOC orders, and the transaction price and trade volume is determined by the intersection of the two curves.

The Designated Market Makers (DMMs) play an important role in the close auction. They set the closing price at a level that satisfies all interest that is willing to participate at a price better than the close auction price, and supply liquidity as needed to offset any remaining auction imbalances that exist at the closing bell. That means market-on-close orders are guaranteed to be executed.

1.2.2 "Guaranteed close" service

Investment banks such as Goldman Sachs, Morgan Stanley, Credit Suisse Group AG, and UBS Group AG, started a "guaranteed close" service around 2016.[10] Investors looking to buy or sell shares of a stock can get a guarantee from the bank to execute their orders at the close price set on the corresponding primary exchange, where the stock is listed. At 4:00 p.m., the bank pairs the buyers with the sellers of the stock. For the unmatched orders, it can either send them to the exchange or take the other side of the trade itself, storing the extra shares or short interest on its books overnight. As an investor, using "guaranteed close" is equivalent to sending a MOC order to the close auction in execution outcomes, but paying a lower fee. People familiar with the matter told us that Goldman Sachs recently cut the fee charged to broker-dealers to zero (buy-side clients still pay a fee).

Trading alongside the client orders, that is, dual trading, is conducted by the banks. This is one way to cover the bank's cost of providing liquidity, and is documented by

[10]We do not have a comprehensive list of the venues that conduct "guaranteed close". For the top ten alternative trading systems by total volume (during both regular hours and market close), see Table 4 in "Staff Report on Algorithmic Trading in U.S. Capital Markets", 2020, SEC.[link]

the following excerpts from the descriptions of the "guaranteed close" service by Morgan Stanley and Goldman Sachs to their clients.

Morgan Stanley:[11]

> "When we accept an order for execution on a guaranteed benchmark basis (for example, a guaranteed opening, closing, volume weighted average price or other guaranteed transaction), we will typically attempt to offset the risk incurred as a result of such guarantee by transacting in the market on a principal basis, or accessing internal liquidity sources, in the benchmark security or a related instrument, although we may choose not to perfectly hedge our exposure."

Goldman Sachs:[12]

> "We offer client facilitation services, which are typically used by clients to obtain liquidity or a guaranteed execution price. When you use our client facilitation services, we may also effect transactions as agent, as principal (including trading as a market maker or liquidity provider to other clients and trading to manage risks resulting from client facilitation activities), or in a mixed capacity."

Also, Morgan Stanley claims to split the gain from trade between the client and itself:

> "In accordance with market standards and best practices, we strive for allocations between client and principal orders that are fair and equitable.[...]Challenges presented by the current market structure and limitations of certain market centers and trading system may in some cases render a precisely even split impracticable."

[11] See "A Message to Morgan Stanley's U.S. Institutional Equity Division Sales & Trading Clients regarding U.S. Equity Order Handling Practices", 2019, [link]

[12] See "Cash Equities Order Handling Procedures of Goldman Sachs (Asia) L.L.C.", 2019,[link]

Who are the users of the "guaranteed close"? Anecdotal evidence suggests the "guaranteed close" venue is used by index-fund managers including Vanguard Group and BlackRock Inc., as well as some smaller broker-dealers (WSJ (2018)). In general, institutional investors and broker-dealers are primary users of alternative trading systems.[13] In the next section, we present evidence that stocks with higher ETF/index fund ownership are more heavily traded at the "guaranteed close".

"Guaranteed close" is different from dark pools in that it has no execution risk, while in the latter matching depends on the availability of counterparties and some orders from the "heavier" side of the market will fail to be executed (Zhu, 2014). Also, it is also in nature different from the recently approved Cboe Market Close program providing a lower-fee venue to trade at the market close price set by NYSE and Nasdaq. In Cboe Market Close, traders can enter, cancel or replace MOC orders only before 3:35pm. After that, the system would match for execution all buy and sell MOC orders entered with execution priority given based on time-received. But any remaining balance of unmatched shares would be cancelled and returned to the traders. That is, Cboe Market Close only pre-matches some non-price-forming orders.[14]

1.3 Off-Exchange Trading and Close Price Informativeness: Evidence from NYSE Fee Cut

In this section, we empirically study the relationship between the change in off-exchange trading fraction and close price informativeness. We exploit the quasi-experimental setting of the NYSE fee cut in January 2018 and the difference-in-differences method to establish causal evidence.

[13] See Section IV in "Staff Report on Algorithmic Trading in U.S. Capital Markets", 2020, SEC.[link]

[14] See WSJ article "SEC Decision on 4 p.m. Closing Trades Deals Blow to NYSE, Nasdaq " at [link], and SEC Release No. 34-88008 at [link].

1.3.1 NYSE close auction fee cut

In January 2018, the NYSE reduced the close auction fee for market-on-close (MOC) orders, which was seen as an attempt to keep clients from choosing the banks' low-cost "guaranteed close" service.[15] Indeed, the NYSE only reduced the fee for MOC orders but kept the fee for limit-on-close (LOC) orders unchanged. The NYSE's claimed intention is to encourage higher volumes of MOC orders at the close. This fee cut was given with short notice. On December 21, 2017, the NYSE announced the plan of fee changes intended to be effective January 2, 2018. On January 8, 2018, NYSE filed with the SEC about the change in close auction fees.[16]

In NYSE's close auctions, the amount of fee broker-dealers need to pay for the MOC orders depend on the trading volume. Specifically, there are three different tiers. Tier 1 rates would be available for a broker-dealer who in the prior three billing months executed (1) an ADV (average daily volume) of MOC activity on the NYSE of at least 0.45% of NYSE CADV (consolidated average daily volume), (2) an ADV of total close activity (MOC/LOC and executions at the close) on the NYSE of at least 0.7% of NYSE CADV , and (3) whose MOC activity comprised at least 35% of the its total close activity. Tier 2 rates would require a lower ADV of MOC activity and ADV of total close activity. Those who don't meet the requirements for Tier 1 and Tier 2 rates are subject to the Non-Tier rate.

The fee cut reduced the Tier 1 rate and Tier 2 rate significantly, but left the Non-Tier rates almost untouched. Specifically, Tier 1 rate is reduced from \$0.0007 to \$0.0004 (a 42.9% drop). Tier 2 rate is reduced from \$0.0008 to \$0.0005 (a 37.5% drop). But Non-Tier rate is reduced from \$0.0011 only to \$0.0010 (a 9% drop).

Given the different changes, we should expect the fee cut to have differential impact on

[15] Alternatively, the fee cut may be seen as a reaction to the threat of CBOE's entry into the close auction [link], which was awaiting SEC decision back in 2018.

[16] See SEC No. 34-82563 [link]. Also, see NYSE trader update [link].

different market participants. The broker-dealers who qualify for Tier 1 or Tier 2 rates may have a client base that has high demand of trading at the market close price – for example, index funds and ETFs. The fee cut for broker-dealers would pass through to them. Stocks with higher ETF/index fund ownership should be more exposed to the fee cut. Similarly, stocks more heavily traded at the bank's venue would also be more exposed to the fee cut. This is because they have higher ETF/index fund ownership, while ETF/index funds are more sensitive to fees and in the meantime received larger NYSE fee cuts.[17]

We note that on January 2, 2018, Nasdaq also implemented a policy change to attract customers in the close auctions – it increased the rebate to LOC orders.[18] We find that this also induced a flow of trading volume into Nasdaq close auction. Off-exchange volume as a fraction of volume at close price declined for Nasdaq stocks. The price informativeness of stocks mostly affected by the policy change also deteriorated. In the internet appendix, we show this set of results for Nasdaq stocks. (see Table B.5). However, the evidence is not directly related to our story regarding market-on-close orders; we therefore restrict our analysis to NYSE stocks.

1.3.2 Data and descriptive findings

Data construction Our main data source is the NYSE millisecond-level trade and quote data (TAQ) spanning from 2012 to 2019. We also leverage the WRDS Intraday Indicator Dataset (WRDS IID) built from the TAQ data.

The close price, open price, and close auction volume for each stock and each day are from WRDS IID. The close auction volume, measured by TAQ trades with the sale condi-

[17] This is why these stocks are traded more heavily at the bank in the first place

[18] "Limit on Close volume entered between 3:50 p.m. ET and immediately prior to 3:55 p.m. will now be included in the 'Member with imbalance shares executed in all securities in the Opening and Closing Cross through one or more of its NASDAQ Market Center MPIDs that represent more than 0.01% Consolidated Volume during the month' tier providing a $0.0028. The updated tier will read 'Members with shares executed in the Opening and Closing Cross comprised of imbalance orders and/or limit on close orders entered after 3:50pm that represent more than 0.01% Consolidated Volume during the month.' " See http://www.nasdaqtrader.com/TraderNews.aspx?id=ETA2017-220

tion of 6 (Closing Print), occurs on a stock's primary listing exchange (SEC DERA, 2017) – for example, an NYSE-listed stock's close volume is executed in NYSE (NYSE stocks can also be traded at Nasdaq closing crosses.). The market close prices are the recorded transaction prices of these trades. Open auction volume is measured as the volume of trades with the sale condition of O (Market Center Opening Trade). The market open prices are measured by the recorded prices of these trades. For each stock each day, we also calculate from TAQ the volume-weighted average price in the last 5 minutes, $p_t^{close5m}$, in the last 15 minutes, $p_t^{close15m}$, and in the first 5 minutes, p_t^{open5m}, of the regular session. For these calculations, we exclude invalid or erroneous trades that were later canceled or changed.

We proxy the trade volume via banks' "guaranteed close" service (henceforth, the "bank volume") by the trade volume using off-exchange market-on-close (MOC) orders. This is without loss of generality, given that banks anecdotally are the major players in providing the "guaranteed close", and that other broker-dealers accepting MOC orders are no different from the banks in the business format. To measure the off-exchange MOC volumes, we consider all the trades from TAQ that are not cancelled or corrected and occur between 4:00 p.m.–4:10 p.m. EST at the official market close price determined by the close auction. This off-exchange MOC volume has been validated against two regulatory datasets with more detailed information – the FINRA Trade Reporting Facility data, and the FINRA-provided Audit Trail data, to ensure that it is from MOC orders.[19] The two regulatory datasets identify off-exchange executions by venue and trace the executions back to the original orders. Our off-exchange MOC volume would be almost identical if measured by the regulatory datasets.

ETF ownership data is obtained from ETF Global. Index fund ownership data is obtained from the CRSP Mutual Fund database. We closely follow Ben-David et al. (2021)[20]

[19] See SEC report SEC DERA (2017).

[20] We thank the authors for making their code publicly available as supplementary data to the Review of Financial Studies.

and Dannhauser and Pontiff (2019) to identify passive index funds in the CRSP Mutual Fund database. Our sample contains 918 index funds in 2016.

For other variables, we obtain market capitalization and earnings announcement days from CRSP. Measures of volatility, liquidity, and retail and institutional order flows are from WRDS IID. They include the trade-based intraday volatility during market hours, time-weighted percent quoted spread during market hours, total trade volume during market hours, total retail trade volume following Boehmer et al. (2021), total volume of trades $\geq$ 20K in value, and total volume of trades $\geq$ 50K in value. The last two variables are proxies for institutional trades. Lee and Radhakrishna (2000) show 53% of institutional trades are above $20,000 in value; Bhattacharya et al. (2007) and Shanthikumar (2003) use $50,000 dollar value-based cutoff. When trades with value exceeding these cut-offs are interpreted as institutional trades, not surprisingly, we find large institutional presence at banks' "guaranteed close", as large-value trades proliferate, although smaller-value trades also exist.[21]

Price informativeness measure Price informativeness is usually measured by how well a price tracks the fundamental value of an asset. Close price is, however, unique. Being the last price of the day, it aggregates all the information of the day and is generated by auctions with substantial liquidity and turnovers. Prices right before the close are not good measures of the fundamental value since they do not incorporate all the information of the day, and in some cases are prone to manipulation given the lower liquidity. Neither are prices generated by after hours trades since these trades are infrequent and slim. We recognize the next day's open price as the fundamental value of a stock, to which we compare the close price. In our model and most theoretical literature, fundamental value of the stock is realized in a future period at which investors can liquidate the stock. Aligning

[21] Figure A.1 in the appendix plots the distribution of sizes of the off-exchange MOC trades for the stock AAPL. Both extremely large orders and smaller trades exist, and large orders are frequent.

with this, for close price, a reasonable measure for the fundamental value it's reflecting would be the next day's open price.[22]

Therefore, our (inverse) measure of the informativeness of the close price is the mean squared error (MSE) of the close price relative to the next day's open price:

$$\text{MSE} = E\left(\frac{p_{t+1}^{\text{open}} - p_t^{\text{close}}}{p_t^{\text{close}}}\right)^2 \tag{1.1}$$

where p_t^{close} is the close price of day t, p_{t+1}^{open} is the open price of day $t+1$. A lower MSE corresponds to better price informativeness.

Note that we scale the difference by the close price itself. This scaling takes into account that prices at higher levels mechanically vary more, and transforms the "closeness" to the fundamental value into return space. Our results are robust to the scaling factor, where we calculate the MSE's using volume-weighted price in the last 5 minutes (p_t^{last5m}), last 15 minutes ($p_t^{last15m}$) or first 5 minutes (p_{t+1}^{open5m}), as the scaling factor.

Descriptive findings Our sample contains common stocks listed on the major exchanges: NYSE (including NYSE MKT and NYSE Arca), Nasdaq and Bats BZX, with a price greater than \$5 and a market capitalization greater than \$100 million at the end of a month. Daily variables are aggregated to monthly by taking averages. Given the presence of outliers, we winsorize the daily overnight returns at 5% tails before taking the monthly averages.[23] We also winsorize monthly observations of all the variables at 2% tails. Our main results come from the difference-in-differences analysis, which uses the sample of 1,489 NYSE stocks (excluding NYSE MKT, NYSE Arca), spanning from January 2015 to December 2019. Table 1.1 reports the summary statistics of all the variables in this

[22]Indeed, next day's open price, beyond the existing information at the market close, also incorporates overnight information, hence is not a perfect measure for the fundamental value *at* the market close. Nevertheless, our results are robust to excluding data within 1 day from earnings announcements days, where overnight information is strong and potentially causes significant changes in the fundamental value.

[23]Our results are qualitatively similar if we winsorize at, for example, 2% tails, but noisier.

sample.

Who trades at the banks' "guaranteed close"? In line with anecdotal evidence, we show that the off-exchange MOC volume of a stock is closely related to passive ETF/index fund ownership in the cross-section. Figure 1.2 shows the bin-scatter plots of off-exchange MOC volume as a fraction of the close price volume against ETF ownership, and against index fund ownership. The binscatter plots are based on OLS regression, controlling for a battery of confounding factors including log(market cap), log(trade volume), volatility, time fixed effects and other controls including log(total retail volume), log(total volume of trades $\geq$ \$20K in value), log(total volume of trades $\geq$ \$50K in value), after-close volume/total volume, and close auction volume/total volume. The results are robust to not including those controls, but noisier. The bin-scatter plots suggest that as ETF ownership or index fund ownership increases by 1 percent, off-exchange MOC volume as a fraction of the close price volume increases by 0.3 percent.

1.3.3 Difference-in-differences: The effects of NYSE fee cut

We present quasi-experimental evidence that the banks' "guaranteed close" improves the informativeness of the close price. To that goal, we use a standard difference-in-differences strategy to estimate the effect of the NYSE close auction fee cut in January 2018 on the MSE of close prices. As discussed in Section 1.3.1, the fee cut significantly reduced the transaction fee in the NYSE close auction for investors with large trading demand at the market close, like ETFs and index funds. Stocks more heavily traded at the bank's venue would be more exposed to the fee cut. They have higher ETF/index fund ownership, while ETF/index funds are more sensitive to fees and in the meantime received larger NYSE fee cuts.[24]

We drop the NYSE stocks that participated as treated stocks in the 2016 Tick Size

[24]This is why these stocks are traded more heavily at the bank in the first place

Pilot Program. The Pilot increased the tick size for select small stocks, which may force the close price to deviate from the fundamental price. The Pilot ends on September 28, 2018, when the tick size requirements are repealed. This may interfere with the NYSE fee cut quasi-experiment.[25] We rank the remaining stocks by the average fraction of close-price volume (i.e., volume traded at the close price) executed off exchange in 2017. The treated group consists of stocks that rank at top 50% – they are more exposed to the NYSE fee cut. The control group consists of stocks that rank at the bottom 50%. Treated-group stocks have on average a market cap of $22 billion, and ETF/index fund ownership of 8.7% and 8.4%, respectively. Control-group stocks have on average a market cap of $7.1 billion, and ETF/index fund ownership of 4.8% and 3.8%, respectively.

Indeed, right after the NYSE fee cut, the treated stocks experienced a drop in off-exchange MOC volume while the stocks in the control group continued the trend before the fee cut. This is salient in panel (a) of Figure 1.3, where we plot the raw means of the fraction of close price volume executed off exchange within the treated and control group. Off-exchange MOC volume increased more rapidly in the treated group prior to the fee cut. But almost immediately after the fee cut, the off-exchange MOC volume of the treated group stopped growing and started declining, while continued to grow at a rate similar as before for the control group. This effect is more salient as we control for a battery of confounding variables including volatility, stock fixed effects, time fixed effects, etc. (See Figure B.1 in the internet appendix.)

The major difficulty in difference-in-differences analyses involves separating out pre-existing trends from the dynamic effects of a policy shock. To avoid confounding the two, we first test for pre-existing trends in the MSE measure of price informativeness.

[25] In fact, our results are robust to including those stocks.

Specifically, we run a regression of the following form.

$$\text{MSE}_{i,t} = \alpha_i + \lambda_t + \sum_k \beta_k \text{Treat}_i \cdot \mathbb{I}_{t=2017m11+k} + \Gamma X_{i,t} + \epsilon_{i,t}. \tag{1.2}$$

where Treat_i is a dummy variable that takes the value one if stock i is in the treated group, and takes the value zero otherwise. $\mathbb{I}_{t=2017m11+k}$ is a dummy variable that takes the value one if time t is k months after 2017m11, and takes the value zero otherwise. Stock fixed effects, time fixed effects, log(market cap), log(total trade volume), volatility, and other control variables, including log(total retail volume), log(total volume of trades $\geq$ \$20K in value), log(total volume of trades $\geq$ \$50K in value), after-close volume/total volume, and close auction volume/total volume, are controlled in this regression.

Panel (b) of Figure 1.3 plots the coefficients β_k and the corresponding confidence intervals. We have three findings. First, the MSEs of the treated and control group moved almost perfectly in tandem from 20 months before the NYSE fee cut, so the parallel trend assumption appears to hold well. Second, from September 2015 to March 2016, there appeared to be a drop in MSE for the treated stocks. This same time period was accompanied by the start of the ramp-up of off-exchange MOC volume of treated stocks relative to control stocks. This seems consistent with that off-exchange MOC activity improves close price informativeness, although a causal interpretation is unwarranted since low MSE stocks may select into the treated group. Third, the MSE of treated stocks increased substantially and remained differentially higher than the control stocks after the fee cut.

Finally, we run the following standard difference-in-difference regression.

$$\text{MSE}_{i,t} = \alpha_i + \lambda_t + \beta \text{Treat}_i \cdot \mathbb{I}_{t\geq 2018m1} + \Gamma X_{i,t} + \epsilon_{i,t} \tag{1.3}$$

where $\mathbb{I}_{t\geq 2018m1}$ is a dummy variable that takes the value one if time t is after January

2018, and all the other variables are as defined before in Equation (1.2).

The coefficient of interest is β, which measures the differential change in MSE for the treated stocks and control stocks, holding constant stock-level time-varying characteristics, as well as stock and time fixed effects. Besides MSE (close price as the scaling factor), we also used as a dependent variable the MSEs using volume-weighted price in the last 5 minutes (p_t^{last5m}), last 15 minutes ($p_t^{last15m}$) or first 5 minutes (p_{t+1}^{open5m}) as the scaling factor, as well as the median of $(\frac{p_{t+1}^{open}-p_t^{close}}{p_t^{close}})^2$ (median SE). We also adopt two placebo variables as dependent variables. They are intraday volatility during market hours, and quoted bid-ask spread during market hours. Both of them are important measures of market conditions for a stock, but given that they are measured during market hours, they should not be affected by the NYSE fee cut. To account for serial correlation and stock-specific random shocks, we cluster standard errors at the stock level in all specifications.

Table 1.2 shows the difference-in-difference regression results. We see that the fee cut increased significantly the MSE measures and the median SE of treated stocks, meaning the informativeness of treated group stocks relative to the control group have been worsened by the shock. Not surprisingly, the fee cut did not seem to affect the intraday volatility and liquidity of the treated stocks.

1.3.4 Robustness checks

We take several additional steps to ensure the validity of our research design and the robustness of our estimates.

Alternative treatment designation One potential concern with the difference-in-difference results is about the designation of the treatment group. We argued that stocks more heavily traded ex-ante at the bank's venue would be more exposed to the fee cut, since they have higher ETF/index fund ownership, while ETF/index funds are more sensitive to fees and in the meantime received larger NYSE fee cuts. We now verify this idea by defining the

treatment group with ownership by ETFs and index funds. Specifically, we classify stocks ranked in the top 50% in the sum of ETF and index-fund ownership into the treated group, and the remaining stocks into the control group. Table B.1 in the internet appendix gives the estimation results which are all consistent with the main results.[26]

Excluding earnings announcement days We take the next day's open price as the fundamental value to which we compare the close price. Next day's open price, beyond the existing information at the market close, also incorporates overnight information, hence is not the fundamental value *at* the market close. For most stocks, earnings announcements are the major overnight information that affects an individual stock's price. One potential concern is that treated stocks may experience stronger earnings news after 2018. To address this, we exclude the data within 1 day, that is, the [-1,0,1] days, from earnings announcements days, before conducting the difference-in-difference estimation. As shown in Table B.2 in the internet appendix, our results are robust to excluding these data.

Matching specification Another concern regarding our estimates is that stocks in the treated and control group might be very different in size and many other characteristics, although they are already all NYSE stocks. To alleviate this concern, we conduct a matched sample approach. We match the stocks based on pre-treatment values of log(market cap), log(trading volume), and intraday volatility. For each stock, the closest matching control stock is chosen (with replacement) according to the Mahalanobis distance of the three variables, to constitute the matched control group. Table B.3 in the internet appendix shows that the matched sample yields quantitatively similar results compared to the nonmatched sample. Table B.4 in the internet appendix shows a set of balance test results for the nonmatched samples and the matched samples.

[26]Quoted spreads of treated stocks are estimated to increased by the fee cut, but not very statistically significant.

1.4 A Model of Dual Trading and Price Discovery at Market Close

In this section, we present a simple model of trading at the market close where traders can trade either on-exchange with the market maker, or off-exchange via a bank. We solve the equilibrium strategies of the market participants and show that the bank's trading activity based on orders received from traders can improve the informativeness of the close price. This provides an explanation of our empirical findings that off-exchange trading improves close price informativeness.

Our result is mainly driven by two differences from the previous literature on dual trading (e.g. Sarkar (1995)). First, there are two trading venues and the proportion of informed orders relative to uninformed orders can be different across the two venues. Second, we model small informed traders with unit demand in a stylized setting following Kyle (1985). The first allows the net total orders received by the bank to contain information that is not priced without dual trading, providing an incentive for the bank to trade even if it does not observe the identity of each trader. The bank trades in the same direction as the net orders received when it expects to have a higher proportion of informed orders compared to the other trading venue, or in the opposite direction to it when its proportion of informed orders is expected to be lower. Such trading activity, without reaction of other traders, increases the number of informed orders relative to uninformed orders on the market. The second ensures that such increase is not fully offset by strategic reactions of other traders. Small informed traders who are budget-constrained and infinitely small compared to the market do not react to changes in market depth. As long as there are some small informed traders on the market, orders by informed traders are no longer proportional to market depth. Then the effect of dual trading cannot be fully offset by strategic reaction of other traders.

In our baseline model, we assume that all traders have unit demand for simplicity.[27] The fractions of traders choosing to trade with the market maker and the bank are determined in a largely exogenous way – these fractions depend only on the transaction fees and convenience costs incurred when traders trade with the bank. This assumption not only simplifies our analysis, but also makes our result more general, that is, it works for any arbitrary sorting of traders between two venues. We generate closed-form solutions to the model and show that such a simple model already has interesting implications. We discuss the robustness of our model to our simplifying assumptions and its applications to other scenarios beyond trading at the market close at the end of this section.

1.4.1 Markets and traders

There are three periods in the model, denoted by $t = 0, 1, 2$. There is a single asset with an uncertain liquidation value $v \sim N(0, \sigma^2)$. The asset value v is publicly revealed at $t = 2$.

There are four types of risk-neutral market participants: a market maker, a bank, a continuum of infinitely many liquidity traders, and a continuum of infinitely many informed traders. The market maker determines the close price p that equals her expected asset value. Both the market maker and the bank provide services that execute orders at the close price for the traders, while the bank also submits orders to the market maker itself. Each trader can buy or sell one unit of the asset if he participates. There are convenience costs for traders to trade with the bank, which represent the cost of building connections and contracting with the bank. Convenience costs are heterogeneous across traders, following twice-differentiable cumulative distribution functions $G_L : [0, \Gamma_L) \to [0, 1]$ and $G_I : [0, \Gamma_I) \to [0, 1]$ for liquidity traders and informed traders respectively. Each informed trader has an alternative trading opportunity $\chi_i \geq 0$, which can be interpreted as a signal for another asset. Convenience costs and alternative trading opportunities are inde-

[27] The assumption of unit demand is frequently observed in models involving venue choice, for example, Zhu (2014); Hendershott and Mendelson (2000)

pendently distributed among informed traders. Traders observe their private features. The distributions of these features are public information.

A trader can submit his order either to the market maker or the bank, both executing his order at the same close price p announced at $t = 2$ by the market maker. The market maker charges a transaction fee μ_c and the bank charges μ_b per unit of asset traded. μ_c is exogenously given and μ_b is chosen by the bank to maximize profit, announced at the beginning of $t = 0$. Traders make their trading decisions after observing the fees. The bank, besides executing orders for the traders, also takes its own trading position and submits orders to the market maker containing the order imbalance received from traders and its own position. All market participants only use market orders.

The figure below shows the timeline of the model. At the beginning of $t = 0$, the bank announces fee μ_b. Some liquidity traders receive liquidity shocks such that L^+ liquidity traders have to buy one unit of the asset, and L^- liquidity traders have to sell one unit of the asset.[28] Assume that the net total liquidity orders $L = L^+ - L^-$ is random and follows a normal distribution $N(0, \sigma_L^2)$. Random variables v, L are independent and whether a liquidity trader receives a positive or negative shock is independent of her convenience cost for trading via the bank. Each informed trader receives a signal about the value of v that is uniformly distributed $s_i \sim U(v - \sigma_s, v + \sigma_s)$. They decide to buy or sell one unit of the asset if the expected profit is higher than the opportunity cost.[29] The assumption of unit demand can reflect capital constraints that limit the informed trader's maximal trade size.[30]

[28] We suppose it is very costly for a liquidity trader not to meet her liquidity need, so that she always trade when receiving a shock. The only decision left to be made is the venue choice.

[29] The assumption of unit demand is frequently observed in models involving venue choice, for example, Zhu (2014); Hendershott and Mendelson (2000)

[30] A helpful way to interpret informed traders with unit demand and opportunity costs here is to regard them as risk-neutral capital constrained investors who invest all cash in the most profitable opportunity.

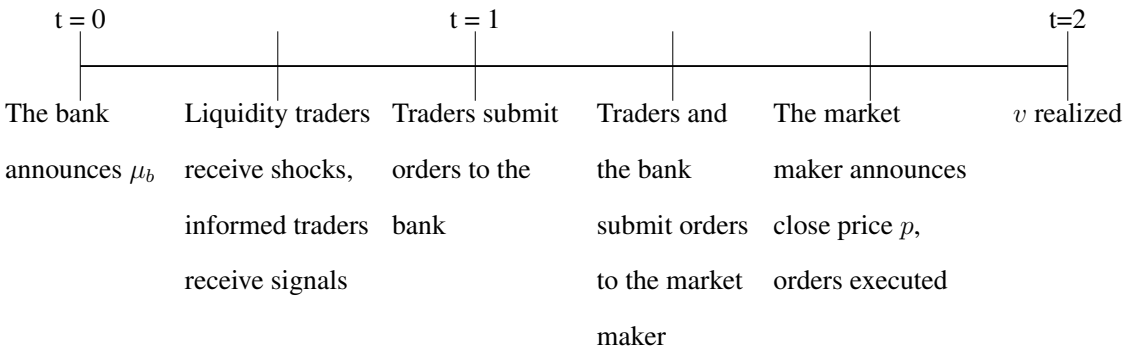


Figure: Model Time Line

At the beginning of $t = 1$, traders choose where to trade and submit their orders. After receiving orders from the traders, the bank can take its own position to trade along or against the net orders received. The bank cannot observe whether an order is informed or not, but knows the composition of its customers given the distributions of traders' features. The bank matches buy and sell orders, including its own position, and submits the remaining order imbalance to the market maker.

At the end of $t = 1$, the close trading ends. The market maker does not observe whether an order is from an informed trader, a liquidity trader or the bank. After collecting all the orders from traders and the bank, the market maker announces the close price p, which equals the expected asset value conditional on the net orders she received. All orders are then executed at this close price p. Then at $t = 2$, the asset's value v is realized and all participants receive their payoffs.

1.4.2 Equilibrium

An equilibrium consists of the quoting strategy of the market maker, the fee charged by the bank, and the trading strategies of the traders and the bank. In equilibrium, the market maker breaks even in expectation. All traders and the bank maximize their expected profits. We solve for the equilibrium backwards along the timeline, starting from trading strategies at $t = 1$ given the bank's fee μ_b announced at $t = 0$.

Traders' venue choices We first characterize the venue choices of the traders. A trader's expected payoff from trading $x_i = \pm 1$ unit of the asset with the market maker is

$$E_i[v - p]x_i - \mu_c \tag{1.4}$$

Likewise, the trader's expected payoff from trading via the bank is

$$E_i[v - p]x_i - (\mu_b + \gamma_i) \tag{1.5}$$

Traders choose the trading venue with higher expected payoff, or equivalently, lower total trading cost. Hence a trader trades via bank if

$$\gamma_i < \mu_c - \mu_b \tag{1.6}$$

and trades with the market maker otherwise. Let α_b be the equilibrium fraction of liquidity traders who choose to trade via the bank. The remaining fraction $1-\alpha_b$ of liquidity traders send their orders to the market maker. Similarly, let θ_b be the fraction of informed traders who trade via the bank, and the remaining fraction $1 - \theta_b$ of informed traders send their orders to the market maker. Then we have $\alpha_b = G_L(\mu_c - \mu_b)$, and $\theta_b = G_I(\mu_c - \mu_b)$.

Let L_b, L_c be the net total orders submitted to the bank and the market maker by the liquidity traders. Assume that L_b, L_c are normally distributed and independent, that is, the net liquidity orders received by the bank and the market maker are independent of each other, and the variance of the net liquidity orders is linear in the fraction of traders who

trade via that venue.[31] Then, since $L \sim N(0, \sigma_L^2)$, we have

$$L_b \sim N(0, \alpha_b \sigma_L^2)$$
$$L_c \sim N\left(0, (1-\alpha_b)\sigma_L^2\right) \tag{1.7}$$

An informed trader trades if the expected profit of trading the asset is higher than that of the alternative trading opportunity and the transaction cost, that is,

$$sign(s_i)E[v-p|s_i] > \max[E[v_i - p_i|\chi_i], \min(\mu_c, \mu_b + \gamma_i)] \tag{1.8}$$

where $E[v_i - p_i|\chi_i]$ is the expected profit of the alternative trading opportunity given χ_i. Assume that $E[v_i - p_i|\chi_i] = E[v-p|\chi_i]$. Guess that $sign(s_i)E[v-p|s_i]$ is strictly increasing in $|s_i|$,[32] then the informed trader trades the asset rather than the alternative opportunity when $|s_i| > \chi_i$. Then we can solve for the net total informed orders for any asset value v and signals $s_i \sim U(v-\sigma_s, v+\sigma_s)$ in the following proposition. All proofs (for other propositions as well) are in the appendix.

Proposition 1. *Let $h(\chi)$ be the mass of informed traders with alternative trading opportunity χ. If $\sigma_s > \mu_c$ and*

$$h(s) = 0 \quad \textit{when} \quad sign(s)E[v-p|s] < \mu_c \tag{1.9}$$
$$h(|s| + \sigma_s) = h(|s| - \sigma_s) \quad \forall |s| \geq \sigma_s \tag{1.10}$$

then there exists a constant a such that the net total informed orders is

$$I = av \tag{1.11}$$

[31] As a robustness check, we find that our main results still hold when L_b and L_c are perfectly correlated.
[32] This is verified by the linear equilibrium we solved in Proposition 2.

The assumptions on alternative trading opportunities here are made for tractability. Due to transaction fee, the fraction of informed traders that trade only increase every time after a certain size of increment in asset value. In general, the aggregate net informed orders I is weakly increasing in the asset value v regardless of the distribution of χ_i. When the sufficient conditions above hold, the net total informed orders is linear in the asset value v. The following lemma shows a simple example of a distribution of χ_i such that the conditions above hold.

Corollary 1. *Assume that there is mass $2\sigma_s a$ of informed traders with alternative trading opportunity $\chi_i = (1+2i)\sigma_s$ for each $i \in \mathbb{N}$, and $\sigma_s > \mu_c$. Then the net total informed orders is*

$$I = av \tag{1.12}$$

Then given the fractions of informed traders trading via each venue shown above, the net total informed orders submitted to the bank (I_b) and the market maker(I_c) are

$$\begin{aligned} I_b &= \theta_b a v \\ I_c &= (1-\theta_b) a v \end{aligned} \tag{1.13}$$

Bank's trading strategy and the close price Given the liquidity and informed orders submitted to each trading venue, we now characterize the bank's trading strategy and the market maker's quoting strategy. The net total orders received by the bank is $u = L_b + \theta_b a v$, with variance $\sigma_u^2 = \alpha_b \sigma_L^2 + \theta_b^2 a^2 \sigma^2$. We focus on the case when $\sigma_u^2 > 0$ in the following discussion, that is, the bank receives a positive measure of orders. The bank takes its own position $X(u)$ based on its expectations given the net orders received and the distributions of traders' features. $X(u)$ can be either along or against the direction of the

net orders the bank receives. The bank's expected trading profit is

$$E[v-p|u]X(u) \tag{1.14}$$

Here we assume that the bank's position $X(u)$ does not incur changes in the transaction fee paid to the market maker to preserve tractability of the model.[33]

Following Kyle (1985), we look for a linear equilibrium of the bank's trading strategy and the close price quoted by the market maker. Guess that the market maker applies a linear price setting rule $p = \lambda y$. Then we can write the bank's trading problem as

$$\max_X E[v-p|u]X = \left(\left[(1-\lambda a)\frac{\theta_b a \sigma^2}{\sigma_u^2} - \lambda \frac{\alpha_b \sigma_L^2}{\sigma_u^2} \right] u - \lambda X \right) X \tag{1.15}$$

By taking the first order condition and combining it with the market maker's actual price setting rule $p = E[v|y]$, where $y = L + av + X(u)$ is the net orders received by the market maker, we get a linear equilibrium as in the following.

Proposition 2. *A linear Nash equilibrium of the model described above is given by*

$$p(y) = \lambda y \tag{1.16}$$

$$X(u) = Ku \tag{1.17}$$

where $\lambda = \frac{(1+K\theta_b)a\sigma^2}{((1+K)^2\alpha_b+(1-\alpha_b))\sigma_L^2+(1+K\theta_b)^2a^2\sigma^2}$, $K = \frac{-B+\sqrt{B^2-4AC}}{2A}$, *with* $A = \sigma_u^2, B = \theta_b a^2\sigma^2 + (2-\theta_b)\frac{\alpha_b}{\theta_b}\sigma_L^2, C = \left(\frac{\alpha_b}{\theta_b} - 1\right)\sigma_L^2$ *when* $\theta_b > 0$, *and* $K = -\frac{1}{2}$ *when* $\theta_b = 0, \alpha_b > 0$.[34]

In this equilibrium, the trading strategy of the bank and the quoting strategy of the market maker are both linear in the net total orders they receive, and determined only by

[33] In reality, it can save transaction fees if the position is against the net order imbalance, or incur additional transaction fees if the position is in the same direction as the net order imbalance.

[34] When $\alpha_b = \theta_b = 0$, we get the trivial equilibrium with no bank where $\lambda = \frac{a\sigma^2}{\sigma_L^2+a^2\sigma^2}$ and K is not defined.

the exogenous parameters and the bank's fee announced at $t = 0$. Some interesting results about the bank's trading strategy are worthwhile to note here. First, K has the same sign as $\theta_b - \alpha_b$. When $\theta_b > \alpha_b$, that is, the fraction of informed traders trading via the bank exceeds that of the liquidity traders, then $K > 0$ — the bank takes its own position in the same direction as the net orders it receives. When $\theta_b < \alpha_b$, that is, the fraction of liquidity traders trading via the bank exceeds that of the informed traders, then $K > 0$ — the bank takes its position in the opposite direction to the net orders it receives. This is because if $\theta_b > \alpha_b$, the bank perceives the orders received to be more informative than orders on the whole market, so following them is profitable. If $\theta_b < \alpha_b$, the bank perceives the orders received to be more noisy than orders on the whole market, so trading against them is profitable due to the price impact of liquidity traders. Second, when $\theta_b = 0$, the bank's trading strategy $K = \frac{1}{2}u$ coincides with that of the dual trading broker in Röell (1990). In this case, the bank only receives liquidity orders, just like in Röell (1990), and makes a profit by providing liquidity to them. Third, when $\alpha_b = 0$, $X = \frac{1}{2}\left(\sqrt{a^2 + 4\frac{\sigma_L^2}{\sigma^2}} - a\right)v$, the bank's trading position does not depend on θ as long as it is positive. This is because when the bank only receives informed orders, it can perfectly infer the value of the asset from just a small mass of informed orders. Then it behaves in the same way as a large informed trader who perfectly observes the asset value.

Fee setting of the bank Given the equilibrium trading strategies of the market participants at $t = 1$ solved above, now we solve for the bank's fee setting problem at $t = 0$. The bank chooses a fee μ_b that maximizes the expected total profit

$$E\left[E[v - p|u]X(u)|\mu_b\right] + E\left[|L_b^+| + |L_b^-| + |\theta_b av||\mu_b\right]\mu_b - E\left[|L_b + \theta_b av||\mu_b\right]\mu_c \tag{1.18}$$

Assume that $E[|L_b^+| + |L_b^-|] = \eta E[|L_b|]$, $\eta \geq 1$, that is, a constant proportion of liquidity orders submitted to the bank are expected to be matched between liquidity buyers and sellers. Substituting previous results for the expected trading profit, we can rewrite the problem as

$$\max_{\mu_b} \lambda(\mu_b) K^2(\mu_b) \sigma_u(\mu_b)^2 + \sqrt{\frac{2}{\pi}} \left(\eta \sqrt{\alpha_b(\mu_b)} \sigma_L + \theta_b(\mu_b) a \sigma \right) \mu_b - \sqrt{\frac{2}{\pi}} \sigma_u(\mu_b) \mu_c \tag{1.19}$$

The problem cannot be solved in closed form generally. However, in order to get the equilibrium results solved above, it is enough to show that the bank does participate with a lower fee than the market maker's in equilibrium, which is shown below.

Proposition 3. *There exists an optimal fee for the bank, $\mu_b \in [0, \mu_c]$. If $\eta > 1$ and $G_L'(0) > 0$, then bank's optimal fee $\mu_b < \mu_c$.*

When the above sufficient conditions are satisfied, the bank can at least make a profit by setting a fee slightly lower than μ_c and match some liquidity orders. Then the bank participates with a lower fee than the market maker's and the equilibrium holds.

1.4.3 Price informativeness

The MSE measure of price informativeness Now we measure the informativeness of the close price p, in terms of how well it reveals the fundamental asset value v. We define that the close price is more informative when it is "closer" to v, and we use the scaled mean square error to measure this closeness.

We calculate the MSE of p and scale it by the variance of the asset value σ^2 to get the scaled MSE

$$\frac{MSE}{\sigma^2} = \frac{E\left[E[(v-p)^2|v]\right]}{\sigma^2} = \frac{1}{\xi + 1} \tag{1.20}$$

where $\xi = \frac{(1+K\theta_b)^2 a^2 \sigma^2}{[(1+K)^2\alpha_b + 1 - \alpha_b]\sigma_L^2}$.

Before discussing the effect of dual trading by the bank on price informativeness, it is helpful to note the following result.

Lemma 1. *The variance of informed orders increases more than the variance of liquidity orders due to dual trading of the bank, that is,*

$$(1 + K\theta_b)^2 \geq (1 + K)^2\alpha_b + 1 - \alpha_b \tag{1.21}$$

Improvement in price informativeness In order to analyze the effect of the bank's guaranteed close service on close price informativeness, we compare the scaled MSE in the equilibrium where bank conducts the guaranteed close service, with the scaled MSE in the equilibrium without the bank. This comparison shows that having the bank conducting the service improves close price informativeness.

Proposition 4. *Having the bank conducting guaranteed close service improves the informativeness of* p.

1.4.4 Changes in the profits of the traders

Now we discuss the effect of dual trading by the bank on the profits of the traders, that is, the expected gains of the informed traders and the expected losses of the liquidity traders. The expected trading profit before fees of an informed trader is

$$\Pi_I = E[E[|v - p||v]] \tag{1.22}$$

$$= \sqrt{\frac{2}{\pi}}\sigma \frac{\left((1 + K)^2\alpha_b + (1 - \alpha_b)\right)\sigma_L^2}{\left((1 + K)^2\alpha_b + (1 - \alpha_b)\right)\sigma_L^2 + (1 + K\theta_b)^2 a^2 \sigma^2} \tag{1.23}$$

The expected trading profit before fees of liquidity traders who trade directly with the

exchange is

$$\Pi_{L,c} = E[E[|v-p||L_c]] \tag{1.24}$$

$$= -\sqrt{\frac{2}{\pi}}(1-\alpha_b)\sigma_L^2 \frac{(1+K\theta_b)a\sigma^2}{((1+K)^2\alpha_b + (1-\alpha_b))\,\sigma_L^2 + (1+K\theta_b)^2 a^2\sigma^2} \tag{1.25}$$

The expected trading profit before fees of liquidity traders who trade with the bank is

$$\Pi_{L,b} = E[E[|v-p||L_b]] \tag{1.26}$$

$$= -\sqrt{\frac{2}{\pi}}\alpha_b\sigma_L^2 \frac{(1+K)(1+K\theta_b)a\sigma^2}{((1+K)^2\alpha_b + (1-\alpha_b))\,\sigma_L^2 + (1+K\theta_b)^2 a^2\sigma^2} \tag{1.27}$$

Note again that the equilibrium results with no bank coincides with the results when we impose $K = 0$, that is, the bank do not trade. Then we compare the profits of the traders with and without the bank and get the following results.

Proposition 5. *The introduction of the bank decreases the expected gain before fees of informed traders.*

It is not surprising that the expected gain of the informed traders decrease given our previous result that price informativeness improved. If the bank receives a larger fraction of informed orders than uninformed orders, it chooses $K > 0$ and competes with the informed traders, reducing their profits. If the bank receives a smaller fraction of informed orders than uninformed orders, it chooses $K < 0$ and reduces market depth, still hurting the profits of the informed even if they do not trade with the bank.

Proposition 6. *The introduction of the bank decreases the expected loss before fees of liquidity traders who still directly trade with the exchange if and only if* $\theta_b > \alpha_b$ *and* $(\theta_b - 3\alpha_b)(\theta_b - 2\alpha_b + \frac{\alpha_b}{\theta_b})\sigma_L^2 < [3\theta_b(\theta_b - \alpha_b) + 4\alpha_b(1-\theta_b)]\,\sigma^2$*. It decreases the expected loss before fees of liquidity traders who trade with the bank if and only if* $\theta_b < \alpha_b$*.*

The welfare effects on the liquidity traders are mixed. For a liquidity trader, the welfare effect largely depends on whether she trades with the bank and whether the bank receives a larger fraction of liquidity orders than informed orders. A liquidity trader benefit from the bank's business if she trades with the bank and it provides liquidity to trades with $K < 0$, or if she does not trade with the bank and it amplifies orders received with $K > 0$. Notice that such difference does not affect the venue choices of the liquidity traders since they are small and price taking. The results are calculated for all liquidity traders who trade with the bank and all who do not, instead of each individual trader, and they are not able to coordinate.

1.4.5 Discussion

Lastly, we discuss the robustness of our results to assumption changes and the generality of our model implications on scenarios beyond trading at the market close. Our result on price informativeness is driven by two forces: dual trading increases informed orders relative to uninformed orders, and such direct effect is not fully offset by reactions of other traders. The former naturally holds for dual trading to be profitable. The later holds when some informed traders are small and do not react to changes in market depth. In the following, we discuss what happens if some specific changes are made.

Large traders For simplicity, all traders in our model above are assumed to be small. Here we discuss how large traders can interact with our model mechanism. If large traders are included, the effect on informativeness depends on assumptions on venues choices of the large trader. Suppose the motivations of the large traders are observed by the bank.[35] Then concerns on price impact drive large informed traders to trade with the market maker and large uninformed traders to trade with the bank, because the bank's trading increases

[35]Previous literature on dual trading usually assume brokers have information about the motivations of the traders, either perfect (e.g., Röell (1990), Sarkar (1995)), or imperfect Fishman and Longstaff (1992))

the price impact of the former and reduces that of the later. In this case, the bank trades against half of the large liquidity traders' orders besides trading on the total orders of the small traders as above. Since there are some small informed traders who do not react to changes in market depth, the reaction of the large informed traders can only partly offset effects of the banks' orders.[36]

The results can be different if the motivations of the large traders are unobserved or if large informed traders have to trade with a dual trading broker as in Sarkar (1995). When a large informed traders' motivation is not observed, he trades more with the bank when $K < 0$ and less when $K > 0$, so the bank does not trade in equilibrium. When a large informed trader has to trade with a dual trader, their interaction actually reduces price informativeness if there are also small informed traders in the model, since both reduce their informed trading due to the presence of small informed traders.

Endogenous venue choices While the venue choices are modeled in a largely exogenous way in our model, this assumption actually makes our result more general, since our discussion includes any arbitrary sorting of traders between the two venues. Any specific setting that models the venue choices endogenously would lead to a special case of our model. A necessary and sufficient condition for price informativeness to improve in our model is that the fraction of informed and uninformed traders trading with the bank is different from that of the whole market. This condition is satisfied as long as there are some differences between the two venues that have different effects on informed and uninformed traders. For example, if we apply our model to a regular darkpool like in Zhu (2014), execution risks make liquidity traders more likely to trade in the darkpool. Some brokers in reality offer different services to sort uninformed and informed traders intentionally.

[36]For example, in a model with a large informed trader, a large uninformed trader, and some small informed traders without access to the bank, the price informativeness is decreasing in market depth and dual trading of the bank decreases market depth and improves price informativeness. This is equivalent to including some small informed traders to the one period model in Kyle (1985) and let the bank reduce the variance of noise orders.

Application to other scenarios Our model is built on the stylized setting following the one period model in Kyle (1985) and can be applied generally beyond just trading at the market close. In order to focus on the role of the bank to explain our empirical findings, we assumed in our model that only the bank conducts dual trading and traders can directly trade with the market maker. However, our result on price informativeness still hold if everyone trades via brokers, as long as there is heterogeneity among the brokers so the fractions of informed traders relative to uninformed traders trading with them are different. We do not discuss continuous trading with multiple prices here due to the feature of trading at the market close, but the mechanism of out model can be general since the main driving forces, that is, heterogeneity in trading venues and small informed traders, are not specific to trading at the market close.

1.5 Conclusion

In this paper, we formally study the impact of "guaranteed close" on the informativeness of the close price, both empirically and theoretically. Using the NYSE close auction fee cut in January 2018 as a policy shock, we provide quasi-experimental evidence that "guaranteed close" improves price discovery. Our unique empirical setting is ideal for testing the effect of dual trading: banks trade on their own accounts after viewing order flows of customers, and a quasi-experimental shock reduces the order flow to the banks; designated market makers clear the market in a single auction and set the price – these features correspond to the framework that the dual-trading literature usually builds on. Our empirical finding cannot be explained by the predictions of the dual-trading literature.

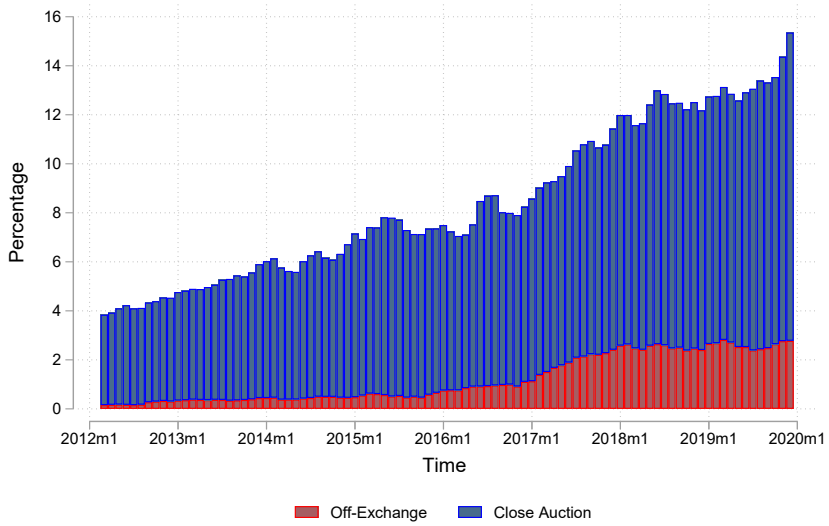
We build a model and provide a novel mechanism through which dual-trading improves price discovery. In our model, traders explicitly choose between two trading venues (i.e., the bank and the close auction) based on transaction costs. As long as the proportions of informed orders relative to uninformed orders are different between two venues, the net

total orders to the bank provide the bank with information to trade profitably. Such trading activity alone amplifies the fraction of informed orders received by the market maker. Unlike previous literature on dual trading that usually assumes all informed traders to be large and sensitive to price impact, our model introduces small informed traders with unit demand who do not react to changes in market depth. Hence the amplification effect of the bank on the fraction of informed orders cannot be fully offset by the reaction of informed traders reducing their orders.

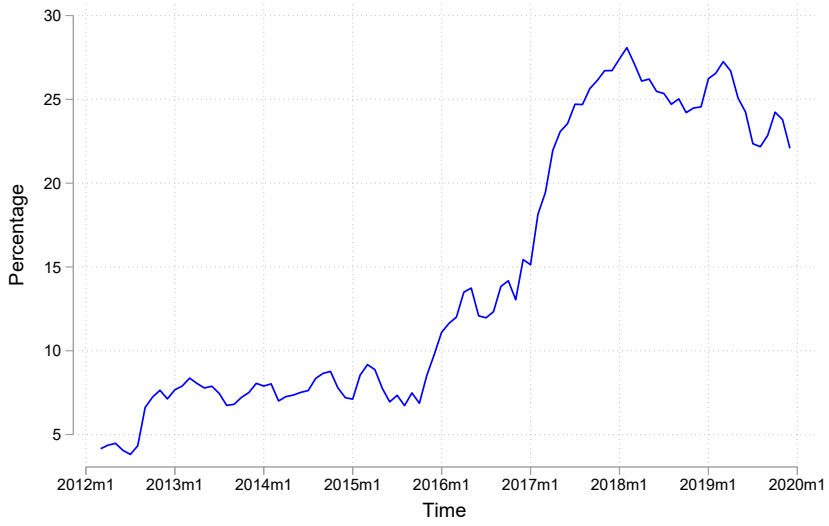
The US brokerage industry has been undergoing rapid changes in its landscapes nowadays. While some brokers still charge commission fee and offer explicit execution price, e-brokers like Robinhood flock to charge zero commission fee and sell order flow data to sophisticated investment firms. The former brokers would be comparable to the close auction in our paper, and the latter the bank's "guaranteed close". Our model and empirical evidence shed light on the effect of these changes on price discovery. But of course, the jury is still out when it comes to the impact of these changes on liquidity, fairness, and aggregate investor welfare.

Figure 1.1: Off-exchange MOC volume and close auction volume of S&P 500 stocks

(a) Off-exchange MOC volume and close auction volume



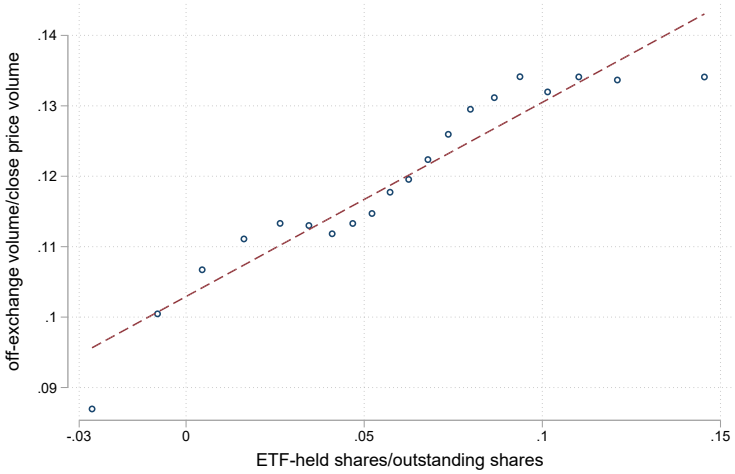
(b) Off-exchange MOC volume/close auction volume



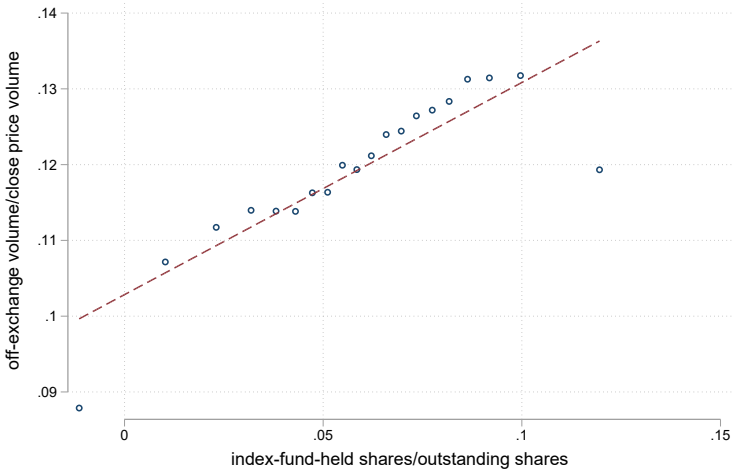
Notes. Panel (a) shows the off-exchange MOC trade volume and close auction volume as a percentage of total trade volume of S&P 500 stocks from 2012m1 to 2019m12. Panel (b) compares off-exchange MOC trade volume with close auction trade volume. In both panels, we first aggregate the volumes of each stock to monthly observations and calculate the percentages, then smooth the time series by taking three-month moving average.

Figure 1.2: Cross-sectional evidence: relationship between ETF/Index-fund ownership and off-exchange MOC volume

(a) ETF ownership vs. Off-exchange MOC volume/close price volume



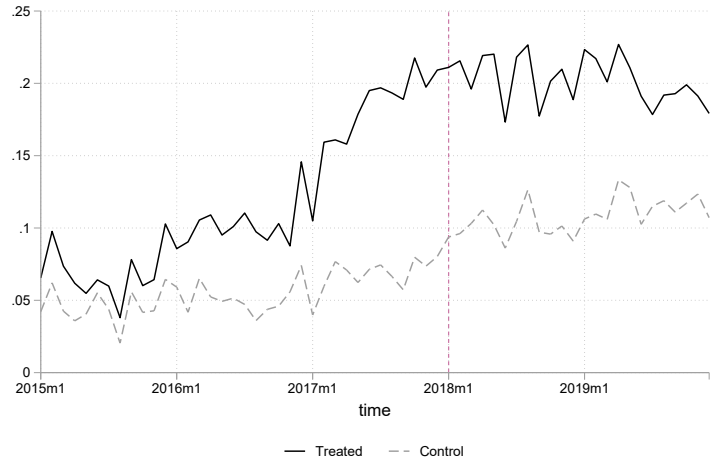
(b) Index fund ownership vs. Off-exchange MOC volume/close price volume



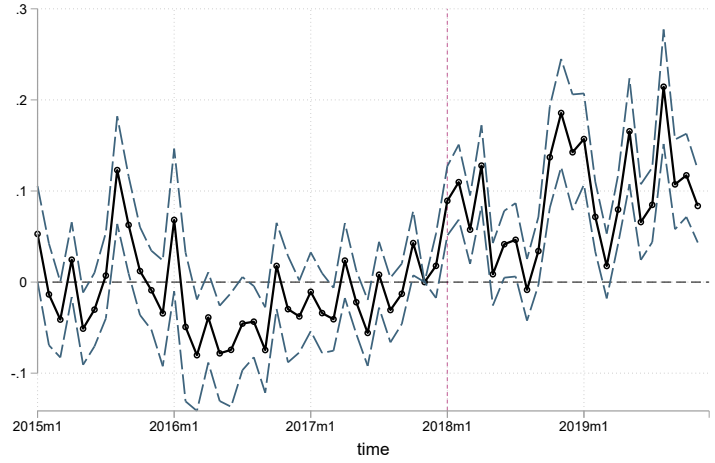
Notes. Panel (a) shows the bin-scatter plot between off-exchange MOC volume as a fraction of total volume traded at close price (the "off-exchange fraction"), and ETF ownership of stocks. Panel (b) shows the bin-scatter plot between the "off-exchange fraction" , and index fund ownership of stocks. The sample consists of 4,091 stocks (common shares), from 2015m1 to 2019m12. For each stock, we aggregate the trade volumes from daily data to monthly observations, and then calculate the fractions. ETF ownership and index fund ownership are the average ownerships in 2016. In both panels, log(market cap), log(total trade volume), volatility, time fixed effects and other control variables (see text for descriptions) are controlled in the regressions.

Figure 1.3: Trends and dynamic responses of off-exchange MOC volume and close price informativeness

(a) First stage: off-exchange MOC volume/close price volume



(b) MSE of close price



Notes. We rank NYSE stocks by the average fraction of close price volume (i.e., volume traded at the close price) executed off exchange in 2017. The treated group consists of stocks that rank at top 50% – they are more exposed to the NYSE fee cut. The control group consists of the remaining stocks. Panel (a) shows the raw means (monthly averages) of the fraction of close price volume executed off exchange of stocks in the treated group versus the control group. Panel (b) plots treatment effect of the NYSE fee cut on price informativeness, that is the DID coefficients β_k and 95% confidence intervals estimated from the model: $\text{MSE}_{i,t} = \alpha_i + \lambda_t + \sum_k \beta_k \text{Treat}_i \cdot \mathbb{I}_{t=2017m11+k} + \Gamma X_{i,t} + \epsilon_{i,t}$. Stock fixed effects, time fixed effects, log(market cap), log(total trade volume), volatility, and other control variables (see text for descriptions) are controlled in this regression. Standard errors are clustered at the stock level.

Table 1.1: Summary statistics: NYSE sample

	N	Mean	SD	10th	50th	90th
MSE ($\times 10^4$)	76,367	0.57	0.58	0.09	0.39	1.30
MSE1 ($\times 10^4$)	76,266	0.77	0.78	0.12	0.52	1.75
MSE2 ($\times 10^4$)	76,184	0.54	0.53	0.09	0.38	1.23
MSE3 ($\times 10^4$)	76,247	0.55	0.54	0.09	0.38	1.24
Median SE ($\times 10^4$)	76,367	0.24	0.47	0.03	0.11	0.51
Off-ex MOC volume/close price volume	76,395	0.11	0.09	0.00	0.09	0.24
Index fund ownership	76,398	0.07	0.04	0.00	0.08	0.11
ETF ownership	76,398	0.06	0.05	0.00	0.07	0.13
Volatility ($\times 10^6$)	76,391	0.86	2.27	0.03	0.16	1.77
Spread ($\times 10^6$)	76,396	15.81	16.14	3.31	10.28	35.26
Market cap ($ bil.)	76,393	9.68	13.73	0.32	3.54	33.62
Total volume ($1,000$ shares)	76,397	1417.72	1839.74	57.68	679.55	4038.47
Total retail volume ($1,000$ shares)	76,357	82.73	135.55	4.84	29.32	227.54
Total volume of trades $\geq$ \$20K in value	76,118	365.89	591.72	11.58	122.32	1046.03
Total volume of trades $\geq$ \$50K in value	75,491	234.78	394.92	9.52	76.23	647.77
After-close volume/total volume	74,549	0.04	0.04	0.01	0.03	0.09
Close auction volume/total volume	76,395	0.08	0.05	0.01	0.08	0.15

Notes. This table reports the summary statistics of all the variables used in this paper. The sample consists of 1,489 NYSE stocks, and spans from 2015m1 to 2019m12. Variables with daily observations are aggregated at a monthly frequency by calculating averages. MSE, MSE1, MSE2, MSE3 are mean squared error measures of price informativeness, calculated from daily data. Specifically, they are the monthly average of $(\frac{p_{t+1}^{open}-p_t^{close}}{p_t^{close}})^2$, of $(\frac{p_{t+1}^{open5m}-p_t^{close}}{p_t^{close}})^2$, of $(\frac{p_{t+1}^{open}-p_t^{close}}{p_t^{close5m}})^2$, and of $(\frac{p_{t+1}^{open}-p_t^{close}}{p_t^{close15m}})^2$, respectively. Median SE is the monthly median of $(\frac{p_{t+1}^{open}-p_t^{close}}{p_t^{close}})^2$. Close price volume is volume traded at close price (i.e., close auction volume + Off-ex MOC volume). Index fund ownership and ETF ownership are the fraction of outstanding shares held on average in 2016 by index funds and ETFs, respectively . Volatility is trade-based intrayday volatility during market hours. Spread is time-weighted percent quoted spread during market hours. Total volume is total trade volume during market hours. Total retail volume is total volume of retail trades during market hours. After-close volume is trade volume after market close and before next day's market open.

Table 1.2: DID estimated effects of NYSE fee cut on price informativeness: Main results

	(1) MSE	(2) MSE1	(3) MSE2	(4) MSE3	(5) Median SE	(6) Volatility	(7) Spread
Treat × Post	0.106***	0.141***	0.100***	0.102***	0.057***	0.028	-0.274
	(10.929)	(10.732)	(11.126)	(11.105)	(6.920)	(0.684)	(-1.713)
Volatility	0.057***	0.130***	0.052***	0.053***	0.023***		3.961***
	(18.402)	(24.741)	(18.619)	(18.488)	(7.347)		(30.386)
log(Volume)	0.138***	0.207***	0.124***	0.127***	0.051***	-1.692***	-3.310***
	(9.090)	(9.954)	(8.890)	(8.921)	(3.482)	(-13.997)	(-11.891)
N	71755	71751	71752	71753	71755	71763	71763
R^2	0.691	0.682	0.702	0.699	0.492	0.773	0.890
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes. The table reports the difference-in-differences regressions estimating the effect of NYSE fee cut on close price informativeness. The sample consists of 1,489 NYSE stocks, and spans from 2015m1 to 2019m12. The dependent variables in columns 1-4, MSE, MSE1, MSE2 and MSE3, are mean squared error measures of price informativeness, calculated from daily data. Specifically, they are the monthly average of $(\frac{p_{t+1}^{open}-p_t^{close}}{p_t^{close}})^2$, of $(\frac{p_{t+1}^{open5m}-p_t^{close}}{p_t^{close}})^2$, of $(\frac{p_{t+1}^{open}-p_t^{close}}{p_t^{close5m}})^2$, and of $(\frac{p_{t+1}^{open}-p_t^{close}}{p_t^{close15m}})^2$, respectively. The dependent variable is the monthly median of $(\frac{p_{t+1}^{open}-p_t^{close}}{p_t^{close}})^2$ in column 5, volatility during market hours in column 6, and time-weighted percent quoted spread during market hours in column 7. Treat is a dummy that takes the value of one if a stock ranks in top 50% in the average fraction of close price volume (i.e., volume traded at the close price) executed off exchange in 2017 . Post is a dummy that takes the value of one if the time is after the NYSE fee cut time (Jan 2018). The control variables are volatility, log(market cap), log(total volume), log(total retail volume), log(total volume of trades $\geq$ 20K in value), log(total volume of trades $\geq$ 50K in value), after-close volume/total volume, close auction volume/total volume, stock fixed effects and time fixed effects. See definitions of these control variables in the text. Standard errors are clustered at the stock level. T statistics are listed in the brackets, and *, **, *** indicate statistical significance at 1%, 5%, and 10%, respectively.

Appendix A: MOC Orders and Off-Exchange Trades

1. Predominant use of MOC orders in close auctions

Table A.1: NYSE Group Order Type Usage (March 2020): Percentage of Matched Volumes

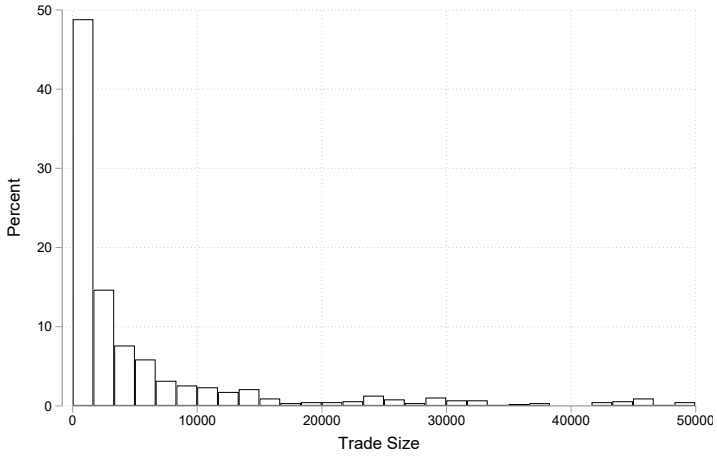
	NYSE	NYSE Arca	NYSE American
Auction	**28.21%**	**5.95%**	**8.07%**
Market-on-Close	16.37%	3.72%	4.61%
Limit-on-Close	3.63%	1.18%	1.59%
Market-on-Open	1.88%	0.55%	0.99%
Limit-on-Open	1.22%	0.49%	0.88%
Closing D-Orders	5.07%	0.00%	0.00%
Closing Offset	0.04%	0.00%	0.00%

Notes: The table reports the percentage of matched total daily volumes constituted by different order types. The table should be interpreted as: in NYSE in Mar 2020, 28.21% of total daily volume happens in auctions, and 16.37% of total daily volume are triggered by market-on-close orders. Source: `https://www.nyse.com/publicdocs/nyse/NYSE_Group_Executed_Order_Type_Usage.xlsx`

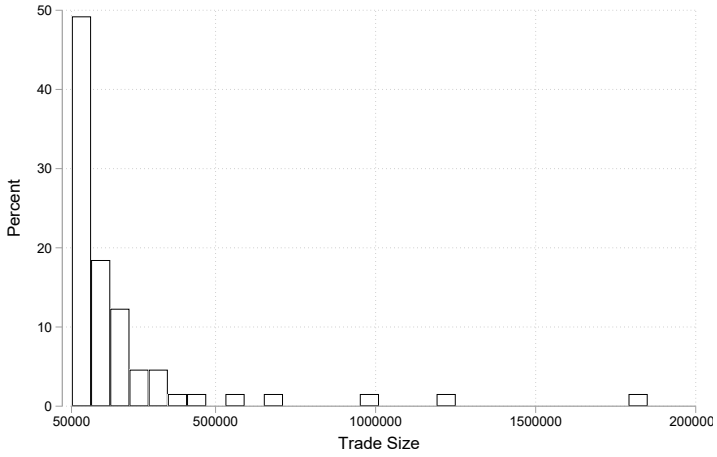
2. Distribution of off-exchange trade size

Figure A.1: Histogram of off-exchange MOC trade size for stock AAPL in Jan 2018

(a) Histogram of off-exchange MOC trade size (AAPL Jan 2018) ($\leq 50,000$ Shares)



(b) Histogram of off-exchange MOC trade size (AAPL Jan 2018) ($> 50,000$ Shares)



Appendix B: Proofs

Proof of Proposition 1

Let $H(|s|)$ be the mass of traders that trade if getting signal s, and $h(\chi)$ be the density of informed traders with alternative trading opportunity χ. By definition,

$$H(|s|) = \begin{cases} 0 & \text{if } sign(s)E[v-p|s] \leq \mu_b \\ G(sign(s)E[v-p|s] - \mu_b) \int_0^s h(\chi)d\chi & \text{if } \mu_b < sign(s)E[v-p|s] \leq \mu_c \\ \int_0^{|s|} h(\chi)d\chi & \text{if } sign(s)E[v-p|s] > \mu_c \end{cases}$$

Then since $s \sim U(v-\sigma_s, v+\sigma_s)$, total net informed orders is $I = \frac{1}{2\sigma_s} \int_{v-\sigma_s}^{v+\sigma_s} H(|s|)ds$. The total net informed orders is linear in the asset value if and only if

$$\frac{dI}{dv} = \frac{1}{2\sigma_s} \left[H(|v+\sigma_s|) - H(|v-\sigma_s|) \right] = a, \ \forall v \tag{1.1}$$

We look for sufficient conditions that does not depend on the bank's endogenous transaction fee μ_b. Assume that

$$h(s) = 0 \quad \text{if} \quad sign(s)E[v-p|s] < \mu_c \tag{1.2}$$

Then the mass of traders that trade if getting signal s can be simplified to $H(|s|) = \int_0^{|s|} h(\chi)d\chi$. And the necessary condition above can be written as

$$h(|s| + \sigma_s) = h(|s| - \sigma_s), \forall s \tag{1.3}$$

Therefore, the conditions 1.2 and 1.3 are sufficient for the total net informed orders to be linear in the asset value.

Proof of Proposition 2

The bank chooses a position $X(u)$ that maximizes its expected profits, given the net order flow $u = L_b + \theta_b a v$. The bank solves the following optimization problem

$$\max_X E[v - p|u]X(u) \tag{1.4}$$

when taking the fee μ_b as given. Here α_b, θ_b are the participation fractions of uninformed and informed traders. Guess that the market maker applies a linear price setting rule $p = \lambda y = \lambda [L + av + X(u)]$. For simplicity, we write $X(u)$ as X for solving the bank's problem.

Knowing that $v \sim N(0, \sigma^2), L \sim N(0, \sigma_L^2)$, the bank's expectation of the value of the asset is

$$\begin{aligned} E[v|u] &= E[v] + \frac{Cov(v,u)}{Var(u)}(u - E[u]) \\ &= \frac{\theta_b a \sigma^2}{\theta_b^2 a^2 \sigma^2 + \alpha_b \sigma_L^2} u \end{aligned} \tag{1.5}$$

And similarly the bank's expectation of aggregate liquidity orders is

$$E[L|u] = \frac{\alpha_b \sigma_L^2}{\theta_b^2 a^2 \sigma^2 + \alpha_b^2 \sigma_L^2} u \tag{1.6}$$

Then we can solve for the bank's problem

$$\begin{aligned} \max_X E[v - p|u]X &= E[v - \lambda [L + av + X] |u]X \\ &= ((1 - \lambda a)E[v|u] - \lambda E[L|u] - \lambda X) X \\ &= \left(\left[(1 - \lambda a) \frac{\theta_b a \sigma^2}{\sigma_u^2} - \lambda \frac{\alpha_b \sigma_L^2}{\sigma_u^2} \right] u - \lambda X \right) X \end{aligned} \tag{1.7}$$

where $\sigma_u^2 = \theta_b^2 a^2 \sigma^2 + \alpha_b \sigma_L^2$. Take first order conditions, we get

$$\begin{aligned} X &= \frac{1}{2\lambda}\left[(1-\lambda a)\frac{\theta_b a \sigma^2}{\sigma_u^2} - \lambda \frac{\alpha_b \sigma_L^2}{\sigma_u^2}\right] u \\ &= \left[\frac{1}{2\lambda}\frac{\theta_b a \sigma^2}{\sigma_u^2} - \frac{1}{2}\frac{\theta_b a^2 \sigma^2 + \alpha_b \sigma_L^2}{\sigma_u^2}\right] u \end{aligned} \tag{1.8}$$

Hence

$$X = K(\lambda) u \tag{1.9}$$

where

$$K(\lambda) = \frac{1}{2\lambda}\frac{\theta_b a \sigma^2}{\sigma_u^2} - \frac{1}{2}\frac{\theta_b a^2 \sigma^2 + \alpha_b \sigma_L^2}{\sigma_u^2} \tag{1.10}$$

Next we derive the price setting rule of the market maker and show that it takes the linear form as we guessed. The market maker receives net order

$$y = L + av + X(u) = L_c + (1 + K(\lambda))L_b + (1 + K(\lambda)\theta_b)av \tag{1.11}$$

and knows that $v \sim N(0, \sigma^2), L \sim N(0, \sigma_L^2)$. Then the close price is set by

$$p = E[v|y] = \frac{(1 + K(\lambda)\theta_b)a\sigma^2}{((1 + K(\lambda))^2 \alpha_b + (1 - \alpha_b))\, \sigma_L^2 + (1 + K(\lambda)\theta_b)^2 a^2 \sigma^2} y \tag{1.12}$$

which can be written as

$$p = \lambda y \tag{1.13}$$

where

$$\lambda = \frac{(1 + K(\lambda)\theta_b)a\sigma^2}{((1 + K(\lambda))^2\alpha_b + (1 - \alpha_b))\,\sigma_L^2 + (1 + K(\lambda)\theta_b)^2 a^2\sigma^2} \tag{1.14}$$

Therefore, the equilibrium is characterised by the bank's optimal strategy (1.10) and the exchange's price setting rule (1.14). Substituting λ as a function of K, i.e., Equation (1.14) into Equation (1.10), we can solve for K as a result of a quadratic equation:

$$0 = \sigma_u^2 K^2 + \left(\theta_b a^2\sigma^2 + (2\alpha_b - \theta_b)\frac{\alpha_b}{\theta_b}\sigma_L^2\right) K + \left(\frac{\alpha_b}{\theta_b} - 1\right)\sigma_L^2 \tag{1.15}$$

where $\sigma_u^2 = \theta_b^2 a^2\sigma^2 + \alpha_b\sigma_L^2$. Then

$$K = \frac{-B + \sqrt{B^2 - 4AC}}{2A} \tag{1.16}$$

where $A = \sigma_u^2, B = \theta_b a^2\sigma^2 + (2 - \theta_b)\frac{\alpha_b}{\theta_b}\sigma_L^2, C = \left(\frac{\alpha_b}{\theta_b} - 1\right)\sigma_L^2$.[37] And

$$\lambda = \frac{(1 + K\theta_b)a\sigma^2}{((1 + K(\lambda))^2\alpha_b + (1 - \alpha_b))\,\sigma_L^2 + (1 + K\theta_b)^2 a^2\sigma^2} \tag{1.17}$$

Both K, λ are constants that only depend on parameters. The proposition is hence proved. □

Proof of Lemma 1

Here we show that

$$\frac{(1 + K\theta_b)^2}{(1 + K)^2\alpha_b + 1 - \alpha_b} \geq 1 \tag{1.18}$$

[37] The other solution of the quadratic equation $\frac{-B-\sqrt{B^2-4AC}}{2A}$ does not optimize the bank's profit here.

where

$$K = \frac{-B + \sqrt{B^2 - 4AC}}{2A} \tag{1.19}$$

where $A = \sigma_u^2, B = \theta_b a^2 \sigma^2 + (2 - \theta_b)\frac{\alpha_b}{\theta_b}\sigma_L^2, C = \left(\frac{\alpha_b}{\theta_b} - 1\right)\sigma_L^2$.

If $\theta_b = \alpha_b$, then $K = 0$ and the LHS equals 1.

If $\theta_b > \alpha_b$, then $K > 0$, the problem is equivalent to

$$\theta_b^2 K + 2\theta_b > \alpha_b K + 2\alpha_b \tag{1.20}$$

The inequality immediately holds when $\theta_b^2 \geq \alpha_b$. When $\theta_b^2 < \alpha_b$, the inequality can be written as

$$K < \frac{2(\theta_b - \alpha_b)}{\alpha_b - \theta_b^2} \tag{1.21}$$

By substituting K as a function of the parameters and rearranging, we can write the inequality as

$$0 < D_1 a^2 \sigma^2 + D_2 \sigma_L^2 \tag{1.22}$$

where

$$D_1 = 2\frac{(\theta_b - \alpha_b)\theta}{(\alpha_b - \theta_b^2)^2}(\theta_b^2 - 2\alpha_b\theta_b + \alpha_b) \tag{1.23}$$

$$D_2 = 4\frac{(\theta_b - \alpha_b)^2}{(\alpha_b - \theta_b^2)^2}\alpha_b + \frac{1}{\alpha_b - \theta_b^2}\left(3\alpha_b - 2\alpha_b\theta_b + \theta_b^2\right)\left(1 - \frac{\alpha_b}{\theta_b}\right) \tag{1.24}$$

Since $D_1 > 0$ and $D_2 > 0$, the inequality holds. If $\theta_b < \alpha_b$, then $K < 0$, the problem can be written as

$$K > \frac{2(\theta_b - \alpha_b)}{\alpha_b - \theta_b^2} \tag{1.25}$$

By substituting K as a function of the parameters and rearranging, we can write the in-

equality as

$$0 > D_1 a^2 \sigma^2 + D_2 \sigma_L^2 \tag{1.26}$$

where

$$D_1 = 2\frac{(\theta_b - \alpha_b)\theta}{(\alpha_b - \theta_b^2)^2}(\theta_b^2 - 2\alpha_b\theta_b + \alpha_b) \tag{1.27}$$

$$D_2 = \frac{\theta_b - \alpha_b}{(\alpha_b - \theta_b^2)^2}\left[3\frac{\alpha_b^2}{\theta_b} - 6\alpha_b^2 + 2\alpha_b\theta_b + 2\alpha_b\theta_b^2 - \theta_b^3\right] \tag{1.28}$$

Since $D_1 < 0$ and $D_2 < 0$, the inequality holds and the proposition is proved.

Proof of Proposition 3

Since the bank's total profit is a continuous function of μ_b, a maximum exists on the compact set $[0, \mu_c]$. When the above conditions are satisfied, the first order derivative of the bank's total profit at $\mu_b = \mu_c$ is negative, so the maximum profit must be positive and $\mu_b < \mu_c$.

Proof of Proposition 4

Observe that the equilibrium in which the bank does not conduct the service coincides the equilibrium in which the bank does, yet the fractions of traders trading via the bank is 0, i.e., $\alpha_b = \theta_b = 0$.

So the scaled MSE in the "no bank" equilibrium is

$$\frac{MSE_{\text{no bank}}}{\sigma^2} = \frac{1}{\xi_{\text{no bank}} + 1} \tag{1.29}$$

where $\xi_{\text{no bank}} = \frac{a^2\sigma^2}{\sigma_L^2}$. Then following Lemma 1, we have $\xi \geq \xi_{\text{no bank}}$. Hence $MSE_{\text{with bank}} \leq MSE_{\text{no bank}}$.

Proof of Proposition 5

An informed trader's expected trading profit before fees in the "no bank" equilibrium coincides with the profit when $K = 0$, that is

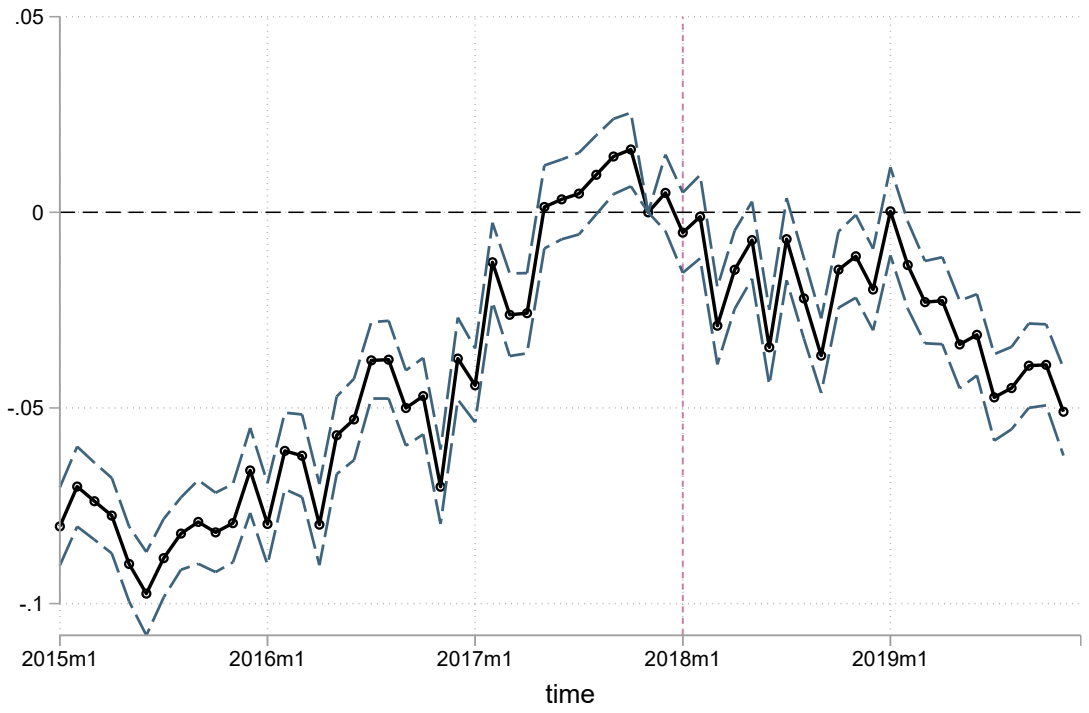
$$\Pi_{I,\text{no bank}} = \sqrt{\frac{2}{\pi}}\sigma\frac{\sigma_L^2}{\sigma_L^2 + a^2\sigma^2} \tag{1.30}$$

Following Lemma 1, we have $\Pi_I < \Pi_{I,\text{no bank}}$.

Appendix D: Internet Appendix

1.6 Robustness of DID Results

Figure B.1: Trend of off-exchange MOC trade volume and dynamic responses to NYSE fee cut



Notes. The figure plots the trend and treatment effect of the NYSE fee cut on off-exchange MOC trade volume, that is the DID coefficients β_k and 95% confidence intervals estimated from the model: (off-ex MOC volume/close price volume)$_{i,t} = \alpha_i + \lambda_t + \sum_k \beta_k \text{Treat}_i \cdot \mathbb{I}_{t=2017m11+k} + \Gamma X_{i,t} + \epsilon_{i,t}$. Stock fixed effects, time fixed effects, log(market cap), log(total trade volume), volatility, and other control variables (see text for descriptions) are controlled in this regression.

Table B.1: DID estimated effects of NYSE fee cut on price informativeness: Designating treatment group by passive ownership

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	MSE	MSE1	MSE2	MSE3	Median SE	Volatility	Spread
Treat × Post	0.088***	0.118***	0.084***	0.086***	0.043***	-0.018	0.338*
	(9.197)	(9.105)	(9.429)	(9.397)	(5.273)	(-0.493)	(2.216)
Volatility	0.057***	0.131***	0.052***	0.053***	0.023***		3.957***
	(18.301)	(24.632)	(18.512)	(18.379)	(7.397)		(30.347)
log(Volume)	0.135***	0.203***	0.121***	0.124***	0.049***	-1.689***	-3.188***
	(8.852)	(9.737)	(8.651)	(8.682)	(3.355)	(-13.799)	(-11.423)
N	71756	71752	71753	71754	71756	71764	71764
R^2	0.690	0.681	0.701	0.698	0.492	0.773	0.890
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes. The table reports the difference-in-differences regressions estimating the effect of NYSE fee cut on close price informativeness. The sample consists of 1,489 NYSE stocks, and spans from 2015m1 to 2019m12. The dependent variables in columns 1-4, MSE, MSE1, MSE2 and MSE3, are mean squared error measures of price informativeness, calculated from daily data. Specifically, they are the monthly average of $(\frac{p_{t+1}^{open}-p_t^{close}}{p_t^{close}})^2$, of $(\frac{p_{t+1}^{open5m}-p_t^{close}}{p_t^{close}})^2$, of $(\frac{p_{t+1}^{open}-p_t^{close}}{p_t^{close5m}})^2$, and of $(\frac{p_{t+1}^{open}-p_t^{close}}{p_t^{close15m}})^2$, respectively. The dependent variable is the monthly median of $(\frac{p_{t+1}^{open}-p_t^{close}}{p_t^{close}})^2$ in column 5, volatility during market hours in column 6, and time-weighted percent quoted spread during market hours in column 7. Treat is a dummy that takes the value of one if a stock ranks in top 50% in the average fraction of shares held by ETFs and index funds in 2017 . Post is a dummy that takes the value of one if the time is after the NYSE fee cut time (Jan 2018). The control variables are volatility, log(market cap), log(total volume), log(total retail volume), log(total volume of trades $\geq$ 20K in value), log(total volume of trades $\geq$ 50K in value), after-close volume/total volume, close auction volume/total volume, stock fixed effects and time fixed effects. See definitions of these control variables in the text. Standard errors are clustered at the stock level. T statistics are listed in the brackets, and *, **, *** indicate statistical significance at 1%, 5%, and 10%, respectively.

Table B.2: DID estimated effects of NYSE fee cut on price informativeness: Excluding earnings announcement days

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	MSE	MSE1	MSE2	MSE3	Median SE	Volatility	Spread
Treat × Post	0.083***	0.118***	0.082***	0.082***	0.057***	0.040	-0.189
	(11.343)	(11.498)	(11.353)	(11.380)	(8.627)	(1.008)	(-1.187)
Volatility	0.040***	0.097***	0.038***	0.039***	0.022***		4.047***
	(19.240)	(24.418)	(19.096)	(19.114)	(10.041)		(31.043)
log(Volume)	0.090***	0.153***	0.086***	0.086***	0.062***	-1.628***	-3.261***
	(8.531)	(10.187)	(8.364)	(8.369)	(5.986)	(-14.084)	(-11.833)
N	71708	71703	71705	71706	71708	71714	71714
R^2	0.736	0.713	0.736	0.737	0.574	0.770	0.890
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes. The table reports the difference-in-differences regressions estimating the effect of NYSE fee cut on close price informativeness. The sample consists of 1,489 NYSE stocks, and spans from 2015m1 to 2019m12. The dependent variables in columns 1-4, MSE, MSE1, MSE2 and MSE3, are mean squared error measures of price informativeness, calculated from daily data, excluding data within 1 day, that is, the [-1,0,1] days, from earnings announcement days. Specifically, they are the monthly average of $(\frac{p_{t+1}^{open}-p_t^{close}}{p_t^{close}})^2$, of $(\frac{p_{t+1}^{open5m}-p_t^{close}}{p_t^{close}})^2$, of $(\frac{p_{t+1}^{open}-p_t^{close}}{p_t^{close5m}})^2$, and of $(\frac{p_{t+1}^{open}-p_t^{close}}{p_t^{close15m}})^2$, respectively. The dependent variable is the monthly median of $(\frac{p_{t+1}^{open}-p_t^{close}}{p_t^{close}})^2$ in column 5, volatility during market hours in column 6, and time-weighted percent quoted spread during market hours in column 7. Treat is a dummy that takes the value of one if a stock ranks in top 50% in the average fraction of close price volume (i.e., volume traded at the close price) executed off exchange in 2017. Post is a dummy that takes the value of one if the time is after the NYSE fee cut time (Jan 2018). The control variables are volatility, log(market cap), log(total volume), log(total retail volume), log(total volume of trades $\geq$ 20K in value), log(total volume of trades $\geq$ 50K in value), after-close volume/total volume, close auction volume/total volume, stock fixed effects and time fixed effects. See definitions of these control variables in the text. Standard errors are clustered at the stock level. T statistics are listed in the brackets, and *, **, *** indicate statistical significance at 1%, 5%, and 10%, respectively.

Table B.3: DID estimated effects of NYSE fee cut on price informativeness: Matching specification

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	MSE	MSE1	MSE2	MSE3	Median SE	Volatility	Spread
Treat × Post	0.088***	0.123***	0.084***	0.085***	0.055***	0.013	0.297**
	(8.593)	(9.213)	(8.787)	(8.763)	(6.271)	(0.828)	(2.622)
Volatility	0.091***	0.178***	0.083***	0.085***	0.052***		4.732***
	(8.163)	(9.707)	(8.192)	(8.185)	(4.050)		(15.753)
log(Volume)	0.179***	0.258***	0.168***	0.170***	0.057**	-0.495***	-3.329***
	(10.366)	(11.485)	(10.452)	(10.414)	(3.265)	(-8.464)	(-14.182)
N	72308	72307	72308	72308	72308	72309	72309
R^2	0.708	0.696	0.720	0.717	0.511	0.635	0.795
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes. The table reports the robustness of the main difference-in-differences estimates to using matched treated and control stocks. We perform the matching on market cap, total trade volume, and volatility. The sample consists of 1,489 NYSE stocks, and spans from 2015m1 to 2019m12. The dependent variables in columns 1-4, MSE, MSE1, MSE2 and MSE3, are mean squared error measures of price informativeness, calculated from daily data, excluding data within 1 day, that is, the [-1,0,1] days, from earnings announcement days. Specifically, they are the monthly average of $(\frac{p_{t+1}^{open}-p_t^{close}}{p_t^{close}})^2$, of $(\frac{p_{t+1}^{open5m}-p_t^{close}}{p_t^{close}})^2$, of $(\frac{p_{t+1}^{open}-p_t^{close}}{p_t^{close5m}})^2$, and of $(\frac{p_{t+1}^{open}-p_t^{close}}{p_t^{close15m}})^2$, respectively. The dependent variable is the monthly median of $(\frac{p_{t+1}^{open}-p_t^{close}}{p_t^{close}})^2$ in column 5, volatility during market hours in column 6, and time-weighted percent quoted spread during market hours in column 7. Treat is a dummy that takes the value of one if a stock ranks in top 50% in the average fraction of close price volume (i.e., volume traded at the close price) executed off exchange in 2017. Post is a dummy that takes the value of one if the time is after the NYSE fee cut time (Jan 2018). The control variables are volatility, log(market cap), log(total volume), log(total retail volume), log(total volume of trades $\geq$ 20K in value), log(total volume of trades $\geq$ 50K in value), after-close volume/total volume, close auction volume/total volume, stock fixed effects and time fixed effects. See definitions of these control variables in the text. Standard errors are clustered at the stock level. T statistics are listed in the brackets, and *, **, *** indicate statistical significance at 1%, 5%, and 10%, respectively.

Table B.4: Distribution of variables and balance of matching

	Treated			Control			Comparison	
	A. Pre-Matching							
	Mean	Variance	Skewness	Mean	Variance	Skewness	Std-diff	Var-ratio
log(Market Cap)	2.199	1.252	-0.325	0.384	2.349	0.553	1.353	0.533
log(Total Volume)	13.896	1.141	-0.199	12.516	2.529	-0.251	1.019	0.451
Volatility	0.159	0.183	12.798	1.247	7.513	3.542	-0.555	0.024
	B. Post-Matching (Mahalanobis distance with replacement)							
	Mean	Variance	Skewness	Mean	Variance	Skewness	Std-diff	Var-ratio
log(Market Cap)	2.199	1.252	-0.325	2.180	1.293	-0.297	0.017	0.968
log(Total Volume)	13.896	1.141	-0.199	13.887	1.099	-0.129	0.008	1.038
Volatility	0.159	0.183	12.798	0.152	0.192	13.238	0.015	0.953

Notes. This table reports distributional test statistics for three variables (pre-treatment values) we use in matching treated stocks to control stocks: log(market cap), log(total volume), and volatility during market hours. The table also assesses balance between treated and control group in the means (using the standardized difference) and in the variances (using the variance ratio).

Table B.5: DID estimated effects of Nasdaq LOC rebate boost on price informativeness: Same specification as main DID results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	MSE	MSE1	MSE2	MSE3	Median SE	Volatility	Spread
Treat × Post	0.139***	0.192***	0.131***	0.133***	0.088***	-0.042	-1.923**
	(12.802)	(9.928)	(13.063)	(13.023)	(12.962)	(-0.108)	(-3.257)
Volatility	0.008***	0.020***	0.007***	0.008***	0.005***		1.593***
	(22.800)	(25.961)	(21.988)	(22.090)	(13.208)		(48.804)
log(Volume)	0.097***	0.086***	0.086***	0.088***	0.066***	-16.384***	-10.308***
	(9.076)	(3.933)	(8.676)	(8.722)	(6.589)	(-23.018)	(-15.794)
N	75159	75103	75149	75158	75159	75170	75170
R^2	0.607	0.556	0.613	0.612	0.402	0.709	0.930
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

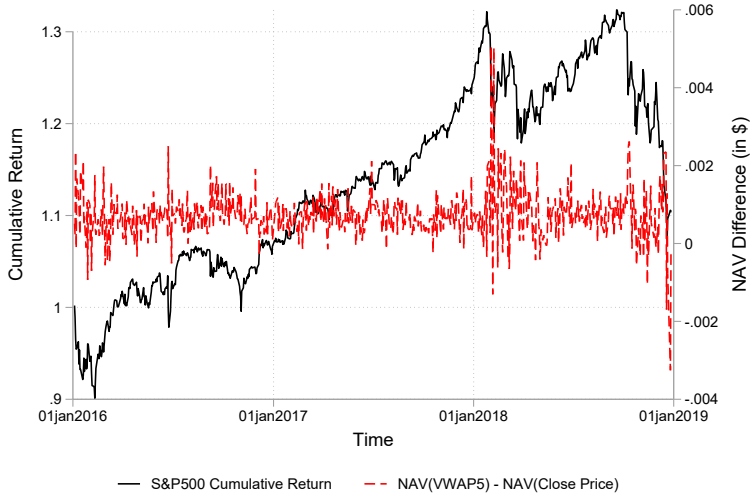
Notes. The table reports the difference-in-differences regressions estimating the effect of Nasdaq LOC rebate boost on close price informativeness. The sample consists of 1,497 Nasdaq stocks, and spans from 2015m1 to 2019m12. The dependent variables in columns 1-4, MSE, MSE1, MSE2 and MSE3, are mean squared error measures of price informativeness, calculated from daily data. Specifically, they are the monthly average of $(\frac{p_{t+1}^{open}-p_t^{close}}{p_t^{close}})^2$, of $(\frac{p_{t+1}^{open5m}-p_t^{close}}{p_t^{close}})^2$, of $(\frac{p_{t+1}^{open}-p_t^{close}}{p_t^{close5m}})^2$, and of $(\frac{p_{t+1}^{open}-p_t^{close}}{p_t^{close15m}})^2$, respectively. The dependent variable is the monthly median of $(\frac{p_{t+1}^{open}-p_t^{close}}{p_t^{close}})^2$ in column 5, volatility during market hours in column 6, and time-weighted percent quoted spread during market hours in column 7. Treat is a dummy that takes the value of one if a stock ranks in top 50% in the average fraction of close price volume (i.e., volume traded at the close price) executed off exchange in 2017 . Post is a dummy that takes the value of one if the time is after the Nasdaq LOC rebate boost time (Jan 2018). The control variables are volatility, log(market cap), log(total volume), log(total retail volume), log(total volume of trades $\geq$ 20K in value), log(total volume of trades $\geq$ 50K in value), after-close volume/total volume, close auction volume/total volume, stock fixed effects and time fixed effects. See definitions of these control variables in the text. Standard errors are clustered at the stock level. T statistics are listed in the brackets, and *, **, *** indicate statistical significance at 1%, 5%, and 10%, respectively.

1.7 Welfare Implications

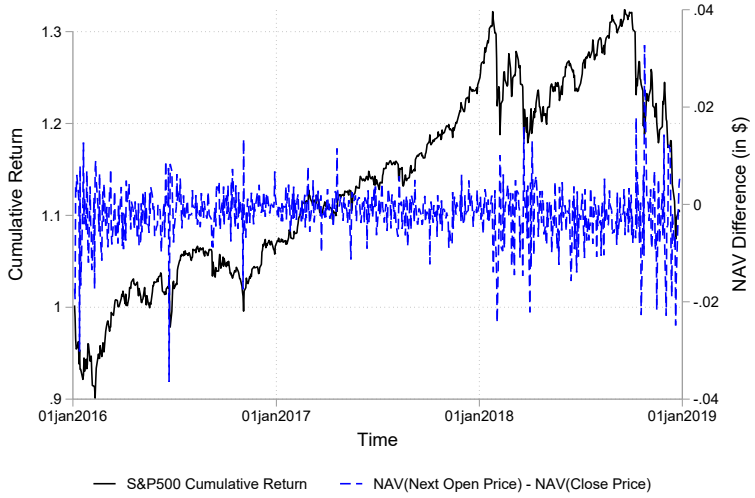
We show that deviations in close prices also have important welfare implications for general investors. Consider an index fund targeting a value-weighted portfolio of SP500 stocks with $1,000 initial NAV. Figure B.2 shows that, the fund's daily NAV would differ by $2 if portfolio stocks are evaluated by the last 5 minute volume-weighted average prices rather than the close prices. The fund's daily NAV would differ by $10 if evaluated by next day's open price.

Figure B.2: NAV differences From Deviations of Close Prices

(a) NAV difference: close price vs. last 5 min VWAP



(b) NAV difference: close price vs. next day's open price



Notes. In Panel (a), the black line shows the NAV of a S&P 500 index fund, while the red line shows the NAV differences if we use last-5-minute VWAP instead of close price for the NAV evaluation. In Panel (b), the black line shows the NAV of a S&P 500 index fund, while the blue line shows the NAV differences if we use next-day open price instead of close price for the NAV evaluation.

Chapter 2

CEO Stress, Aging, and Death

Mark Borgschulte, Marius Guenzel, Canyao Liu, and Ulrike Malmendier

Abstract

We assess the long-term effects of changing job demands on aging and mortality. First, we analyze the effect of distress shocks during the Great Recession on visible signs of aging in CEOs. Applying neural-network based machine-learning techniques to pre- and post-distress pictures, we estimate a one-year increase in apparent age over the following decade. Second, using data on the lifespans of CEOs since the 1980s, we estimate a 1.2-year decrease in response to an industry-distress shock, and a two-year increase when insulated from market discipline via anti-takeover laws. The estimated health costs are significant, also relative to known health risks.

2.1 Introduction

Job demands and work-related stress are increasingly recognized as key determinants of population health and well-being.[1] As Kaplan and Schulhofer-Wohl (2018) document, the amount of stress experienced at work has steadily grown since at least the 1950s, even as shifts in the composition of occupations have reduced job-related physical pain and tiredness for the average worker. Stress has been proposed as an explanation for pro-cyclical mortality and the efficacy of child-tax credits (Ruhm (2000); Milligan and Stabile (2011)). Health researchers argue that stress, and the damage it causes, is the mechanism underlying many health disparities (Cutler et al. (2006); Pickett and Wilkinson (2015); Puterman et al. (2016); Snyder-Mackler et al. (2020)).

Yet, there is little quasi-experimental evidence that links job demands and stressors at work directly to health outcomes. While stress arising from social hierarchies, especially in the workplace, has been proposed as an explanation for the strong relationship between socioeconomic status and life expectancy, causal evidence, for example, on the effect of promotions is limited and reaches mixed conclusions (Boyce and Oswald (2012); Anderson and Marmot (2012); Johnston and Lee (2013)).

A key reason for the lack of causal evidence is that it is challenging to disentangle the health effects of job stressors from those of income losses and financial hardship (Smith (1999)). In this paper, we overcome these identification hurdles by focusing on the top managers of large publicly traded companies. The CEOs in this sample are wealthy and unlikely to be affected by financial hardships even if they lost their job. Thus, the setting allows us to isolate direct effects on health from indirect effects due to financial constraints.

The top corporate job is a suitable setting to analyze work-related stress as CEOs work long hours, make high-stakes decisions such as layoffs or plant closures, and face uncer-

[1] See, e. g., Marmot (2005) and Ganster and Rosen (2013). A vast literature in psychology, medicine, and biology associates chronic stress with changes in hormone levels, brain function, cardiovascular health, DNA, and deleterious health outcomes (McEwen (1998); Epel et al. (2004); Sapolsky (2005)).

tainty in times of crisis (Bandiera et al. (2020); Porter and Nohria (2018)). They are closely monitored and criticized when their firm is underperforming, and media frequently reports on "overworked [and] overstressed" CEOs.[2] Other, lower-ranked and non-corporate positions might of course entail even higher levels of stress—think of "life-or-death" jobs, such as emergency room doctors and airline pilots, or of minimum-wage jobs with rigid schedules, such as delivery drivers. While our analysis does not speak to the question of which occupations come with the highest personal cost,

the high stakes of the CEO position, and public and plausibly exogenous variation in the economic conditions faced by different CEOs help establish and quantify the influence of job demands on health outcomes.

In order to investigate the link between work-related stress and health, we assemble two measures of health outcomes: visible patterns of aging and life expectancy. Visible signs of aging are widely used by clinicians to assess patient health, and validated as biological markers predicting health outcomes and mortality at least since the Baltimore longitudinal study (Borkan and Norris (1980)). Building on the measurement of so-called "perceived" or "apparent" age from facial photographs in prior medical studies (e. g., Christensen et al. (2004) and Christensen et al. (2009)), we construct a new dataset of photographs of CEOs' faces. We utilize recent visual machine learning (ML) techniques to determine apparent age, and estimate the influence of work-related stress. The ML techniques are a promising avenue for the assessment of work-induced strains in broader samples and, to the best of our knowledge, we are the first to introduce them into the economic literature. We also estimate the influence of work-related stressors on mortality directly, in a new data set on the lifespan of CEOs. Our finding on visible signs of aging and mortality are consistent with each other.

[2]See CNN's Route to the Top segment (cnn.com/2010/business/03/12/ceo.health.warning/index). Cf. also Harvard Business Review on "How Top CEOs Cope with Constant Stress" (hbr.org/2011/04/how-top-ceos-cope-with-constan) and expert psychologists offering "Strategies for CEOs to reduce stress" (vistage.com/research-center/personal-development/20200402-ceo-stress).

We capture variation in work-related stress exploiting periods of industry-wide distress and variation in the intensity of CEO monitoring due to changing corporate-governance regulation. We define industry distress based on a 30% median firm stock-price decline over a two-year horizon, similar to prior work (Opler and Titman (1994); Acharya et al. (2007); Babina (2020)). Prior work has shown that industry shocks and financial distress affect CEO pay and turnover (Bertrand and Mullainathan (2001); Garvey and Milbourn (2006); Jenter and Lewellen (2015)). We identify variation in CEO monitoring from the staggered passage of anti-takeover laws across U.S. states in the mid-1980s, following a large literature in corporate finance. The laws shielded CEOs from market discipline by making hostile takeovers more difficult, reducing CEOs' job demands and allowing them to "enjoy the quiet life" (Bertrand and Mullainathan (2003)).[3] Thus, they constitute a separate and oppositely-signed change in job demands compared to industry distress. Both measures of job demands build on the notion of distress and protection from stress in the economic literature (and on the popular notion of stress), rather than a biomedical analysis of stress in the sense of measuring adrenaline or cortisol levels.[4]

Our analysis has three main parts. The first part exploits industry-level distress shocks during the Great Recession and relates them to visible signs of accelerated aging, identified by neural-network based ML estimations. In the second part, we continue to exploit industry-level distress shocks, now from longer and earlier time periods, and relate them to CEO mortality. In the third part, we continue to focus on CEO mortality, and study the effect of variation in the intensity of CEO monitoring due to corporate-governance legislation.

[3]The prevailing view in law and economics at the time of the passage of the laws was that the "continuous threat of takeover" is an important means to counteract lagging managerial performance (Easterbrook and Fischel (1981)). Other experts made similar arguments at the time. Scharfstein (1988) develops a formal model in which the threat of a takeover disciplines management, and then-U.S. Supreme Court Justice Byron White's opinion in *Edgar vs. MITE* emphasizes "[t]he incentive the tender offer mechanism provides incumbent management to perform well."

[4]Stress arises from experiencing demands without sufficient resources to cope (Lazarus and Folkman (1984)). Biomedically, changes in hormones and other bodily processes due to stress can cause long-term damage and accelerate aging (Brondolo et al. (2017); Franceschi et al. (2018); Kennedy et al. (2014)).

In the first part of the paper, we document relatively immediate health implications of industry crises as captured by visible signs of aging in the faces of CEOs. We employ the ML algorithms from Antipov et al. (2016) that are specifically designed to estimate a person's apparent age, i. e., how old a person looks, rather than a person's biological age. The software, trained on more than 250,000 pictures, is the winner of the *2016 ChaLearn Looking At People* competition in the *apparent-age estimation* track.[5]

We collect a sample of 3,086 pictures of the 2006 *Fortune 1000* CEOs from different points during their tenure (the "CEO Apparent Aging Data") to estimate differential apparent-aging in response to industry-level distress exposure to the financial crisis. Using a difference-in-differences design, we estimate that CEOs look about one year older in post-crisis years if their industry experienced a severe decline in 2007-2008 relative to CEOs in other industries. The estimated difference between distressed and non-distressed CEOs increases over time and amounts to 1.178 years for pictures taken five years and more after the onset of the crisis. We include a detailed description of the procedure and examine issues that have been shown to impact the use of visual machine learning in other settings (Wang and Kosinski (2018); Dotsch et al. (2016); Agüera y Arcas et al. (2018)). To the best of our knowledge, this represents the first application of visual machine learning to a quasi-experimental research design. Our application illustrates its potential for the study of health and aging to complement standard measures based on mortality, hospital admissions, or survey responses.

In the second and third part of the analysis, we study CEO mortality effects associated with industry distress and anti-takeover protection. For these analyses, we extend the data from Gibbons and Murphy (1992) comprised of CEOs in the *Forbes* Executive Compensation Surveys from 1975 to 1991, and merge it with hand-collected data on the exact dates of birth and death of more than 1,600 CEOs of large U.S. firms (the "CEO Mortality Data"). We also add pay, tenure, and firm characteristics to the data. We restrict

[5]The certification effect is roughly comparable to a first-tier publication in other academic fields.

all analyses to CEOs appointed before the enactment of the anti-takeover laws we study in the third part of the analysis, to address the concern that their passage altered the selection of CEOs.

In the analysis exploiting industry distress shocks, we estimate a hazard regression model. Controlling for a CEO's biological age, time trends, industry affiliation, and firm location, we show that industry distress significantly increases CEOs' mortality hazard by about 15%. These results are robust to an array of alternative specifications, including models with CEO birth-cohort fixed effects and birth-cohort-specific age controls.

The estimated mortality effect sizes are large. The effect of experiencing industry distress is equivalent to reducing a CEO's biological age by 1.2 years. The effect size is even slightly larger if we use life tables instead of the estimated CEO age effects to calculate the biological age equivalent of industry distress in terms of mortality hazard. We can also compare the estimated mortality effects to known health threats. For example, smoking until age 30 is associated with a reduction in longevity by roughly one year, and lifelong smoking with a reduction by ten years and more (General (2014); Jha et al. (2013)).

In the analysis exploiting changes in corporate governance regulation, the identifying variation comes from the staggered passage of anti-takeover laws across U.S. states in the mid-1980s. A large prior literature has used the passage of theses laws as proxies for less intense monitoring and shown the effects on CEO behavior. For example, CEOs became less tough in wage negotiations, and their rate of plant closures as well as plant creations decreased (Bertrand and Mullainathan (2003)); they undertook value-destroying actions that reduce their firms' risk of distress (Gormley and Matsa (2016)); patent count and quality decreased (Atanassov (2013)); and managers reduced their stock ownership (Cheng et al. (2004)). While some later studies question whether the passage of anti-takeover laws in fact reduced hostile takeover activity (e. g., Cain et al. (2017)), it

arguably constituted a significant shift in managers' *perception* of their job environment.[6] Consistent with those prior findings, we estimate significantly positive effects on CEO health. Using analogous hazard models as above, we find that anti-takeover laws significantly increase the life expectancy of incumbent CEOs. One additional year under lenient governance lowers mortality rates by four to five percent, corresponding to an overall life expectancy gain of two years for an average protected CEO in the sample. The slightly larger mortality effect sizes, compared to the distress effects, might reflect the more permanent changes in monitoring intensity, relative to temporary distress-induced changes in job demands. The results are robust to the same robustness specifications as above, as well as alternative subsampling and classifications of anti-takeover laws that account for other firm or state anti-takeover provisions, exclude lobbying and opt-out firms, or cut data differently based on firms' industry affiliation or state of incorporation (cf. Cain et al. (2017); Karpoff and Wittry (2018)).

We find no evidence of a compensating differential in the form of lower pay for CEOs as a result of being permanently protected from hostile takeovers. The analysis of pay builds on Bertrand and Mullainathan (1998) and predictions in Edmans and Gabaix (2011)'s Edmans and Gabaix (2011) CEO market model. This may indicate that not all parties fully account for the health implications of job demands.[7]

Our findings establish significant health consequences arising from variation in job demands. Our application to the CEO context allows to distinguish work-related stress from financial and job insecurity, but is of interest in its own right for at least two reasons. First, CEOs bear the ultimate responsibility for the firm and its employees. Given their

[6]Consistent with CEOs' perceptions of anti-takeover laws changing job demands, we find suggestive evidence that takeover-protected CEOs remain on the job for longer. We also show, though, that prolonged tenure cannot explain the estimated longevity effects as the survival gains are concentrated in the first few years of anti-takeover law exposure, with prolonged exposure resulting from prolonged tenure not leading to incremental survival gains. Our results are also robust to using CEOs' *predicted* rather than actual anti-takeover law exposure length (see Section 2.5 for details).

[7]We note, though, that prior literature has generally struggled to find evidence of compensating differentials outside of select settings and carefully designed experiments (e. g., Mas and Pallais (2017); Lavetti (2020)).

overarching importance for firm performance and job stability of workers, it matters how incentives and performance affect CEOs personally. Second, the health implications of CEOs' job demands affect their ability to stay on the job and, if anticipated, their willingness to select into the CEO job. Our analysis might thus speak to the prevalence of certain CEO characteristics and possible feedback effects: Are aspiring CEOs (over-)confident about their health? Are women vastly underrepresented in the C-suite not only because of discrimination but also because they (correctly) anticipate the health costs of assuming such positions? Our findings motivate further exploration of possible selection effects related to the health implications of job demands.

Our paper contributes to the literatures examining CEOs and firms, and separately, health and labor market linkages. A recent literature sheds light on CEOs' demanding job and time requirements (Bandiera et al. (2020); Bandiera et al. (2018); Porter and Nohria (2018)). Working these long hours requires physical stamina, and consistent with our hypothesis, previous work has documented that CEO health is an input into production. Bennedsen et al. (2020) study the negative effect of CEO hospitalizations on firm performance. Keloharju et al. (2020) find that corporate boards in Scandinavia factor CEO health into CEO appointment and retention decisions. None of these papers, however, examines the effect of CEO job demands on CEOs' health trajectories. To the best of our knowledge, we are the first to explore quasi-random variations to establish significant health costs, in terms of the mortality and visible aging. The only prior work on executives' health outcomes is Yen and Benham (1986), who compare the age-adjusted mortality rates of 125 executives in the banking industry to those in other industries. Our significantly larger sample and quasi-experimental design allows to control for industry-specific selection into job environments, and to implement a rigorous survival analysis.

Second, our paper contributes to the literature on the health effects of stress, socioeconomic status, and financial insecurity. In health and labor economics, stress has been

proposed as an explanation for the association between job loss and higher mortality (Sullivan and Von Wachter (2009)); the health benefits of the EITC (Evans and Garthwaite (2014)), unemployment insurance (Kuka (2020)), and access to health care (Koijen and Van Nieuwerburgh (2020)); and early-life health disparities (Camacho (2008); Black et al. (2016)). Stress is also implicated in the intergenerational persistence of poverty (Aizer et al. (2016); Persson and Rossin-Slater (2018); East et al. (2017)). To the best of our knowledge, the only paper that relates quasi-random increases in job demands directly to health outcomes is Hummels et al. (2016), who document the negative impact of trade shocks on workers' stress, injury, and illness. Evolving worker-firm interactions may have eroded protections from competition in the product market, likely exposing workers to higher levels of stress (Bertrand (2004)).

Turning from the general or poorer populations to wealthier populations, income appears to play a small role in health disparities among the already-wealthy, though Engelberg and Parsons (2016) document a link between stock-market crashes and hospital admissions, especially for anxiety and panic disorders. Social factors, such as the prestige of winning a Nobel prize or an election may be protective (Rablen and Oswald (2008); Cesarini et al. (2016); Borgschulte and Vogler (2019)). Our paper offers complementary evidence that work-related stressors impose long-term health costs, even for successful and wealthy individuals.

In the remainder of the paper, Section 2.2 describes the data and discusses the identifying variation. Section 2.3 presents the results on apparent aging and industry-wide distress shocks. The results pertaining to life expectancy and distress shocks are in Section 2.4, and life expectancy and corporate governance regulation in Section 2.5. Section 2.6 concludes.

2.2 CEO Datasets and Variation in CEO Job Demands

A. CEO Apparent Aging Data

To study visible signs of aging in CEOs' faces, we collect pictures of CEOs of the firms in the 2006 *Fortune 1000* list. Using a relatively recent CEO sample is necessary as both picture availability and quality have substantially improved over time. We focus on the 2006 cohort to exploit differential exposure to industry shocks during the Great Recession.

We search for five pictures from the beginning of a CEO's tenure and two additional pictures every four years after that. The main challenge is finding *dated* pictures in order to compare CEOs' apparent age to their true age. (Pictures from LinkedIn, for instance, do not satisfy this criterion as we generally have no information on when LinkedIn profile pictures were taken.) In addition, we aim for pictures that are taken in daily life, such as at social events or conferences, rather than posed pictures. The most useful source given these criteria is gettyimages.com, followed by Google Images. We are able to find at least two pictures from different points in time during or after their tenure for 463 CEOs, of whom 452 are male and 447 are White. Among the sixteen non-White CEOs, seven are African-American, two are Hispanic or Latinx, and seven are Asian (including Indian).[8] The final data includes 3,086 pictures; we call this the CEO Apparent Aging Data.

Table I provides the summary statistics for this data set, both at the image level (Panel A) and the CEO level (Panel B). The median picture is from 2009. On average, we are able to find about 7 pictures of a CEO (conditional on finding at least two pictures), and the average CEO is 55.54 years old in 2006 and has been in office for eight years. 65% of the CEOs in the data experienced an industry distress shock during the financial crisis (see Section II.C for details on the definitions of our job demand measures). In our analyses, we will also control for any prior industry shocks, which is zero at the median and

[8] We collect this information through Google searches and Wikipedia.

0.54 on average. The majority of CEOs head firms in the manufacturing, transportation, communications, electricity and gas, and finance industries (Panel C). We will discuss the picture-level characteristics from Panel A in Section III.A.

B. CEO Mortality Data

To study CEO mortality, we conduct a mortality followup for the universe of CEOs included in the *Forbes* Executive Compensation Surveys from 1975 to 1991, as originally collected by Gibbons and Murphy (1992).[9] These surveys are derived from corporate proxy statements and include the executives serving in the largest U.S. firms. We choose 1975 as the start year given the source of identifying variation in the third part of our analysis, i. e., the timing of anti-takeover laws (see Section II.C), and also in line with prior studies in this literature.[10] We include all firms with a PERMNO identifier in CRSP. The initial sample comprises 2,720 CEOs employed by 1,501 firms.

We manually search for (i) the exact dates of CEOs' birth, (ii) whether a CEO has died, and (iii) the date of death if the CEO has passed away. All CEOs who did not pass away by the cutoff date of October 1st, 2017 are treated as censored. Our main source of birth and death information is Ancestry.com, which links historical birth and death records from the U.S. Census, the Social Security Death Index, birth certificates, and other historical sources. To ensure that we have identified the correct person, we validate Ancestry's information with online and newspapers searches, e. g., on birth place, elementary school, or city of residence. Identifying a person as alive turns out to be more difficult as there is little coverage of retired CEOs. We classify a CEO as alive whenever recent sources confirm the alive status, such as newspaper articles or websites that listing the CEO as a board member, sponsor, donor, or chair of an organization or event.[11] We

[9] We are very grateful to Kevin J. Murphy for providing the data.

[10] Bertrand and Mullainathan (2003), Giroud and Mueller (2010), Gormley and Matsa (2016) all start their sample in the mid-1970s. Our results are robust to varying the start year (see Section V.E).

[11] We use sources dated 01/2010 or later to infer alive status since recent coverage of a retired CEO

obtain the birth and death information for 2,361 CEOs from 1,352 firms in the post-1975 sample, implying a finding rate of 87%. We test and confirm that the availability of birth and death information is not correlated with our measures of variation in work-related stressors, namely, industry distress experience and incorporation in a state that passed a BC law. We call this the CEO Mortality Data.

We augment the CEO Mortality Data with several key variables. We identify the historical states of incorporation during CEOs' tenure, needed to measure CEOs' exposure to anti-takeover laws. Since CRSP/Compustat backfills the current state of incorporation, we access its historical Comphist and Compustat Snapshot data as well as incorporation data recorded at issuances and merger events in the SDC database. In case of discrepancies, we use firms' 10-Ks and other SEC filings, legal documents, and news articles to identify the correct information. We correct the state of incorporation in 169 cases, which is 6.7% of the initial sample with state-of-incorporation information (2,514 CEOs). Out of the sample of 2,361 CEOs with birth and death information, we are able to identify the historical state of incorporation for 2,209 CEOs.

We collect tenure information for all individuals in the CEO Mortality Data to fill the gaps and correct misrecorded data in the *Forbes* Executive Compensation Surveys. We use Execucomp, online searches, and especially the *New York Times* Business People section, which frequently reports on executive changes in our sample firms. When the exact month of a CEO transition is missing, we use the "mid-year convention" motivated by the relatively uniform distribution of CEO starting months in Execucomp (Eisfeldt and Kuhnen (2013)). We further restrict the sample to CEOs whose firm was included in CRSP during the time of their tenure (1,900 CEOs).[12]

Finally, we address selection concerns with regard to variation in job demands. For

makes it very likely that news outlets would also have reported their passing had it occurred by the cutoff date (October 1st, 2017). Our results are robust to ending our sample in 2010 (Sections IV.E and V.E), and to restricting the sample period for CEOs classified as alive as of 10/2017 to end in 01/2010.

[12]Relative to the previously mentioned restriction to firms with a PERMNO in CRSP, we drop CEOs who served, for instance, before their firm went public.

example, it would confound our analyses if less resilient managers, i. e., those more prone to health ailments, became more likely to seek the CEO position when governance regulation makes CEO monitoring more lenient. To alleviate such concerns, we focus on CEOs appointed prior to the enactment of the business combination laws as our main sample (1,605 CEOs). That said, our results are robust to being estimated on the enlarged sample of 1,900 CEOs. Similarly, in the analysis exploiting industry distress, we consider the effects on the mortality of CEOs appointed before the distress shock, whether or not they left their position during the crisis.

Table II presents the summary statistics. The median CEO in this sample was born in 1925, became CEO at age 52, and served as CEO for nine years. There is relatively large heterogeneity in tenure, with an interdecile range of 17 years. Non-integer values result from CEOs not starting or ending their tenure or stepping down at the end of the year. Seventy-one percent of our CEOs have passed away by the censoring date (October 1st, 2017). The median CEO died at age 83, and passed away in 2006. Forty percent of CEOs witness industry distress during their tenure. Throughout, we will focus on a binary rather than a cumulative distress measure, to avoid endogeneity from picking up additional industry shocks long into a CEO's tenure.[13] Similarly, about forty percent of CEOs are protected by a BC law at some point during their tenure.

C. Variation in Job Demands

We exploit two sources of variation in job demands, industry-wide distress shocks and the implementation of state-level anti-takeover protection. The empirical analysis examines how both types of variation affect aging patterns (in the CEO Apparent Aging Data) and survival patterns (in the CEO Mortality Data).

Industry-Wide Distress Shocks. Industry distress shocks induce a temporary increase

[13]Long-serving CEOs are associated with longer lifespans in the data. Multiple distress shocks are rare in our sample, with more than 80% of CEOs experiencing at most one distress shock during their tenure.

in job demands. Similarly to Opler and Titman (1994), Babina (2020), and Acharya et al. (2007), we define an industry as distressed in year t if the median firm's two-year stock return (forward-looking) is less than -30%. As in Babina (2020), we generate the annual industries-in-distress panel (i) restricting to single-segment CRSP/Compustat firms, i. e., dropping firms with multiple reported segments in the Compustat Business Segment Database; (ii) dropping firms if the reported single segment sales differ from those in Compustat by more than 5%; (iii) restricting to firms with sales of at least $20 million; and (iv) excluding industry-years with fewer than four firms.[14] This annual distress panel serves as the foundation for our mortality analysis (Section 2.4). For the apparent-aging analysis of the 2006 *Fortune 500* CEO cohort, we specifically examine industry distress during the Great Recession (see Section 2.3 for additional detail). Following prior work, we use 3-digit SIC classes to measure industry affiliation and rely on historical SIC codes for the firms in our sample.

Anti-Takeover Laws. Anti-takeover statutes increase the hurdles for hostile takeovers and help protect a CEO's job. They induce a shift in the opposite direction of industry distress, and are of a more permanent nature. Thus, they constitute a useful alternative approach to analyzing the health consequences of a CEO's job demands.

Following prior literature, we focus on the second-generation anti-takeover laws, which states started passing in the mid-1980s after the first-generation laws were struck down by courts in the 1970s and early 1980s (cf. Cheng et al. (2004); Cain et al. (2017)). The statutes included Business Combination (BC) laws, Control Share Acquisition, Fair Price, and Directors' Duties laws, and Poison Pills. We follow prior literature and first focus on BC laws as the most potent type of statutes, but will return to the other types of laws in Section V.E. BC laws significantly reduced the threat of hostile takeovers by imposing a moratorium on large shareholder conducting certain transactions with the firm, usually for

[14] Sections 2.3 and 2.4 also discuss more restrictive distress definitions, using industry returns in conjunction with sales growth, or specific recession periods.

a period of three to five years.

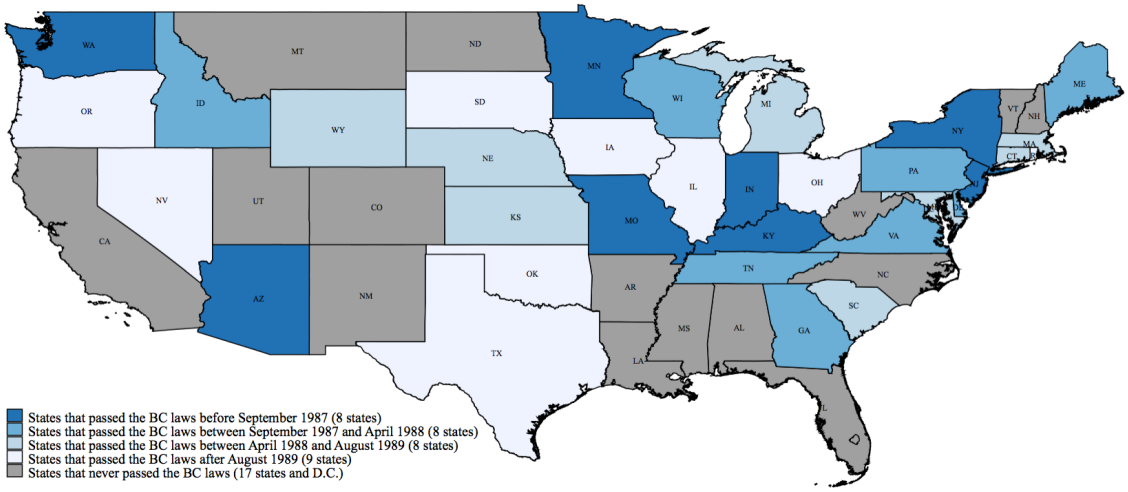


Figure 1.—Introduction of Business Combination laws over time.

Notes: The map omits the states of Alaska and Hawaii, which never passed a BC law.

An advantage of using anti-takeover laws as identifying variation is that these laws applied based on the state of incorporation, not the state of firms' headquarters or operation. The frequent discrepancies between firms' location and state of incorporation enables us to assess the impact of the laws while controlling for shocks to the local economy.

Figure 1 visualizes the staggered introduction of BC laws across states.[15] The map illustrates the variation across both time and states as a source of identification: 33 states passed a BC law between 1985 and 1997, with most laws being passed in 1987-1989. Consistent with prior literature, the most common state of incorporation in our dataset is Delaware. Other common states include New York and Ohio.

2.3. Industry-Wide Distress Shocks and Apparent Aging

In the first step of our analysis, our focus in on non-fatal manifestations of CEOs' health associated with demanding job environments. Research in medicine and biology has es-

[15] Appendix-Figure D.1 contains a similar map based on the earliest enactment of any of the five types of second-generation anti-takeover laws listed above.

tablished numerous links between stress and signs of visible aging, such as hair loss (Choi et al. (2021)), hair whitening Zhang et al. (2020), and inflammation, which in turn accelerates skin aging (Heidt et al. (2014); Kim et al. (2013)). Moreover, visible signs of aging predict mortality in longitudinal studies. For example, the apparent age of Danish twins, rated from facial photographs, predicts short-term mortality over the next two years (Christensen et al. (2004)) and long-term survival of twins aged 70 and older (Christensen et al. (2009)).[16]

We ask whether experiencing industry distress predicts accelerated apparent-aging in the CEO Apparent Aging Data. We exploit CEOs' differential exposure to industry shocks during the Great Recession in a difference-in-differences framework.

A. Apparent-Age Estimation

To analyze visible CEO aging, we make use of recent advances in machine learning on estimating people's age. While most of the earlier generations of age-estimation software focused on a person's *biological*, i. e., "true" age (Antipov et al. (2016)), recent research aims at estimating a person's *apparent* age, i. e., how old a person looks. Progress in this area has been made possible by the development of deep learning in convolutional neural networks (CNNs) and the increased availability of large datasets of facial images with associated true and apparent ages, the latter estimated by people.

We provide a detailed discussion of CNNs and the software training steps in Appendix B.1, and give a brief summary here. A CNN is a neural network which employs the method of convolution, i. e., of transforming the data by sliding (or, convolving) over it using a slider matrix, to abstractly determine intermediate features about the data such as edges or corners.

We use a machine-learning based software (Antipov et al. (2016)) that has been specifi-

[16]Christensen et al. (2009) also establish strong correlations between apparent age and physical functioning (e. g., strength and endurance tests), cognitive functioning (e. g., verbal fluency and recall), and leucocyte telomere length (which is associated with ageing-related diseases and mortality).

cally developed for the problem of apparent-age estimation. This software is the winner of the *2016 Looking At People apparent-age estimation* competition. The software is based on Oxford's Visual Geometry Group deep CNN architecture.

In a first step, the software was trained on more than 250,000 pictures with information on people's true age using the Internet Movie Database and pictures from Wikipedia. In a second step, it was fine-tuned for apparent-age estimation using a newly available dataset of 5,613 facial pictures, each of which was rated by at least ten people in terms of the person's age. The addition of fine-tuning on this apparent-age data is particularly important; this step led to the software's largest accuracy improvement (of more than 20%) in the apparent-age estimation of the competition data (see Table 2 in Antipov et al. (2016) and Appendix 1).

Both the distribution of true ages used for training and human age estimations used for software fine-tuning covers people from all age groups, including elderly people. The output of the neural network is a 100×1 vector of probabilities associated with all apparent ages from 0 to 99 years. The apparent-age point estimate is derived by multiplying each apparent age with its probability. The software also carries out an eleven-fold cross-validation, drawing 5,113 images for each training and 500 (non-overlapping) images for each validation sample. The ultimate output is the average apparent-age estimation of the eleven models.

Figure 2 provides graphical evidence of the distributions and correlations of biological and apparent ages for our sample of 3,086 pictures in the CEO Apparent Aging Data. Panel (a) shows that the distributions of apparent and biological ages largely overlap, though the apparent-age distribution is somewhat shifted to the left. That is, on average, the software estimates CEOs to look younger than their biological age. This reflects that CEOs have high socioeconomic status (SES), have better access to health care, can afford healthier food, and live longer than the average population (see Table II discussed in Section II.B, and cf. Chetty et al. (2016)). Our results below on the effect of industry shocks on CEO

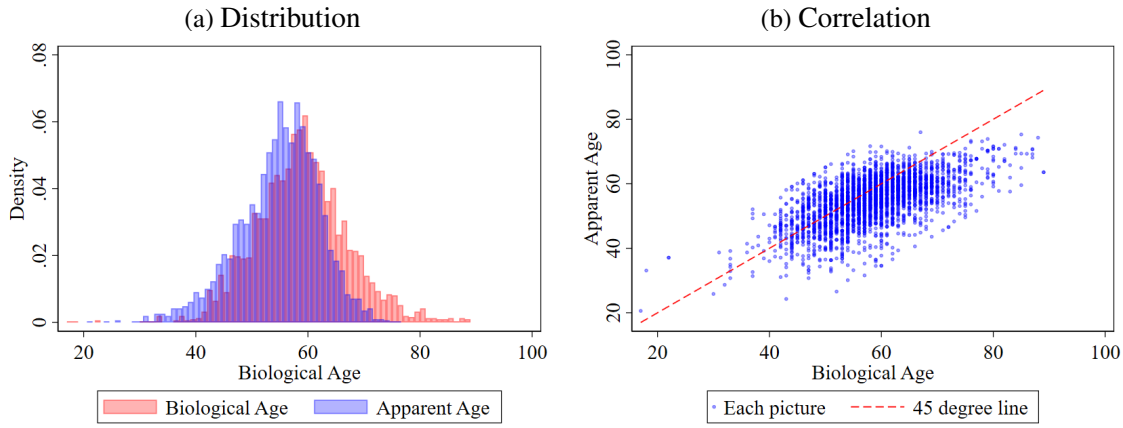


Figure 2.—CEO apparent and biological age.

Notes: The figure plots apparent and biological ages for our sample of 3,086 CEO images. Panel (a) shows the CEO apparent-age distribution in blue, and the biological-age distribution in red, with the overlapping areas appearing as purple. Panel (b) shows a scatter plot of CEOs' apparent age against biological age. The dotted line represents the 45-degree-line. We winsorize the apparent age by first winsorizing the top and bottom 0.5% of the difference between apparent and biological age, and then adding this winsorized difference to the biological age.

aging do not rely on comparisons between CEOs and the general population but entail solely *within-CEO* comparisons.

Panel (b) shows a scatter plot of CEOs' apparent age against biological age, confirming a high correlation between the two concepts of age, but also a greater mass below than above the 45-degree-line. In this figure and in the regression analysis below, we winsorize the estimated apparent-age variable to ensure that the outliers in age estimation do not affect the results. To do that, we first winsorize the top and bottom 0.5% of the difference between apparent and biological age, and then add this winsorized difference to the biological age.

To illustrate the proposed channel from industry shocks to aging, we first discuss a specific example. James Donald was the CEO of Starbucks from April 2005 until January 2008, when he was fired after Starbucks' stock had plunged by more than 40% over the preceding year. The top of Figure 3 shows two pictures of Donald: the left one was taken on December 8, 2004, before his appointment at Starbucks, and the right one 4.42 years

later, on May 11, 2009, after his dismissal. Donald was 50.76 years old in the first picture, and 55.18 years in the second. The machine-learning based aging software predicts his age in the earlier picture at 53.47 years, and in the later picture as 60.45 years. Thus, for both pictures, the software determines that he looks older than his true age. Most importantly, the software estimates that he aged by 6.98 years, i. e., 2.5 years more than actual time passed.

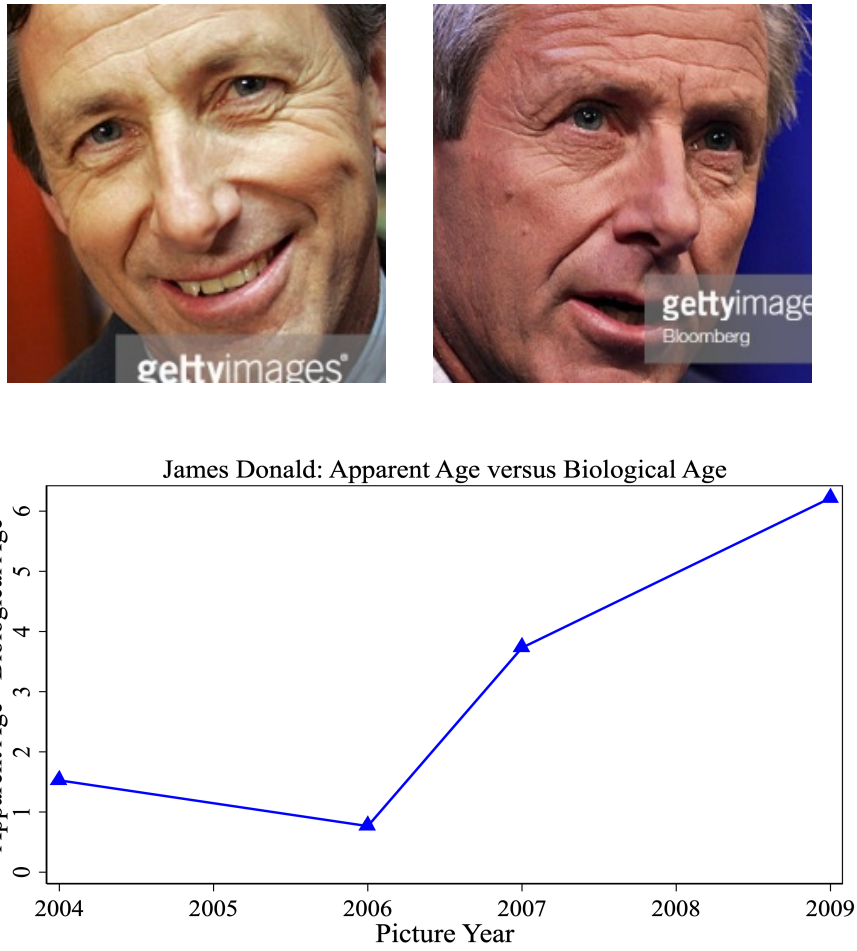


Figure 3.—Sample pictures (James Donald, CEO of Starbucks from 2005 to 2008).

Notes: The first two pictures show James Donald, CEO of Starbucks from 2005 to 2008. Based on data from Ancestry.com, Donald was born on March 5, 1954. The picture on the left was taken on December 8, 2004, that on the right on Monday, May 11, 2009. Biological ages: 50.76 and 55.18 years, respectively. Apparent ages based on aging software: 53.47 and 60.45 years, respectively. The figure at the bottom shows how James Donald's apparent age compares to his true age over time based on 20 pictures collected for the period from 2004 to 2010.

Turning to the full set of 20 pictures of Donald that we are able to collect for the period from three years before to three years after the onset of the crisis in 2007, i. e., 2004-2010, we find that the mean difference between his apparent and his biological age is 0.96 years prior to 2007 and increases to 4.97 years from 2007 on. The bottom half of Figure 3 summarizes these estimates and visualizes the jump in Donald's apparent versus biological age in 2007 as well as the continued aging effects after the crisis. The example typifies our approach, especially in light of Donald's struggles during his final year as Starbucks' CEO.

The example also allows us to discuss concerns one may have regarding picture heterogeneity in the CEO Apparent Aging Data. For example, the lighting in the two pictures seems to be different, and the left picture, with Donald smiling into the camera, might be from a more staged setting than the right one. More broadly, researchers have pointed to the importance of accounting for picture context and facial positioning in other settings, such as in inferring people's character, attractiveness, or sexual orientation from facial images (Wang and Kosinski (2018); Dotsch et al. (2016); Agüera y Arcas et al. (2018)). While the *image pre-processing* and *fine-tuning* steps described in Appendix B.1 help account for such image heterogeneity, we go one step further and manually assess all pictures along seven dimensions: *logo*, *side face*, *professional*, *magazine*, *natural*, *natural lighting*, and *glasses*. For *logo*, we construct an indicator variable that takes value 1 if there is a logo (for instance, the "gettyimages" logo) on the face in the picture. For *side face*, the indicator is 1 if the CEO in the picture shows a side face instead of front. For *professional*, the indicator takes on 1 if the CEO is in work mode, say wearing business clothes, and 0 if in casual mode, say wearing a short-sleeved shirt, T-shirt, etc. For *magazine*, the variable takes on 1 if the picture is from a magazine cover. For *natural*, the variable reflects whether the CEO expects the picture or not, i. e., whether it is natural posing or a photo call. For *natural lighting*, the variable reflects whether the lighting feels natural (with light from all directions) or unusual, e.g., black and white, stage lighting, etc. The variable *glasses* takes

on 1 if the CEO in the picture wears glasses.

By controlling for all of these variables in our estimations, we further alleviate concerns about spurious correlations between picture characteristics and changes in apparent age.

B. Difference-in-Differences Results

We formalize our analysis of job-induced apparent aging in a difference-in-differences design, analyzing differences in visible signs of aging between CEOs whose industry was in distress during the financial crisis in 2007 and 2008 versus those whose industry was not in distress. As detailed in Section II.C, we use three-digit SIC codes and a 30% decline in equity value criterion to identify firms that experienced an industry shock during the crisis years. This approach classifies 79 out of a total of 149 industries represented in the CEO Apparent Aging Data as distressed during at least one of the crisis years 2007 and 2008. Industries classified as distressed during these years include real estate and banking. Non-distressed industries include agriculture, food products, and utilities.

To account for CEOs departing from their job during the Great Recession, potentially introducing selection bias, we identify treated CEOs based on intended exposure. That is, we define the treatment variable, $Industry\ Distress_j$, as equal to 1 if CEO j's firm operates in an industry that was distressed in 2007, 2008, or both years, regardless of whether the CEO stepped down between 2006 and 2008. In particular, $Industry\ Distress_j$ is encoded as 1 for a CEO departing in 2007 and whose firm's industry was distressed in 2008.[17]

We start from plotting the difference in aging trends between distressed and non-distressed CEOs in Figure 4. For this graphical illustration, we bin our data into nine roughly equal-sized groups of pictures from the beginning of the sample period to the end,

[17] Regressing actual 2007-2008 industry shock exposure on intended exposure yields a coefficient of 0.92 (F-statistic of 331.66).

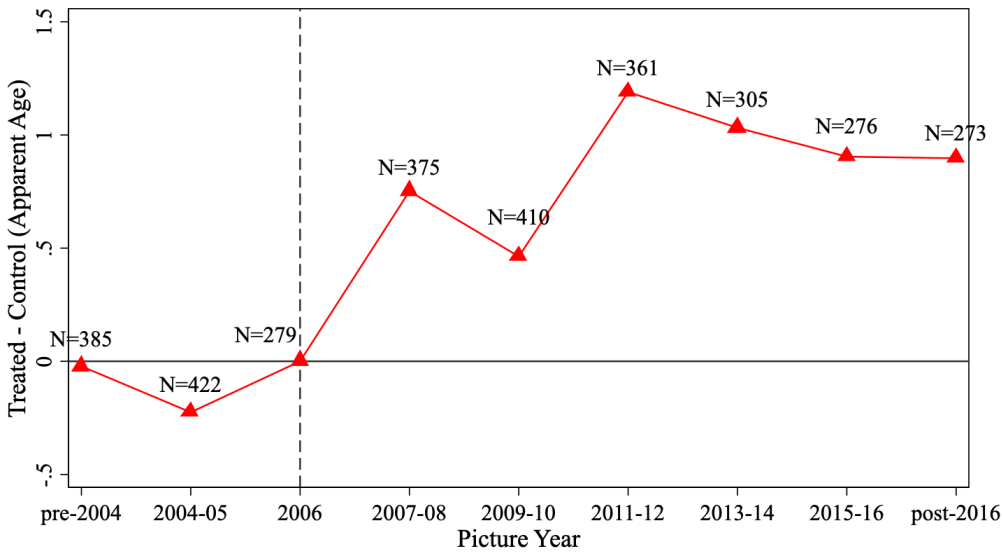


Figure 4.—Differences in apparent aging between CEOs with and without industry distress exposure during the Great Recession.

Notes: This figure depicts the estimated coefficients β_2 of the interaction terms between the time-period indicators and the $Industry\ Distress$ indicator from estimating equation (2.1), where $Industry\ Distress$ is equal to 1 if the CEO's firm was exposed to industry-wide distress during 2007 or 2008. N denotes the number of pictures for each time period. We winsorize the estimated apparent-age variable by first winsorizing the top and bottom 0.5% of the difference between apparent and biological age, and then adding this winsorized difference to the biological age. Observations are weighted by the inverse of the number of pictures per CEO.

$t \in T = \{\text{pre-2004}, 2004\text{-}05, \text{ ..., post-2016}\}$, and estimate the following difference-in-difference model:

$$\begin{aligned} \textit{Apparent Age}_{i,j,t} = \ & \beta_0 + \beta_1 \textit{ Biological Age}_{i,j,t} + \sum_{t \in T} \beta_{2,t} \textit{ Industry Distress}_j \times 1_t \\ & + \beta_3' X_{i,j,t} + \delta_t + \theta_j + \varepsilon_{i,j,t} \end{aligned} \tag{2.1}$$

where i represents a picture, j represents a CEO, and t represents a time bin. 1_t are time indicators, where the t^{th} indicator is equal to 1 for pictures taken at time t. They are interacted with $Industry\ Distress_j$, so that the interaction is 1 if the firm of CEO j shown in picture i was distressed in 2007 or 2008. The vector of control variables $X_{i,j,t}$ includes the number of industry shocks a CEO experienced before 2006 and CEO tenure until 2006. We also include CEO fixed effects θ_j and time fixed effects δ_t. The CEO fixed

effects absorb any time-invariant CEO facial characteristics such as facial shape. The time fixed effects absorb time trends, such as improving picture quality. While the aging software has been trained on a large number of faces and pictures of differing quality, these fixed effects tighten the identification further (and absorb the main effects of the time-industry shock interaction in the regression). We note that for either of these variables to potentially affect the estimation in the first place, they would have to systematically affect the software's age estimate (rather than introducing noise) and be correlated with industry distress experience. As discussed above, we additionally include extensive controls for picture setting and characteristics.

Figure 4 plots the estimates of vector $\beta_2 = (\beta_{2,\text{pre-2004}}, ..., \beta_{2,t}, ..., \beta_{t,\text{post-2016}})$, capturing the apparent-age differences between the treated group and the control group at the different points in time, after controlling for the biological age and other covariates. We see that the difference in apparent age between future distressed and non-distressed CEOs is small and stable over time before the crisis, consistent with the notion that aging in both groups follows parallel pre-trends. After the onset of the Great Recession, however, the apparent-age difference increases markedly, first to about half a year, and then to a full year. It stays and stabilizes at a high level of about one year of apparent-age difference after around five years post-crisis. In other words, exposure to industry distress significantly accelerates aging over the next few years, with the apparent-age difference stabilizing at one year.

The large estimated difference in aging post-crisis is robust to estimating the standard difference-in-differences regression model:

$$\begin{aligned} \textit{Apparent Age}_{i,j,t} =& \beta_0 + \beta_1 \textit{Biological Age}_{i,j,t} + \beta_2 \textit{Industry Distress}_j \times 1_{\{t>2006\}} \\ &+ \beta_3' X_{i,j,t} + \delta_t + \theta_j + \varepsilon_{i,j,t} \end{aligned} \tag{2.2}$$

where i represents a picture, j represents a CEO, and t represents a calendar year. We

continue to code $Industry\ Distress$ as an indicator of intended industry-distress exposure during the Great Recession to account for possible selection bias. The vector of control variables, $X_{i,j,t}$, is the same as in estimating equation 2.1, and δ_t and θ_j capture the year and CEO fixed effects, respectively. The key coefficient of interest is β_2, indicating the difference in how old CEOs look in post-crisis years depending on whether they personally experienced industry shocks during 2007 to 2008.

Table III presents the regression results. In column (1), the coefficient on the interaction term between $Industry\ Distress$ and the post-2006 indicator, $1_{\{t>2006\}}$, is 0.948, indicating that CEOs look around one year older during and post-crisis if they experienced industry distress shocks between 2007 and 2008. In column (2), we add the extensive set of picture controls described above ("logo," "side face," "professional," "magazine," "natural," "natural lighting," and "glasses"). This barely changes the coefficient on the post-treatment interaction term (now 0.978, significant at 5%).

In columns (3) and (4), we split the post-period into two sub-periods, capturing pictures taken between 2007 and 2011 and since 2012, respectively. Our estimates indicate that longer-horizon effects are driving our results: in this specification, we estimate a distress-induced apparent-aging effect of around 0.8 years over a five-year horizon that is insignificant but increases to about 1.2 years over longer horizons and becomes significant. Again, the estimated effects are very similar whether or not we include the additional picture controls.

The fact that CEO aging effects appear to be permanent also ameliorates potential concerns about the media or firms engaging in "picture management" that is correlated with distress exposure and could thus affect the apparent-age estimates. For example, the media might select pictures that show the CEO in distress (and hence possibly looking older) during the crisis, while firms might engage in "CEO appearance management" and try to put forward particularly positive pictures. The significant (and larger) coefficient over the longer horizon helps dispel such concerns. By 2012, it is unlikely that variation in distress

exposure during the 2007/08 financial crisis still affects picture selection, especially since more than 50% of CEOs have departed from their position.

In sum, the apparent-aging analysis provides robust evidence that increased job demands in the form of industry distress accelerate visual aging. Our other results on CEO mortality in Sections 2.4 and 2.5 below suggest that the appearance of aging may presage a shorter lifespan for CEOs whose industries experienced downturns in the Great Recession.

C. Robustness Tests

We perform a series of additional tests, with all tables relegated to Appendix B.2. First, we verify that all results are similar when we estimate the difference-in-differences model on the non-winsorized sample (Appendix-Table B.1). Second, we consider using a more restrictive distress definition that requires negative industry sales growth, as in the robustness tests in Acharya et al. (2007), and in Opler and Titman (1994) and Babina (2020). Around 29% of the CEOs are still classified as experiencing distress under this more stringent definition, and the estimation continues to yield economically and statistically significant aging effects (Appendix-Table B.2). If anything, the estimated effect of industry distress on apparent aging becomes slightly larger, with the differential-aging coefficient increasing from 0.948 to 1.173 in column (1) and from 0.978 to 1.064 in column (2). Moreoever, when splitting the post-crisis period into a five-year and a longer horizon in columns (3) and (4), we now estimate significant effects both for both subperiods.

Lastly, we verify that our results are not affected by differential finding rates of pictures depending on whether CEOs experienced distress during the crisis. For example, if experiencing industry distress shocks makes CEOs more likely to step down earlier, it may be more difficult to find recent, post-tenure pictures. Appendix-Figure B.3 depicts the average number of pictures per CEO we find in each year, split by whether a CEO experienced industry distress shocks in 2007-2008. In general, the finding rates closely follow

each other over time, though there is a small divergence after 2015. Therefore, we repeat our analysis restricting our sample to the years up to 2015, as shown in Appendix-Table B.3. The size and significance of the coefficients on the interaction terms remain similar across all columns.

2.4. Industry-Wide Distress Shocks and Life Expectancy

In the remaining two parts of the analysis, we move from apparent aging to mortality. This section examines the link between mortality and industry distress, and Section 2.5 the link between mortality and governance regulation. For both sets of analyses, we use the CEO Mortality Data described in Section II.B, which allows us to analyze mortality outcomes and accommodates the timing of variation in anti-takeover laws.

A. Empirical Strategy

We employ the Cox (1972) proportional hazards model to estimate the effect of variation in job demands on longevity. CEOs enter the analysis ("become at risk") in the year they are appointed, and they exit at death or the censoring date. We capture the effect of variation in CEOs' exposure to industry distress by estimating

$$\ln\ \lambda(t|Industry\ Distress_{i,t}, X_{i,t}) = \ln\ \lambda_0(t) + \beta\, Industry\ Distress_{i,t} + \delta' X_{i,t}, \quad (2.3)$$

where λ is the hazard rate and λ_0 is the baseline hazard rate. $Industry\ Distress_{i,t}$ is an indicator variable equal to 1 if CEO i has experienced distress, i. e., an industry-wide 30%-decline in equity value over a two-year horizon, by year t. Note that, when a CEO steps down, the value of the distress indicator remains constant at its value at departure. $X_{i,t}$ is a vector of control variables. In our main specifications, we consecutively add as controls biological age, time trends (linear or fixed effects), and firm location (state of headquarters) and industry fixed effects. We later present robustness specifications

with birth-cohort fixed effects and birth-cohort-specific age controls. We cluster standard errors at the three-digit SIC code level, at which industry shocks are defined (Abadie et al. (2017)).

B. Graphical Evidence

Before presenting the main results, we provide graphical evidence on the mortality effects of variation in CEOs' industry distress exposure. Figure 5 plots Kaplan-Meier survival graphs, split by whether CEOs experienced industry distress during their tenure. The non-parametric Kaplan-Meier estimator discretizes time into intervals $t_1, ..., t_J$, and is defined as $\widehat{\lambda_j^{KM}} = \frac{f_j}{r_j}$, where f_j is the number of spells ending at time t_j and r_j is the number of spells that are at risk at the beginning of time t_j. We plot unadjusted survival in Figure 5, where the vertical axis shows the survival rate, and the horizontal axis the time elapsed (in years) since becoming CEO. In lieu of age controls (which are included in the regression analysis), we limit the sample to CEOs appointed up to ten years prior to the retirement age of 65 and stepping down up to ten years after the retirement age.[18]

Figure 5 shows that CEOs who were exposed to industry distress during their tenure have significantly worse long-run survival patterns than never-distressed CEOs. The survival line for distressed CEOs included in the figure is visibly left-shifted at longer horizons. For example, 25 years after a CEO's appointment, about 50 percent of distressed CEOs have passed away, whereas this takes closer to 30 years for the non-distressed CEO group.

The survival plot offers first, suggestive evidence that experiencing industry distress as CEO is associated with adverse consequences in terms of life expectancy. Our hazard analysis below formalizes the observed patterns.

[18] Appendix-Figure C.1 plots the retirement hazard for CEOs in our main sample, confirming a large spike in retirements at age 65, consistent with the evidence in Jenter and Lewellen (2015). While CEOs may continue to work after stepping down, few (34 in total) become CEO at another firm in our sample.

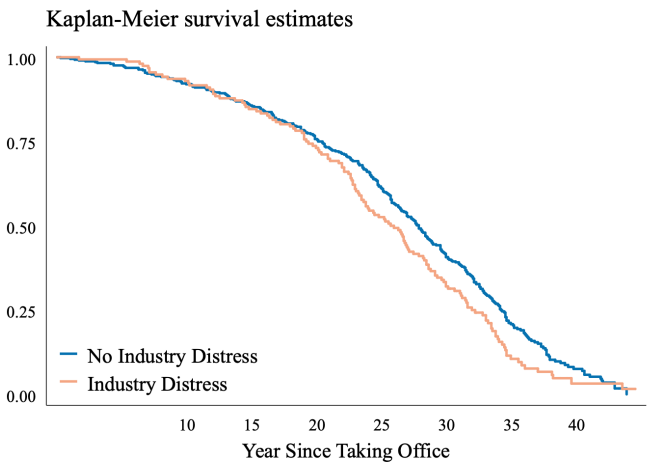


Figure 5.—Kaplan-Meier survival estimates.

Notes: This figure shows Kaplan-Meier survival plots for the relation between industry distress experience and longevity.

The vertical axis shows the fraction of CEOs who are still alive. The horizontal axis reflects time elapsed (in years) since a person became CEO. The figure compares the survival of CEOs who never experienced industry-wide distress during their tenure (dark blue) to those who experienced distress (orange). The figure is limited to CEOs appointed up to ten years prior to the retirement age of 65 and stepping down up to ten years after the retirement age.

C. Main Results

Table IV shows the hazard model results on the relationship between industry distress and CEOs' mortality rates, based on estimating equation (2.3). All columns show the estimated hazard coefficients (β, δ), such that a coefficient greater than zero indicates that the risk of failure (death) is positively associated with a given variable.

The specification in column (1) estimates the model with solely the industry-distress indicator and a linear control for age, without including any fixed effects. In column (2), we add a linear control for time trends as well as location fixed effects based on firms' state of headquarters, which absorb state-level characteristics such as general business conditions, pollution, and eating habits to the extent that these are time-invariant. In both specifications, the estimated hazard coefficient on the distress indicator is statistically and economically significant, amounting to 0.123 and 0.145, respectively. The results do not

change when we make comparisons within Fama and French (1997) 49 industries in column (3) and replace the linear time control with year fixed effects in column (4). The coefficient estimates on distress exposure are almost unchanged, 0.146 and 0.137, both significant at 5%. Averaging the estimates across columns, distress experience increases CEOs' log mortality hazard by 0.138, i.e., the mortality hazard by approximately 13.8%.[19]

Turning to the control variables, the effect of age is significantly positive across specifications, reflecting that older people have a higher risk of dying. Note that the linear age term is motivated by the Gompertz (1825) "law of mortality," i. e., the empirical regularity that the risk of dying follows a geometric increase after middle age (Olshansky and Carnes (1997)). In untabulated results, we obtain virtually identical estimates when including higher-order age terms. The linear time control in columns (2) and (3) is close to zero and insignificant, suggesting no general time trends in the survival of CEOs over the sample period. Nevertheless, we further address concerns around improvements in life expectancy over time in Section IV.D, where we test and confirm the robustness of our results to including birth-year fixed effects and age-birth-cohort interactions.

Economic Significance. One way to evaluate the magnitude of the estimated distress effect on longevity is relative to other predictors of CEO life expectancy in our hazard model, in particular CEO age: How much "older" does a CEO become due to distress exposure in terms of mortality hazard? The estimated effect of age on death hazard in our sample (averaging across all columns) is 0.118, which implies an 11.8% increase in the mortality hazard per year of age. This means that the effect of industry distress exposure corresponds to the effect of a 1.2-year shift in age ($0.138/0.118 \approx 1.2$).

This "in-sample" comparison has the advantage that it is directly based on data from the sample CEOs. Alternatively, we can compare our estimated distress hazard with mortality statistics of the general U.S. population, acknowledging that statistics derived for

[19] Without approximating, the average coefficient of 0.138 implies a hazard ratio of $\exp(0.138) = 1.148$, i. e., a 14.8% increase in the mortality hazard.

high SES groups would be ideal. For example, at age 57 (the median CEO age in our sample), the one-year mortality rate of a male American born in 1925 (the median birth year in our sample) is 1.366% (Human Mortality Database (2019)). Industry distress experience as CEO pushes this rate up to 1.5709%, which is roughly the mortality rate of a male born in 1925 at age 58.5, i. e., when one and a half years older. Also with respect to the general population, Sullivan and Von Wachter (2009) estimate that job displacement at age 40 increases the mortality hazard by 10–15% and reduces life expectancy by one to one and a half years.

Another benchmark for comparison are other known health threats. For example, smoking until age 30 is associated with a reduction in longevity by roughly one year (Jha et al. (2013)). The reduction in life expectancy from industry distress exposure is thus slightly larger than the reduction from smoking in the first three decades of one's life.

In sum, unexpected changes in the work environment and job demands of CEOs arising from industry-wide distress have substantial health consequences not only in terms of short- to medium-term visible aging but also in terms of long-term mortality.

D. CEO Birth Cohort Robustness

As alluded to above, one concern with the results from Table IV may be that the year and age controls are insufficient to control for trends in longevity (even though we find the linear time control and higher-order age controls to be insignificant). To further examine potential confounds around heterogeneity in life expectancy over time, and this being correlated with industry distress exposure, Table V estimates alternative specifications accounting for CEO birth cohorts. Specifically, Columns (1) and (2) directly include CEO birth-year fixed effects in lieu of year controls or fixed effects. The industry distress coefficients remain similar compared to those in Table IV and, if anything, become statistically and economically slightly more significant. For example, the estimation in

column (2) implies a distress-induced mortality effect equivalent to a 1.5-year increase in age. Columns (3) and (4) revert to year fixed effects but allow the effect of age on mortality to be cohort-specific: we sort CEOs into quintiles based on birth year, and allow for separate age estimates. While there are small differences in age effects across CEO cohorts, the industry distress coefficients are again barely affected and remain statistically and economically significant.

E. Additional Robustness Tests

We perform several additional robustness tests, with all tables relegated to Appendix 2. First, we re-estimate equation (2.3) when including as additional control variables CEO pay (from Gibbons and Murphy (1992)) and firm size (assets and employees from CRSP and Compustat). The estimated hazard coefficients on distress exposure become slightly larger and remain significant (Appendix-Table C.1). In terms of economic magnitude, the distress effect is now roughly equivalent to being 1.3 years older.

Second, we re-estimate the model on an extended sample that includes the 295 CEOs we had dropped from the analysis in the final step of Section II.B (Appendix-Table C.2, with Panel A including the baseline controls and Panel B the additional controls from above). The estimated coefficients remain similar throughout and the estimated distress effect, averaging across columns, is now equivalent to being 0.9 years older.

We also estimate the effect of industry distress when varying the censoring year for CEOs' alive status. This check can help alleviate concerns that we may have failed to identify some deaths, hence granting "extra years" of life to these CEOs (cf. footnote 11). The coefficients remain stable as we gradually move the censoring date from Oct. 1, 2017 to Dec. 31, 2010 (Appendix-Figure C.2).

Finally, we have also explored specific recession periods similar to the analysis in Section 2.3, such as the 1987 stock-market downturn or the 1981-82 recession. However, fewer than five percent of the CEOs in our sample experienced either of these shocks so

that we lack statistical power when applying the same methodology of comparing CEOs who did and did not experience these industry downturns in their firms. While the corresponding estimates indicate that CEOs who experienced these shocks tend to have a higher mortality hazard, they are not statistically significant. We have also again considered the more restrictive return- and sales-based distress definition from Section III.C. In our CEO mortality sample preceding the Great Recession, this definition classifies fewer than five percent of CEOs as distressed and substantially increases standard errors. Nonetheless, we estimate similar effect sizes as above, with the average point estimate across specifications implying a distress effect that corresponds to a seven-month increase in age.

2.5. Corporate Monitoring and Life Expectancy

In the final step of our analysis, we exploit the staggered passage of anti-takeover laws across U.S. states in the mid-1980s as the source of identifying variation to study CEO mortality effects.

A. Empirical Strategy

Our main analysis continues to use the Cox (1972) proportional hazards model. First, we estimate a modified version of equation (2.3):

$$\ln\ \lambda(t|BC_{i,t}, X_{i,t}) = \ln\ \lambda_0(t)\ + \beta\, I(BC_{i,t}) + \delta' X_{i,t}, \tag{2.4}$$

where $I(BC_{i,t})$ is an indicator variable equal to 1 if CEO i has been exposed to a BC law by year t. As in Section IV.A, $X_{i,t}$ includes biological age, time trends (or fixed effects), and state-of-headquarters and industry fixed effects. We also verify again that all of our results are robust to specifications that account for CEO birth cohorts.

Second, we test for differential effects depending on differential BC law exposure intensity. Given that the laws led to a permanent corporate governance regime change, rather

than a temporary exposure (as in the case of industry distress), we replace the indicator $I(BC_{i,t})$ with a measure $BC_{i,t}$ that counts the exposure length in years until year t:[20]

$$\ln\ \lambda(t|BC_{i,t}, X_{i,t}) = \ln\ \lambda_0(t)\ + \beta\, BC_{i,t} + \delta' X_{i,t}. \tag{2.5}$$

As with the distress indicator in the previous section, the values of the BC law indicator and BC law exposure length variables remain constant once a CEO has stepped down. We now cluster standard errors at the state-of-incorporation level, given that the BC laws applied based on firms' state of incorporation (Abadie et al. (2017)).

B. Graphical Evidence

We again start by plotting Kaplan-Meier survival graphs, separating between BC-insulated and non-insulated CEOs. As in Figure 5, we focus on CEOs who started and ended their tenure within a ten-year period around the retirement age of 65.

Panel (a) of Figure 6 plots the survival lines of this set of CEOs comparing those who became CEO in the 1970s and were never shielded by a BC law, those who became CEO in the 1980s and were never shielded by a BC law, and those who became CEO in the 1980s and were eventually insulated by BC law protection during their tenure.[21]

Two results emerge. First, the survival patterns of the 1970s and 1980s cohorts without BC exposure are remarkably similar, allaying concerns that our BC-law-mortality results pick up *general* changes in survival patterns between the 1970s and 1980s. Second, the survival line for the 1980s cohorts with BC exposure is visibly right-shifted compared to

[20] We calculate this measure up to daily precision levels. For example, Delaware's BC law was adopted on 2/2/1988, and a CEO's exposure in Delaware in 1988 is calculated as $BC_{i,1988} = \frac{365-\text{doy}(2/2/1988)}{365} = 0.92$. Relative to the binary measure, the cumulative exposure measure is more prone to endogeneity concerns for long-serving CEOs. We address this concern in Section V.E, where we examine initial vs. incremental exposure as well as predicted exposure length.

[21] For the 1970s cohorts, maximal elapsed time since our sample start is $t = 47.75$ (time elapsed between 1/1/1970 and the censoring date, 10/1/2017). Similarly, for the 1980s cohorts, maximal elapsed time is $t = 37.75$. We restrict the graph to periods when at least 30 CEOs in either cohort group are uncensored, explaining the slightly differential ends of the survival lines (after 36 and 45 years, respectively).

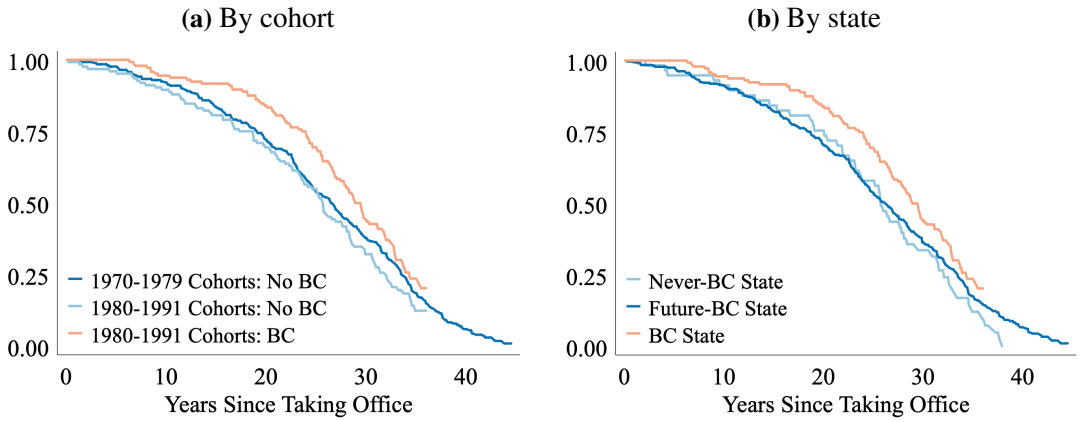


Figure 6.—Kaplan-Meier survival estimates.

Notes: This figure shows Kaplan-Meier survival plots for the relation between anti-takeover law protection and longevity. The vertical axis shows the fraction of CEOs who are still alive. The horizontal axis reflects time elapsed (in years) since a person became CEO. Panel (a) compares the survival of CEOs starting in the 1970s who never served under a BC law (light blue) to those who became CEO in the 1980s and never served under a BC law (dark blue) and those who became CEO in the 1980s and were eventually exposed to a BC law (orange). Panel (b) splits the CEOs from Panel (a) based on whether their state never passed a BC law (light blue), passed a BC law after the CEO stepped down (dark blue), or passed a BC while in office. As in Figure 5, both panels focus on CEOs appointed up to ten years prior to the retirement age of 65 and stepping down up to ten years after the retirement age.

the No-BC-cohorts. For example, 20 years after a CEO's appointment, about 25 percent of CEOs in the 1980s cohorts without BC exposure have died, whereas it takes closer to 25–30 years for CEOs in the 1980s cohorts with BC exposure.

One possibility is that the patterns in Panel (a) might pick up systematic differences between BC and non-BC states—despite the fact that these laws apply based on state of incorporation as opposed to firms' location. To examine this graphically, Panel (b) reshuffles CEOs in Panel (a)'s No-BC-cohorts, grouping them instead by whether their state eventually enacted a BC law after the CEO stepped down (dark blue) or not (light blue). The survival lines for these groups are virtually identical, and only CEOs in BC states *with* BC exposure (orange) show a more beneficial survival curve. Thus, there is no evidence of BC states being inherently different prior to BC enactment. We also note that all our estimations will include location fixed effects and are robust to using state of

incorporation fixed effects.

The survival plots suggest significant adverse consequences in terms of life expectancy associated with serving under more stringent corporate governance regimes. We next turn to the hazard model based results that control for other determinants of mortality.

C. Main Results

Table VI shows the hazard model results relating BC law exposure to CEO mortality rates based on the estimating equations (2.4) and (2.5). In Columns (1) through (4), we summarize the total effect of the BC laws with the indicator $I(BC_{i,t})$ for CEO i having been exposed to a BC law by time t. In Columns (5) through (8), we estimate the linear (in hazards) effect in years of exposure to more lenient corporate governance, $BC_{i,t}$. The control variables and fixed effects in the left four and right four columns are the same as in the distress-mortality analysis of Table IV.

Across all columns, we estimate a statistically and economically strong effect of BC law protection on mortality. For the BC indicator specifications, the hazard coefficients range from -0.225 to -0.269 as we move from just a linear control for age in column (1) to specifications with linear or fixed effects time controls, state of headquarters fixed effects (recall that BC laws apply instead based on firms' state of incorporation), and industry fixed effects in columns (2) to (4). For the cumulative BC exposure measure in columns (5) to (8), the coefficients range from -0.041 to -0.046, indicating that a one-year increase in exposure to more lenient governance is estimated to reduce a CEO's mortality risk by 4.3% averaging across columns. For a CEO with a typical BC law exposure, both BC exposure measures imply very similar effects on longevity.[22]

The estimated effects of control variables on mortality rates mirror those in Table IV, with the linear time control being insignificant and age strongly predicting mortality rates.

[22]The cumulative measure estimates a 19% shift in mortality hazard associated with the median BC exposure of 4.4 years ($4.3\% \times 4.4 = 18.92\%$), close to the 22-27% shift estimated with the BC indicator.

As before, adding higher-order age terms yields nearly identical results (untabulated).

Economic Significance. The estimates imply meaningful effect sizes. First, we use again the "in-sample" approach from Section IV.C. Based on the hazard coefficient estimates from the most conservative specification in column (4), -0.252 for BC exposure and and 0.117 for age, the effect of BC law protection on mortality is equivalent to being about two years younger. This effect size is of similar order of magnitude, but slightly larger than the estimated effect size of exposure to industry distress, which corresponded to about 1.2 years in CEO age. The somewhat larger mortality effect might reflect the more permanent nature of the BC law experiences relative to industry shock experiences.

Second, we use the general U.S. population life tables approach. Here, the median exposure to lenient governance of 4.4 years pushes the 1.366% mortality rate of males born in 1925 rate down to 1.119%, which is roughly the mortality rate of males born in 1925 at age 54, i. e., when three years younger.

Overall, we find variation in job demands based on anti-takeover laws to be associated with substantial mortality effect sizes, also compared to other determinants of longevity and known health risks.

D. CEO Birth Cohort Robustness

Mirroring Section IV.D, we estimate alternative specifications that account for different CEO birth cohorts.

This is of particular importance in the context of the BC law analysis, in light of the fact that BC laws disproportionately protected more recent CEOs who were born later and are thus younger on average.

We re-estimate estimating equations (4) and (5) including CEO birth-year fixed effects (columns (1) and (2)) and age-by-cohort controls (columns (3) and (4)). As shown in Table VII, the hazard coefficients on the indicator and cumulative BC measures are barely

affected when adding these alternative fixed effects and controls, imply very similar effect sizes, and continue to be significant at 5% or 1%. The robustness of the results alleviate potential concerns around differences between BC-protected and non-protected CEOs resulting from systematic age or cohort differences.

E. Additional Robustness Tests

Our results are robust to a series of additional tests, some analogous to the robustness checks of the distress-longevity relation in Section IV.E, and some specific to the use of anti-takeover laws as identifying variation. For brevity, we only provide a brief overview of the tests here and present a detailed discussion of the tests, where applicable, in Appendix 2.

Other Specifications and Sample Choices. Mirroring the robustness tests in Section IV.E, we test and confirm that our results are robust to including CEO pay, firm assets, and number of employees as additional control variables (Panel A of Appendix-Table D.1), and robust to different censoring date choices (Appendix-Figure D.2). While it is standard in the BC law literature to assign location fixed effects based on firms' headquarters (cf. Gormley and Matsa (2016)), the results are also robust to using state-of-incorporation fixed effects (Panel B of Appendix-Table D.1). The stability of the estimates suggests that firm locations do not affect the relationship between governance and longevity. Alternatively, the results are robust to keeping state of headquarters fixed effects but dropping CEOs who stepped down significantly before the passage the BC laws (Appendix-Figure D.3), which further ameliorates potential concerns specific to the BC law results around differences between BC-protected and non-protected CEOs.

Other Anti-Takeover Laws. The results are also robust to using the first-time enactment of any of the five second-generation anti-takeover laws as identifying variation (Appendix-Table D.2). This test highlights that our results should be interpreted more broadly, apply-

ing to different corporate governance mechanisms rather than narrowly to BC laws.

Karpoff–Wittry and Related Tests. All results are robust to extensive robustness checks proposed in Karpoff and Wittry (2018) to account for firms lobbying for the passage of BC laws or opting-out, as well as confounding effects of firm-level defenses and first-generation anti-takeover laws (Appendix-Tables D.3 and D.4). Additionally, the results are robust to data cuts based on state of incorporation and industry affiliation (Appendix-Table D.5).

Nonlinear Exposure Effects and Predicted Length of Exposure. The final two robustness checks address concerns that the cumulative BC specification in equation (2.5) picks up CEOs' endogenous selection into a long tenure. We first address this by examining separately the mortality effects of initial years and incremental years of BC law exposure. Columns (1) to (4) in Appendix-Table D.6 re-estimate the last four columns of Table VI, splitting $BC_{i,t}$ into below- and above-median exposure, $BC_{i,t}^{(\min\text{-}p50)}$ and $BC_{i,t}^{(p51\text{-}\max)}$, with the above-median exposure variable picking up incremental exposure in addition to initial exposure.[23] Across columns, the hazard coefficient on below-median BC exposure is strongly significant. By contrast, the coefficient on above-median BC exposure is close to one and insignificant. These results imply that the estimated survival gains are driven by the initial years of reduced monitoring, rather than by the tails of long-tenured CEOs.

To further address any remaining endogeneity concerns related to BC law exposure intensity, we estimate a hazard model using a CEO's predicted, rather than true, length of BC exposure. The prediction model only uses information from prior to the BC law passage and, in a nutshell, predicts CEOs' remaining tenure over time (from which predicted BC exposure is derived; see the appendix for full details). The results in columns (5) to (8) of Appendix-Table D.6 corroborate our baseline findings. Predicted BC exposure is

[23] For example, for a CEO with a current BC exposure of four years, $BC_{i,t}^{(\min\text{-}p50)}$ would take the value 4, and $BC_{i,t}^{(p51\text{-}\max)}$ the value 0. In the following year $(t+1)$, $BC_{i,t+1}^{(\min\text{-}p50)}$ would be set to 4.4, and $BC_{i,t+1}^{(p51\text{-}\max)}$ to 0.6. In year $t+2$, $BC_{i,t+2}^{(\min\text{-}p50)}$ remains at 4.4, and $BC_{i,t+2}^{(p51\text{-}\max)}$ increases to 1.6.

estimated to significantly affect CEOs' mortality rates. The estimated hazard coefficients range from -0.049 to -0.059 and are very similar to those in Table VI. While the standard errors, now bootstrapped because we are using a generated regressor, are slightly larger compared to Table VI, the coefficient of interest remains significant in all columns, either at 1% or 5%. A regression of true BC exposure on predicted exposure yields a coefficient of 1.21, which indicates that the prediction well approximates the true exposure. The estimated effects remain sizable if we divide them by 1.21. For instance, using the coefficient in column (8) of Appendix-Table D.6 yields $-0.049/1.21 = 0.040$.[24]

F. Business Combination Laws and CEO Pay

The permanent corporate governance regime changes induced by BC laws also lend themselves to studying the extent to which managers account for the health implications of their jobs, for example in negotiating pay.

As laid out in detail in Appendix 2, we conduct a simple calibration exercise that builds on the literature on the value of a statistical life (Viscusi and Aldy (2003)). We calculate that, if CEO pay reflects working conditions, then a reduction in mortality risk of 4.1% per year of BC exposure (as estimated in column (5) of Table VI) would imply a CEO pay change between -2% and -9%.[25]

By contrast, when we turn from the theoretical calibration to the empirical relationship between BC law protection and pay, we estimate a positive, albeit mostly insignificant effect of BC law passage on pay (Appendix-Table D.7). The estimates indicate a pay increase of around 4.1–8.7%.[26] The apparent lack of a compensating differential casts doubt on whether all parties fully account for the health implications of different governance

[24] In Appendix 2, we also analyze how tenure responds to anti-takeover laws. We find suggestive evidence of moderate increases in tenure under BC laws, consistent with the laws affecting managers' perceptions of job demands.

[25] We thank Xavier Gabaix for suggesting this calibration exercise.

[26] In comparing the results to the earlier work (Bertrand and Mullainathan (1998)), which estimated a (more significant) 5.4 percent pay increase, it is important to note that our analysis is conducted on our CEO Mortality Data, a CEO-level sample, and restricts the sample to incumbent, pre-BC CEOs.

regimes.

2.6. Conclusion

In this paper, we assess the health consequences of being exposed to increased job demands and a more stressful work environment while in a high-profile CEO position. We analyze the consequences for CEOs' aging and mortality using two sources of variation in job demands, industry-wide distress shocks and the staggered introduction of anti-takeover laws.

We first show that industry distress is reflected in short- to medium-term signs of adverse health consequences, namely faster visible aging. To the best of our knowledge, we are the first to collect and utilize panel data of facial images and apply machine-learning based apparent-age estimation software in social-science research. Implementing a difference-in-differences design that exploits variation in industry distress during the Great Recession, we estimate that CEOs who experienced industry distress during the 2007-2008 financial crisis look roughly one year older than those whose industry did not suffer the same level of distress. The effect of distress on aging becomes slightly larger over time, increasing to 1.178 years if we analyze pictures from 2012 and afterwards.

We then document, using an earlier CEO sample, more long-term adverse health outcomes associated with strenuous job demands. CEOs who experienced periods of industry-wide distress during their tenure die significantly earlier. We estimate a mortality effect corresponding to that of a 1.2-year increase in biological age. In line with these results, we observe significant improvements in life expectancy for CEOs who became shielded by an anti-takeover law during their tenure.

In sum, our results indicate that financial distress and stricter corporate governance regimes—the latter of which are generally viewed as desirable and welfare-improving—impose significant personal health costs to CEOs. While we lack direct physical or medical

measures of heightened stress, the evidence implies that economic downturns and stricter governance constitute a substantial personal cost for CEOs in terms of their health and life expectancy. As such, our findings also contribute to the literature on the trade-offs between managerial incentives and private benefits arising from the separation of ownership and control. We document and quantify a previously unnoticed yet important cost—personal health cost—associated with serving under strict corporate governance.

Our findings suggest further avenues of investigation. One open question is whether managers fully account for these personal costs as they progress in their careers and how these costs affect selection into service as a CEO. Are there other dimensions of compensation which may respond to these job demands? Are some high-ability candidates for a Forbes-level CEO career more aware of these consequences than others and select out? Additionally, which jobs and hierarchy levels come with the largest adverse health consequences, also in light of looming financial hardships?

Another promising avenue is the more fine-grained identification of stressors. What aspects of individual job situations and which decisions tend to have the largest adverse health consequences, for either management or regular employees: pending layoffs and downsizing; restructurings; hostile merger attempts? Likewise, heightened workplace stress can also adversely affect other aspects of life, including marriage, divorce rates, parenting, and alcohol consumption. We leave these topics for future research.

Table I
SUMMARY STATISTICS OF CEO APPARENT AGING DATA

This table shows summary statistics for the CEO apparent-aging analyses. *Industry Distress* during 2007-2008 is an indicator for distress experience during these years. *Industry Distress* pre-2006 counts the number of industry distress experiences prior to 2006.

	Panel A: Image-Level Statistics					
	N	Mean	SD	P10	P50	P90
Picture Year	3,086	2008.69	6.16	2003	2009	2016
Biological Age	3,086	58.33	8.29	48	58	69
Industry Distress (2007-2008)	3,086	0.66	0.47	0	1	1
Apparent Age	3,086	54.90	7.23	45	56	63
Logo in the Picture	3,086	0.18	0.38	0	0	1
Side Face	3,086	0.19	0.39	0	0	1
Professional Clothes	3,086	0.97	0.17	1	1	1
Magazine Shot	3,086	0.01	0.08	0	0	0
Natural Pose	3,086	0.70	0.46	0	1	1
Natural Lighting	3,086	0.16	0.37	0	0	1
Glasses	3,086	0.33	0.47	0	0	1
	Panel B: CEO-Level Statistics					
	N	Mean	SD	P10	P50	P90
No. of Pictures per CEO	463	7.35	4.51	3	6	13
Biological Age in 2006	463	55.54	6.55	47	56	63
Tenure (Pre-2006)	463	8.00	7.73	2	6	17
Industry Distress (2007-2008)	463	0.65	0.48	0	1	1
Industry Distress (Pre-2006)	463	0.54	1.13	0	0	2
	Panel C: Industry Distribution					
Industry (Number of CEOs)	Manufacturing (180)		Finance, Insur, Real Estate (65)			
	Retail (53)		Services (44)		Others (50)	
	Trans.; Commns.; Elec., Gas, and Sanitary Services (71)					

Table II
SUMMARY STATISTICS OF CEO MORTALITY DATA

	N	Mean	SD	P10	P50	P90
Birth Year	1,605	1925	8.96	1914	1925	1937
Dead (by October 2017)	1,605	0.71	0.45	0	1	1
Year of Death	1,140	2004	9.98	1989	2006	2016
Age at Death	1,140	81.95	9.92	67.58	83.42	93.50
Age Taking Office	1,605	51.63	6.95	43	52	60
Year Taking Office	1,605	1977	7.21	1968	1977	1986
Tenure	1,605	10.62	6.86	3	9.08	20
Industry Distress	1,605	0.40	0.49	0	0	1
BC	1,605	2.21	4.19	0	0	8.24
BC | BC>0	625	5.68	5.05	0.54	4.41	12.37

NOTE. — This table shows summary statistics for the CEO longevity analyses. All variables are defined at the CEO level. BC denotes years of exposure to business combination laws. *Industry Distress* is an indicator variable that equals one if a CEO experienced industry-wide distress during his tenure.

Table III
INDUSTRY DISTRESS AND CEO AGING

Dependent Variable: $Apparent\ Age_{i,j,t}$				
	(1)	(2)	(3)	(4)
Industry Distress $\times\ 1_{\{t>2006\}}$	0.948*	0.978**		
	[0.484]	[0.478]		
Industry Distress $\times\ 1_{\{2006<t<2012\}}$			0.790	0.799
			[0.533]	[0.525]
Industry Distress $\times\ 1_{\{t\geq 2012\}}$			1.178**	1.193**
			[0.547]	[0.538]
Biological Age	0.915***	0.910***	0.943***	0.938***
	[0.092]	[0.092]	[0.094]	[0.093]
CEO FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Picture Controls		Y		Y
Number of CEOs	463	463	463	463
Observations	3,086	3,086	3,086	3,086

NOTE. — This table shows OLS estimates of the effect of industry distress exposure during the Great Recession on CEO apparent age. We winsorize the estimated apparent age by first winsorizing the top and bottom 0.5% of the difference between apparent and biological age and then adding this winsorized difference to the biological age. We weight observations by the inverse of the number of pictures per CEO. Standard errors, clustered at the industry level, are shown in brackets. *, **, and *** denote significance at the 10, 5, and 1 percent level, respectively.

Table IV
INDUSTRY DISTRESS AND MORTALITY

Dependent Variable: $Death_{i,t}$				
	(1)	(2)	(3)	(4)
Industry Distress	0.123*	0.145**	0.146**	0.137**
	[0.065]	[0.062]	[0.066]	[0.066]
Age	0.112***	0.115***	0.123***	0.123***
	[0.006]	[0.006]	[0.007]	[0.007]
Year		0.000	-0.003	
		[0.006]	[0.006]	
Location FE (HQ)		Y	Y	Y
FF49 FE			Y	Y
Year FE				Y
Number of CEOs	1,605	1,605	1,605	1,605
Observations	50,530	50,530	50,530	50,530

NOTE. — This table shows hazard coefficients estimated from a Cox (1972) proportional hazards model. The dependent variable is an indicator that equals one if the CEO dies in a given year. The main independent variable $Industry\ Distress$ is an indicator of a CEO's exposure to industry distress shocks. All variables are defined in Appendix 2.6. Standard errors, clustered at the industry level, are shown in brackets. *, **, and *** denote significance at the 10, 5, and 1 percent level, respectively.

Table V
INDUSTRY DISTRESS AND MORTALITY: BIRTH COHORT ROBUSTNESS

Dependent Variable: $Death_{i,t}$				
	(1)	(2)	(3)	(4)
Industry Distress	0.160***	0.188***	0.139**	0.143**
	[0.061]	[0.067]	[0.063]	[0.067]
Age	0.120***	0.126***		
	[0.007]	[0.008]		
Age × Birth Cohort 1 (oldest)			0.087***	0.092***
			[0.011]	[0.011]
Age × Birth Cohort 2			0.086***	0.090***
			[0.012]	[0.012]
Age × Birth Cohort 3			0.084***	0.088***
			[0.012]	[0.012]
Age × Birth Cohort 4			0.080***	0.084***
			[0.013]	[0.013]
Age × Birth Cohort 5 (youngest)			0.075***	0.079***
			[0.015]	[0.014]
Location FE (HQ)	Y	Y	Y	Y
FF49 FE		Y		Y
Year FE			Y	Y
Birth Year FE	Y	Y		
Number of CEOs	1,605	1,605	1,605	1,605
Observations	50,530	50,530	50,530	50,530

NOTE. — This table shows hazard coefficients estimated as in Table IV but with birth-year fixed effects in lieu of year controls or fixed effects or allowing for birth cohort-specific age effects. Birth cohorts are defined by sorting CEOs into quintiles by birth year. The dependent variable is an indicator that equals one if the CEO dies in a given year. The main independent variable $Industry\ Distress$ is an indicator of a CEO's exposure to industry distress shocks. All variables are defined in Appendix 2.6. Standard errors, clustered at the industry level, are shown in brackets. *, **, and *** denote significance at the 10, 5, and 1 percent level, respectively.

Table VI
BUSINESS COMBINATION LAWS AND MORTALITY

Dependent Variable: $Death_{i,t}$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
I(BC)	-0.225***	-0.269***	-0.262***	-0.252***				
	[0.068]	[0.081]	[0.088]	[0.087]				
BC					-0.041***	-0.046***	-0.042***	-0.042***
					[0.006]	[0.005]	[0.006]	[0.006]
Age	0.107***	0.108***	0.116***	0.117***	0.105***	0.105***	0.114***	0.115***
	[0.005]	[0.006]	[0.004]	[0.004]	[0.006]	[0.006]	[0.005]	[0.005]
Year		0.005	0.002			0.005	0.001	
		[0.004]	[0.005]			[0.004]	[0.004]	
Location FE (HQ)		Y	Y	Y		Y	Y	Y
FF49 FE			Y	Y			Y	Y
Year FE				Y				Y
Number of CEOs	1,605	1,605	1,605	1,605	1,605	1,605	1,605	1,605
Observations	50,530	50,530	50,530	50,530	50,530	50,530	50,530	50,530

NOTE. — This table shows hazard coefficients estimated from a Cox (1972) proportional hazards model. The dependent variable is an indicator that equals one if the CEO dies in a given year. The main independent variables are a binary indicator of BC law exposure, $I(BC)$, in the left four columns and a count variable of years of exposure, BC, in the right four columns. All variables are defined in Appendix 2.6. Standard errors, clustered at the state-of-incorporation level, are shown in brackets. *, **, and *** denote significance at the 10, 5, and 1 percent level, respectively.

Table VII

BUSINESS COMBINATION LAWS AND MORTALITY: BIRTH COHORT ROBUSTNESS

DEPENDENT VARIABLE: $Death_{i,t}$				
	(1)	(2)	(3)	(4)
I(BC)	-0.256**		-0.231**	
	[0.100]		[0.108]	
BC		-0.045***		-0.035***
		[0.006]		[0.006]
Age	0.120***	0.117***		
	[0.004]	[0.005]		
Age × Birth Cohort 1 (oldest)			0.091***	0.088***
			[0.010]	[0.010]
Age × Birth Cohort 2			0.089***	0.086***
			[0.010]	[0.010]
Age × Birth Cohort 3			0.088***	0.084***
			[0.011]	[0.010]
Age × Birth Cohort 4			0.084***	0.081***
			[0.013]	[0.012]
Age × Birth Cohort 5 (youngest)			0.079***	0.076***
			[0.013]	[0.012]
Location FE (HQ)	Y	Y	Y	Y
FF49 FE	Y	Y	Y	Y
Year FE			Y	Y
Birth Year FE	Y	Y		
Number of CEOs	1,605	1,605	1,605	1,605
Observations	50,530	50,530	50,530	50,530

NOTE. — This table shows hazard coefficients estimated as in Table VI but with birth-year fixed effects in lieu of year controls or fixed effects or allowing for birth cohort-specific age effects. Birth cohorts are defined by sorting CEOs into quintiles by birth year. The dependent variable is an indicator that equals one if the CEO dies in a given year. The main independent variables are a binary indicator of BC law exposure, $I(BC)$, in columns (1) and (3), and a count variable of years of exposure, BC, in columns (2) and (4). All variables are defined in Appendix 2.6. Standard errors, clustered at the state-of-incorporation level, are shown in brackets. *, **, and *** denote significance at the 10, 5, and 1 percent level, respectively.

Appendix A. Variable Definitions

Variable Name	Definition
$Apparent\ Age_{i,t}$	How old CEO i looks in year t. The apparent age is estimated using a machine-learning based software by Antipov et al. (2016) that has been specifically developed for apparent-age estimation. See Appendix B.1 for additional detail.
$(Biological)\ Age_{i,t}$	CEO i's age in year t.
$Tenure_{i,t}$	CEO i's cumulative tenure (in years) at time t.
Birth Year	CEO's year of birth.
Dead (by Oct. 2017)	Indicator for whether a CEO has passed away by October 1st, 2017.
Year of Death	CEO's year of death, calculated up to monthly level (e.g. 2010.5 for a person who dies on 6/30/2010).
Age Taking Office	CEO's age when appointed as CEO.
Year Taking Office	Year in which a CEO is appointed.
$Industry\ Distress_{i,t}$	Indicator equal to 1 if CEO i is exposed to an industry shock by year t. Industry shock is defined as median two-year stock return (forward-looking) of firms in the same industry below -30%. As in Babina (2020), we (i) use SIC3 industry classes, (ii) restrict to single-segment CRSP/Compustat firms, i. e., drop firms with multiple segments in the Compustat Business Segment Database (CBSD), (iii) drop firms if the reported single-segment sales differ from those in Compustat by more than 5%, (iv) restrict to firms with sales of at least \$20m, and (v) exclude industry-years with fewer than four firms. We use firms' modal SIC across CRSP, Compustat, and CBSD, and the latter in case of a tie.
$I(BC_{i,t})$	Indicator equal to 1 if CEO i is insulated by a BC law in year t; remains at 1 in all subsequent years $\tau > t$, including after CEO departure.
$BC_{i,t}$	CEO i's cumulative exposure to a BC law during tenure up to time t (in years); remains constant after CEO departure.
$BC_{i,t}^{(\min\text{-}p50)}$	CEO i's below-median (4.4 years) cumulative BC law exposure during tenure up to time t (in years); remains constant after CEO departure.
$BC_{i,t}^{(p51\text{-}\max)}$	CEO i's above-median (4.4 years) cumulative BC law exposure during tenure up to time t (in years); remains constant after CEO departure.
$I(FL_{i,t})$	Indicator equal to 1 if CEO i is insulated by the first-time enactment of a 2nd generation anti-takeover law (FL) in year t; constant after CEO departure.
$FL_{i,t}$	CEO i's cumulative exposure to the first-time enactment of a 2nd generation anti-takeover law (FL) during tenure up to time t (in years); constant after CEO departure.
$Year_{i,t}$	Year of a subspell; used in hazard models when linearly controlling for time.
$Pay_{i,t}$	CEO i's total pay in year t (from Gibbons and Murphy (1992)).
$Assets_{j,t}$	Firm j's total assets in year t (from Compustat); missing data is interpolated.
$Employees_{j,t}$	Firm j's total number of employees in year t (from Compustat); missing data is interpolated.

Appendix B. Industry-Wide Distress Shocks and Apparent Aging: Apparent-Age Estimation Details and Robustness Tests

1. Apparent-Age Estimation

Our goal is to trace visible signs of aging in CEOs' faces. That is, we are interested in how old a person *looks*, which is referred to as the person's *apparent age*. By contrast, biological age describes how old a person is (time elapsed since birth) and will in general differ from a person's apparent age. To implement this analysis, we use machine learning based software by Antipov et al. (2016), henceforth referred to as the ABBD software. This software was specifically developed for the purpose of apparent-age estimation, and it was the winning solution of the second edition of the *ChaLearn Looking At People* competition in the *apparent-age estimation* track.

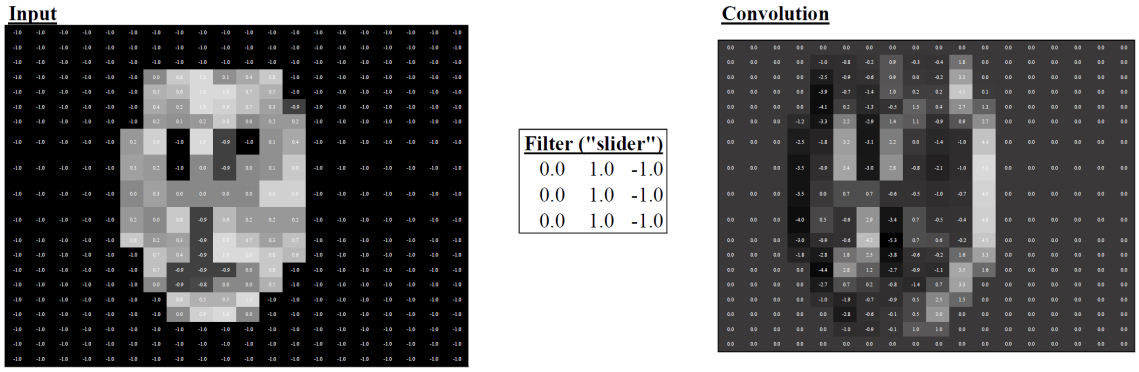
At the core of ABBD's apparent-age estimation tool is the training of a *convolutional neural network* (CNN). A CNN is a special class of *neural networks* that is particularly useful for image recognition and computer vision problems. A neural network is a system that learns to perform a task by studying training data.[27] It is architectured with three classes of layers: input, output, and hidden layers. The input layer receives the external data being evaluated, and the output data contains the network's response to the input. The in-between layers are the hidden layers, which abstractly determine intermediate features about the data. A CNN is a neural network in which some of the hidden layers employ the method of convolution, i. e., of transforming the input by sliding (or, convolving) over it, to detect patterns (such as edges or corners), which are then passed on to the next layer.

Appendix-Figure B.1 provides a simplified example of how convolution works in CNNs. Here, the fictional input is a shape that is roughly recognizable as a face (numbers between -1 and 1 determine pixel color). The filter matrix slides over the input and produces the output as the sum of element-wise matrix multiplication of $3x3$ pixel regions

[27] The task is referred to as *supervised learning* if the data is labeled (annotated), as is our training data.

with the filter matrix. As can be seen in the convoluted output, this specific filter matrix identifies right vertical edges. Convolutional layers further along in a system may be able to detect more advanced patterns such as, in our application, eyes or wrinkles.

CNNs have become widely popular over the past ten to twenty years, with numerous applications, in particular to image recognition and classification. In an influential article on *deep learning*[28] published in *Nature*, LeCun et al. (2015) summarize that CNNs have "brought about a revolution in computer vision" and "breakthroughs in processing images, video, speech and audio," and they are "now the dominant approach for almost all recognition and detection tasks."



Appendix-Figure B.1.—Simplified example of convolution.

Notes. This figure shows a simplified example of convolution. The fictional input image (left) is roughly recognizable as a face. Each cell (pixel) is encoded with a number that determines its color (between -1.0-black and $+1.0$-white). The output image (right) is obtained through convolution. The 3×3 filter matrix (center) slides over each possible 3×3 region in the input image and outputs the sum of element-wise matrix multiplication of these $3x3$ image regions and the filter matrix. Example inspired by material by Jeremy Howard (youtube.com/watch?v=V2h3IOBDvrA) and deeplizard (deeplizard.com/learn/video/YRhxdVk_sIs).

ABBD's apparent-age estimation software starts from a pre-trained version of a state-of-the-art CNN for face recognition called VGG-16,[29] and involves two key steps: *training* and *fine-tuning* of the CNN. In a first step, this CNN is trained on a large dataset of more

[28] A neural network is considered *deep* if it has multiple hidden layers.

[29] Introduced by Simonyan and Zisserman (2014), VGG-16 is a deep CNN. ABBD's software uses a VGG-16 version by Parkhi et al. (2015), which was trained for the purposes of face recognition (identifying identities from facial images) on 2.6 million images. Both works have been widely used and cited.

than 250,000 facial images from the IMDb (Internet Movie Database) and Wikipedia, which also contains information on the biological age of the person. The training step is implemented by minimizing the mean absolute error between predicted age and biological age. In a second step, the software is fine-tuned for apparent-age estimation on a unique dataset of 5,613 facial images, which also contains information on the *apparent* age of the person in each picture. The information on people's apparent age consists of at least 10 age estimates (per picture) by humans, which were specifically collected for the *ChaLearn Looking At People* competition. The fine-tuning step is implemented by minimizing a metric that penalizes deviations from the average (human) age estimate more when the disagreement about the person's apparent age is low.[30] Training and fine-tuning essentially mean that the software learns to estimate the age of the people in the two datasets using the information on biological and apparent age by adapting learning parameters in the hidden layers.

ABBD's software and apparent-age estimation tool have a variety of notable features:

Age distribution in training datasets. Both the IMDb-Wikipedia data and the dataset employed for human-based fine-tuning include people from all age groups, and in particular people aged 50 and above. This ensures that the software is trained and fine-tuned on data that includes people with similar facial characteristics as our CEOs, such as with regard to baldness patterns, hair color, and wrinkle development. For reference, the CEO at the 10^{th} (50^{th}, 90^{th}) percentile in our dataset is 47 (56, 63) years old in 2006 (see Table I).

Image pre-processing. Before feeding the pictures into the CNN for training and fine-tuning, ABBD "standardize" them, a process they label picture pre-processing. Specifically, they use existing software solutions to detect, scale, and align the face in each image,

[30] The metric is defined as $\epsilon = 1 - \exp\left(-\frac{(\hat{x}-\mu)^2}{2\sigma^2}\right)$, where $\hat{x}$ is the predicted apparent age, and μ and σ are the image-level mean and standard deviation of across the human-based age estimates.

and resize each image to 224×224 pixels. Intuitively, standardizing images reduces the noise present when training and fine-tuning the software and improves performance (cf. Table 2 in Antipov et al. (2016)). The software's performance on the *ChaLearn Looking At People* competition data improves by approximately 1% as a result of image-preprocessing (cf. Table 2 in Antipov et al. (2016)).

ABBD's trained software does not include image pre-processing code (and can, in fact, be applied to "raw images" so long as they are resized). We nonetheless replicate some of their pre-processing steps in order to increase the similarity between our CEO images and the images used for software training. We use the Python-based "face_recognition" package[31] to detect the picture region showing the CEO's face, extract the face, center it in the image, and resize the image to 224×224 pixels. Note that any remaining differences to ABBD's image pre-processing might increase the noise in our apparent-age estimates, but not introduce bias as any potential systematic differences in pre-processing steps would need to be correlated with industry shock exposure during the Great Recession. Before pre-processing a picture, we make sure that the image contains only the face of the CEO. If a picture contains multiple faces, such as a CEO with their partner, other managers, or a journalist, we first manually crop the picture and keep only the portion that shows the CEO.

Appendix-Figure B.2 shows several examples of pre-processed facial images. Panel (a) shows pre-processed images used to train ABBD's software. One can see that they differ in terms of "tint" and background. For example, the leftmost picture has a bluish tint and dark background, whereas the rightmost picture has a yellowish tint and light background. This underscores the spectrum of image characteristics the software is "exposed" to while being trained for apparent-age estimation. Panel (b) shows pre-processed CEO images from our sample. Again, there are differences in terms of tint and background, so it is worth reiterating that these are image features that the software can learn to take into account in

[31] The full package documentation is at github.com/ageitgey/face_recognition/blob/master/README.md.

its estimation during the training stage. Furthermore, comparing images across the two panels illustrates that our implementation of the image pre-processing step indeed leads to similar results compared to ABBD's original implementation on the training datasets.

(a) Training sample

(b) CEO sample

Appendix-Figure B.2.—Examples of pre-processed images.

Notes. This figure shows examples of pre-processed facial images. Panel (a) shows examples of images used in the training of the apparent-age estimation software. Panel (b) shows examples of pre-processed CEO images from our sample.

Accuracy gains from software fine-tuning. As described above, ABBD's software development includes a fine-tuning step using a dataset on human-based age estimates. Across all training and image pre-processing steps, the fine-tuning on this apparent age led to the biggest accuracy improvement on the competition data, amounting to more than 20% (cf. Table 2 in Antipov et al. (2016)). This underscores the importance of using a software specifically trained for apparent-age estimation, rather than an "off-the-shelf" software solely trained on images annotated with people's biological age, for our study of CEO visual aging.

Cross-validation. Rather than training one CNN on the 5,613 training images, ABBD's apparent-age estimation merges eleven CNNs, which were trained using eleven-fold cross-validation. Cross-validation is a popular technique in prediction problems. As part of

the training step, a portion of the data (the *validation sample*) is set aside for out-of-sample tests, i. e., tests on data the algorithm was not trained on. Moreover, instead of fixing the validation sample, it is common to train separate models using non-overlapping validation samples and to then average the results. In ABBD's implementation, each of the eleven "sub-CNNs" uses 5,113 images for training and 500 (non-overlapping) images for validation; this corresponds to a near-complete partition of the full training data into equal-sized validation samples ($5,613/11 \approx 500$). Each sub-model outputs a 100×1 vector of probabilities associated with all apparent ages between 0 and 99 years. ABBD's final solution, on which our analyses are based, uses the average of the probabilities across all sub-models.

Data augmentation. In the fine-tuning step of the software development, ABBD use five-times data augmentation to reduce overfitting. This is a popular technique to enlarge the training (or fine-tuning) sample, i. e., to allow the software to learn on more data. Specifically, each apparent-age annotated image is fed into the algorithm jointly with five modified versions: the mirrored image, a rotated image (± 5 degree), a horizontally shifted image ($\pm 5\%$), and a scaled image ($\pm 5\%$). To see the potential benefit of data augmentation in our application, suppose that among the fine-tuning sample of 5,613 images, people who look older happen to look slightly to the upper right. Including mirrored and rotated images in the fine-tuning step reduces the likelihood that the software may learn to associate apparent age with camera angle.[32] In our application, data augmentation also further alleviates concerns about effects of slight differences in image pre-processing.

To match the steps during training, ABBD's final solution uses the same image modifications also on new images that are fed into the tool, i. e., it estimates different apparent ages for each image in our CEO sample based on the original image and modified images as outlined above. The final apparent age is the average across the different estimates.

[32]These specific image modifications assume that there is *no* intrinsic relation between apparent age and camera angle. This appears reasonable but highlights that data augmentation choices involve judgment.

2. Additional Figures and Tables

This section contains all robustness figures and tables on industry-wide distress shocks and apparent aging.

Appendix-Figure B.3.—Average number of pictures per CEO across years.

Notes: This figure depicts the average number of pictures per CEO we are able to collect each year for the group of CEOs that experienced industry shocks during 2007-2008 and the group that did not. The two black vertical lines indicate the years 2006 and 2008.

APPENDIX-TABLE B.1
INDUSTRY DISTRESS AND CEO AGING – NO WINSORIZATION

Dependent Variable: $Apparent\ Age_{i,j,t}$				
	(1)	(2)	(3)	(4)
Industry Distress $\times\ 1_{\{t>2006\}}$	0.940*	0.978**		
	[0.496]	[0.491]		
Industry Distress $\times\ 1_{\{2006<t<2012\}}$			0.807	0.841
			[0.536]	[0.531]
Industry Distress $\times\ 1_{\{t\geq 2012\}}$			1.135*	1.178**
			[0.576]	[0.564]
Biological Age	0.912***	0.908***	0.944***	0.940***
	[0.093]	[0.093]	[0.095]	[0.095]
CEO FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Picture Controls		Y		Y
Number of CEOs	463	463	463	463
Observations	3,086	3,086	3,086	3,086

APPENDIX-TABLE B.2
INDUSTRY DISTRESS AND CEO AGING – RESTRICTIVE INDUSTRY DISTRESS DEFINITION

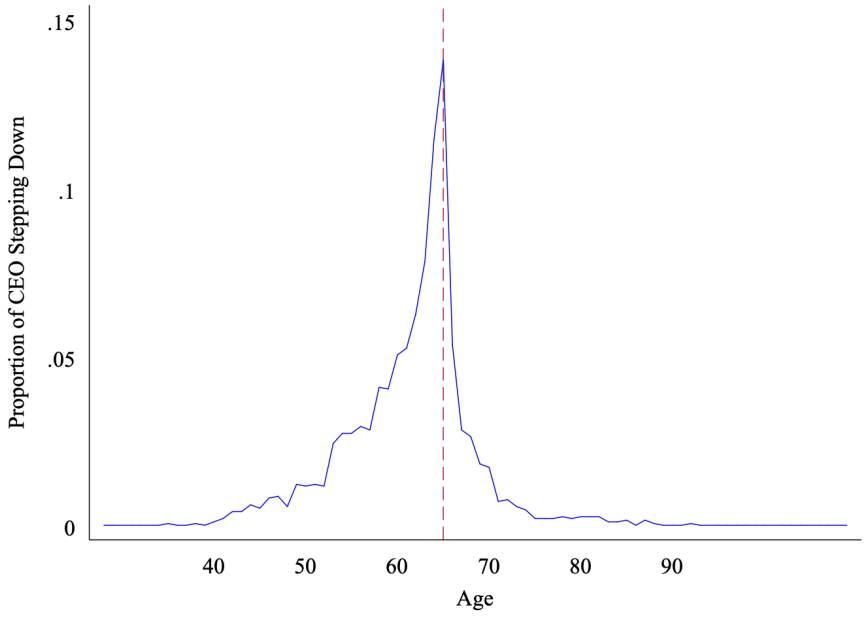
Dependent Variable: $Apparent\ Age_{i,j,t}$				
	(1)	(2)	(3)	(4)
Industry Distress $\times\ 1_{\{t>2006\}}$	1.173**	1.064**		
	[0.469]	[0.463]		
Industry Distress $\times\ 1_{\{2006<t<2012\}}$			1.102**	1.007*
			[0.504]	[0.510]
Industry Distress $\times\ 1_{\{t\geq 2012\}}$			1.261**	1.135**
			[0.563]	[0.551]
Biological Age	1.272***	1.274***	1.268***	1.269***
	[0.021]	[0.022]	[0.027]	[0.028]
CEO FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Picture Controls		Y		Y
Number of CEOs	463	463	463	463
Observations	3,086	3,086	3,086	3,086

APPENDIX-TABLE B.3
INDUSTRY DISTRESS AND CEO AGING – PRE-2016 SAMPLE

Dependent Variable: $Apparent\ Age_{i,j,t}$				
	(1)	(2)	(3)	(4)
Industry Distress $\times\ 1_{\{t>2006\}}$	0.816*	0.839*		
	[0.487]	[0.485]		
Industry Distress $\times\ 1_{\{2006<t<2012\}}$			0.604	0.622
			[0.535]	[0.525]
Industry Distress $\times\ 1_{\{t\geq 2012\}}$			1.323**	1.360**
			[0.566]	[0.568]
Biological Age	0.952***	0.943***	0.983***	0.974***
	[0.095]	[0.096]	[0.094]	[0.093]
CEO FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
Picture Controls		Y		Y
Number of CEOs	463	463	463	463
Observations	3,086	3,086	3,086	3,086

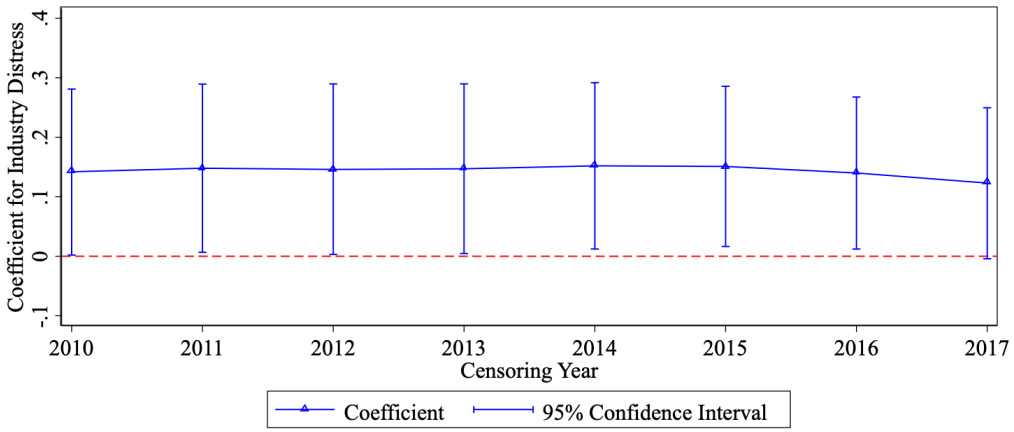
Appendix C. Industry-Wide Distress Shocks and Life Expectancy: Robustness Tests

This appendix presents the robustness tests of the relation between industry distress and CEOs' life expectancy referenced in Section 2.4.



Appendix-Figure C.1. —Proportion of CEOs stepping down by age.

Notes: This figure depicts the proportion of CEOs stepping down at each age. The vertical dashed line indicates age 65.



Appendix-Figure C.2.—Estimated effect of industry distress when varying the censoring year.

Notes: This figure shows the estimated coefficients on the industry distress indicator variable using the specification from Table IV, column 1, but varying the censoring date. In the main analysis, the cutoff date is Oct. 1, 2017, i. e., CEOs who did not pass away before this date are treated as censored. The alternative censoring dates are Dec. 31, 2016; Dec. 31, 2015; ...; up to Dec. 31, 2010. The number of CEOs in the sample remains unchanged when varying the cutoff, i.e. $N = 1,605$.

APPENDIX-TABLE C.1
INDUSTRY DISTRESS AND MORTALITY – ADDITIONAL CONTROLS

Dependent Variable: $Death_{i,t}$				
	(1)	(2)	(3)	(4)
Industry Distress	0.117*	0.150**	0.178***	0.169**
	[0.066]	[0.062]	[0.069]	[0.070]
Age	0.111***	0.113***	0.121***	0.121***
	[0.005]	[0.005]	[0.006]	[0.006]
Year		0.001	0.002	
		[0.006]	[0.006]	
ln(Pay)	-0.066*	-0.059	-0.019	-0.022
	[0.037]	[0.037]	[0.048]	[0.049]
ln(Assets)	0.012	0.006	-0.080*	-0.086*
	[0.031]	[0.034]	[0.049]	[0.049]
ln(Employees)	0.003	0.007	0.070	0.072
	[0.038]	[0.038]	[0.053]	[0.053]
Location FE (HQ)		Y	Y	Y
FF49 FE			Y	Y
Year FE				Y
Number of CEOs	1,553	1,553	1,553	1,553
Observations	49,052	49,052	49,052	49,052

NOTE. — This table reports hazard coefficients estimated as in Table IV but with additional controls for CEO pay, assets, and employees. The dependent variable is an indicator that equals one if the CEO dies in a given year. The main independent variable $Industry\ Distress$ is an indicator of a CEO's exposure to industry distress shocks. All variables are defined in Appendix 2.6. Standard errors, clustered at the industry level, are shown in brackets. *, **, and *** denote significance at the 10, 5, and 1 percent level, respectively.

APPENDIX-TABLE C.2
INDUSTRY DISTRESS AND MORTALITY – ADDITIONAL CEOS

Dependent Variable: $Death_{i,t}$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Baseline Controls				Panel B: Additional Controls			
Industry Distress	0.075	0.095*	0.110**	0.103*	0.072	0.097*	0.141**	0.133**
	[0.062]	[0.056]	[0.054]	[0.055]	[0.063]	[0.057]	[0.060]	[0.060]
Age	0.113***	0.116***	0.123***	0.123***	0.112***	0.115***	0.122***	0.122***
	[0.006]	[0.006]	[0.007]	[0.007]	[0.005]	[0.005]	[0.006]	[0.006]
Year		-0.001	-0.003			-0.001	0.001	
		[0.005]	[0.005]			[0.005]	[0.006]	
ln(Pay)					-0.050	-0.043	0.002	-0.001
					[0.036]	[0.037]	[0.044]	[0.045]
ln(Assets)					0.020	0.015	-0.084*	-0.088*
					[0.030]	[0.032]	[0.049]	[0.049]
ln(Employees)					-0.020	-0.022	0.048	0.049
					[0.036]	[0.035]	[0.051]	[0.051]
Location FE (HQ)		Y	Y	Y		Y	Y	Y
FF49 FE			Y	Y			Y	Y
Year FE				Y				Y
Number of CEOs	1,900	1,900	1,900	1,900	1,818	1,818	1,818	1,818
Observations	58,034	58,034	58,034	58,034	55,796	55,796	55,796	55,796

NOTE. — This table reports hazard coefficients estimated as in Table IV but using an extended CEO sample as described in Section II.B. The dependent variable is an indicator that equals one if the CEO dies in a given year. The main independent variable $Industry\ Distress$ is an indicator of a CEO's exposure to industry distress shocks. Panel A includes the baseline controls from Table IV and Panel B adds additional additional controls for CEO pay, assets, and employees as in Appendix-Table C.1. All variables are defined in Appendix 2.6. Standard errors, clustered at the industry level, are shown in brackets. *, **, and *** denote significance at the 10, 5, and 1 percent level, respectively.

Appendix D. Corporate Monitoring and Life Expectancy: Robustness Tests

This appendix presents the robustness tests of the relation between anti-takeover laws and CEOs' life expectancy referenced in Section 2.5. We first present additional details on BC-specific robustness tests and then report additional figures and tables.

Other Anti-Takeover Laws: First-Time Passage of Second-Generation Anti-Takeover Laws

Our main analysis exploits the enactment of BC laws as they have been shown to create substantial conflicts of interest between managers and shareholders (Bertrand and Mullainathan (2003), Gormley and Matsa (2016)). Some researchers have questioned whether BC laws were the most important legal development impacting corporate governance at the time (see the discussion in Cain et al. (2017) and Karpoff and Wittry (2018)). Here, we replicate our analyses for other anti-takeover legislation from the 1980s that induced plausibly exogenous variation in corporate monitoring intensity.

In addition to BC laws, four other types of anti-takeover laws were passed by individual states since the 1980s: (1) Control Share Acquisition laws prohibited acquirers of large equity stakes from using their voting rights, making it more difficult for hostile acquirers to gain control. (2) Fair Price laws required acquirers to pay a fair price for shares acquired in a takeover attempt. Fair could mean, for example, the highest price paid by the acquirer for shares of the target within the last 24 months (cf. Cheng et al. (2004)). (3) Directors' Duties laws extended the board members' duties to incorporate the interests of non-investor stakeholders, even if not necessarily maximizing shareholder value. (4) Poison Pill laws guaranteed that the firms had the right to use poison pill takeover defenses. We refer to the first of these five laws (including BC laws) passed by a state as the *First Law* (*FL*). Anti-takeover law exposure is similar when jointly looking at all five second-generation laws. For example, conditional on any *FL* exposure, the median CEO

experiences 4.45 years, close to the 4.41 years in the BC-based analysis.

Appendix-Figure D.1 visualizes the *FL* enactment by states over time.

Appendix-Table D.2 re-estimates Table VI using *FL* enactment as identifying variation. We limit the sample to the 1,510 CEOs who are appointed in years prior to the *FL* enactment of any of the five second-generation anti-takeover laws. Consistent with our main findings, we estimate a significant increase in longevity for CEOs under less stringent governance regimes. The estimated effect sizes are very similar to our main specification using BC laws. For example, for the specifications in Panel A based on cumulative law exposure, the hazard coefficients range from -0.039 to -0.046, compared to -0.041 to -0.046 in Table VI. As Panel B shows, the *FL* results are also robust to including the additional CEO and firm level controls from Panel A of Appendix-Table D.1.

Karpoff-Wittry and Related Tests: Institutional and Legal Context of the Anti-Takeover Laws

Karpoff and Wittry (2018) propose several robustness tests to address endogenous firm responses to anti-takeover laws, which we implement in Appendix-Table D.3. For all sample restrictions, we follow the procedure suggested in Karpoff and Wittry (2018): In Panel A, we remove the 46 firms identified by these authors as having lobbied for the passage of the second-generation laws. In Panel B, we use Institutional Shareholder Services (ISS) Governance (formerly, RiskMetrics) data from 1990 to 2017 to identify firms that opted out of coverage by the laws and exclude them from the analysis. In Panel C, we exclude firm-years in which firms had adopted firm-level anti-takeover defenses. We identify firms with firm-level defenses combining ISS data with data provided to us by Cremers and Ferrell (2014), which extends the Gompers et al. (2003) G-index backwards to 1977-1989. We back out whether firms used firm-level defenses in 1977-1989 by "subtracting" the state-wide laws from the G-index, which combines firm- and state-level defenses. Firm-level defenses include Golden Parachutes and Cumulative Voting (cf. Gompers et al. (2003) for

details).

In all subsamples, the hazard coefficient on BC exposure remains significant at 1%, both when using the indicator and the count variable for BC experience. In addition, the hazard coefficient estimates are nearly unchanged, ranging from -0.218 (Panel B, column 1) to -0.276 (Panel A, column 3) for the indicator version, and from -0.041 (Panel A, column 5) to -0.047 (Panel C, column 2) for the count version.

Karpoff and Wittry (2018) also point to possible confounding effects of first-generation anti-takeover laws. They raise the concern that firms without BC exposure might experience lenient governance before 1982 because first-generation anti-takeover laws effectively lost their effect only starting from June 1982 after the *Edgar v. MITE* ruling. We address this concern in Appendix-Table D.4 through three cuts of the data. In subsample A, we drop all CEO-years prior to 1982, i. e., we restrict the sample to years from 1982 onward (albeit including the post-1982 years for CEOs who stepped down prior to 1982). In subsample B, we drop all CEOs who stepped down prior to 1982, i. e., we restrict the sample to CEOs who served during the "post-first-law period" (including CEO-years prior to 1982). Note that in terms of number of CEOs remaining, subsample B is more restrictive than subsample A. In subsample C, we restrict the sample to CEOs who began their tenure in or after 1982, i. e., subsample C is a subset of subsample B. In all subsamples, we continue to estimate negative hazard coefficients for both the indicator and cumulative BC exposure variables, similar in size to those in the main table. The coefficients remain significant at 1% in subsamples A and B as well as in the most restrictive subsample C when using the indicator BC variable. In the latter sample, we lose statistical power when using the cumulative BC exposure (standard errors quintuple), though the point estimate remains similar.

Finally, in a last set of robustness checks, we move beyond the tests suggested in Karpoff and Wittry (2018) and create sub-samples based on firms' state of incorporation and

industry affiliation, inspired by similar robustness checks in Giroud and Mueller (2010) and Gormley and Matsa (2016). In Appendix-Table D.5, we exclude firms that are incorporated in Delaware or in New York, the two most common states of incorporation in our sample (Panel A); firms in the Banking industry (Panel B); or firms in the Utilities industry (Panel C). In all three panels, the hazard coefficient estimates on binary and cumulative BC exposure are barely affected by these data cuts.

Predicted Length of Exposure: Prediction Model Details

To purge the per-year estimates in the right four columns of Table VI of possible endogeneity in the length of exposure, we construct a measure of predicted BC law exposure, and relate predicted, rather than true exposure, to CEO mortality rates.[33] To this end, we proceed in three steps. First, we estimate a prediction model for CEO tenure; we then construct predicted BC exposure; and finally we re-estimate the hazard regressions using predicted BC exposure as the independent variable.

We first predict for every CEO-year, including years after the passage of a BC law:

$$RemainTenure_{i,t} = X'_{i,t}A + e_{i,t}. \tag{D.1}$$

The control variables are an age cubic, tenure cubic, the CEO's cumulative exposure to the BC law until year t, $BC_{i,t}$, and fixed effects for industry, year, headquarters state, birth year, and tenure start-year. Denoting as t^* the year when the BC law is passed, we use the predicted remaining tenure at t^* from equation (D.1) to construct CEOs' predicted exposure to BC laws,

$$\widehat{BC}^*_i = I(BCLawPassed_{s(i),t}) \times \widehat{RemainTenure}_{i,t^*}, \tag{D.2}$$

where $I(BCLawPassed_{s(i),t}) = 1$ for CEO i in state $s(i)$ at $t \geq t^*$. $\widehat{RemainTenure}_{i,t^*}$ is backward-looking, i. e., constructed using information from years up to t^*.

[33] Directly controlling for realized tenure would result in a "bad control" problem and introduce bias (Angrist and Pischke (2008)). .

Using this variable, we construct a CEO's predicted cumulative BC exposure until year t, $\widehat{BC}_{i,t}$ as (i) $\widehat{BC}_{i,t} = 0 \; \forall t$ in the control group; (ii) $\widehat{BC}_{i,t} = 0 \; \forall t < t^*$ if not yet treated; and (iii) $\widehat{BC}_{i,t} = \min\{k+1, \widehat{BC}_i^*\}$ for each year t following t^*, with $t = t^* + k$. Note that k is allowed to be fractional if the BC law goes into effect in the middle of the year.

We then use the predicted cumulative exposure in the following hazard estimations:

$$\lambda(t|\widehat{BC}_{i,t}, X_{i,t}) = \lambda_0(t) \exp\{\beta \, \widehat{BC}_{i,t} + \delta' X_{i,t}\} \tag{D.3}$$

Since this approach involves a generated regressor, we report bootstrapped standard errors, using the block bootstrap method (a block is a state of incorporation cluster) with 500 iterations.

Business Combination Laws and CEO Pay

As discussed in the main text, our aging and mortality results also prompt the question whether parties account for the health consequences of (permanent) changes in job demands. We explore this in the context of the permanent corporate governance regime change induced by BC laws.

The theoretical prediction on the link between BC laws and CEO pay is in fact unclear, as also noted by Bertrand and Mullainathan (1998). On the one hand, a model of compensating differentials would predict a decrease in pay as CEOs' working conditions improve and imposed health costs are reduced. In line with such a channel, Edmans and Gabaix (2011) present a theoretical model of the CEO market in which lower effort—which is isomorphic to lower job demands—is compensated by lower pay. On the other hand, a model of skimming would predict that CEOs use the increase in autonomy to extract additional private benefits in the form of higher compensation. It is thus an empirical question as to which effect dominates in our specific context.

Before estimating the empirical relation, it is useful to first calibrate what effect size we would expect if compensation for health ramifications were the primary channel em-

pirically. In their meta-analysis of the literature on the value of a statistical life (VSL), Viscusi and Aldy (2003) report an estimate around \$6.7 million (in 2000 dollars) for a person with income of around \$26,000, and an income elasticity for the VSL of around 0.5. Applied to our CEO sample, this translates into a VSL of around \$47.3 million.[34] Given a baseline mortality rate of 1.366% for 60-year-olds born in 1925 (Human Mortality Database (2019)), a reduction in mortality risk of 4.1% per year of BC exposure (column (5) in Table VI) implies a CEO pay change between -2% and -9%, depending on whether the wage adjustment reflects the entire BC-induced mortality risk shift over the expected remaining lifespan or solely the shift over the remaining years while serving as CEO.[35]

With these calibrated effects in mind, Table D.7 presents the results on the relation between CEO pay and BC law exposure. In column (1), we estimate linear regressions of CEO pay on the BC indicator and the same controls and fixed effects as in the hazard analyses. This specification excludes any post-treatment outcomes from the right-hand side and parallels the survival analysis. In column (2), we add the control variables used in Bertrand and Mullainathan (1998): tenure, firm assets, and employees. We note that these controls may themselves be affected by the reform and therefore absorb the effect of the anti-takeover laws. Finally, in column (3), we add firm fixed effects (in place of industry fixed effects), as in the baseline specification of Bertrand and Mullainathan (1998):

$$\ln(Pay_{i,j,t}) = \alpha_t + \beta_j + \gamma\, I(BC_{i,t}) + \delta' X_{i,j,t} + e_{i,j,t}$$

where i represents a CEO, j represents a firm, and t represents a calendar year.

We estimate a positive, albeit mostly insignificant treatment effect. The estimates in-

[34] Given an average CEO pay of \$1.3 million (in 2000 dollars) in our sample, we can calculate the implied VSL for the average CEO as $VSL_{CEO} = \exp(0.5 \times (\ln(\$1.3\text{m}) - \ln(\$26\text{k})) + \ln(\$6.7\text{m})) = \$47.3\text{m}$.

[35] The calculations are based on an average length of BC exposure of 5.68 years (Table II), an average time of 24.77 years between onset of BC exposure and death, and an average annual CEO pay of \$1.3 million in 2000 dollars). For example, if we assume that the wage adjustment reflects the mortality risk shift over the expected remaining lifespan, we can calculate the pay change as $(-24.77/5.68) \times (4.1\% \times 1.366\% \times \$47.3\text{mn})/\$1.3\text{mn} = -9\%$.

dicate a pay increase between 4.1–8.7%. Only the estimate in column (2) is marginally significant. In comparing the results to the earlier work, which estimated a (more significant) 5.4 percent pay increase, it is important to note that our analysis is conducted on a CEO-level sample, and restricts the sample to incumbent, pre-BC CEOs.

The evidence speaks against a compensating reduction in pay, but is instead suggestive of additional rents (higher pay). However, any resulting wealth increases are unlikely to explain the longevity results, given that the literature has found little evidence of a causal relation of income and life expectancy for wealthy individuals (Cesarini et al. (2016)). Where evidence has been found of an effect of wealth on health, it appears to work through reductions in stress (Schwandt (2018)). The apparent lack of a compensating differential casts doubt on whether all parties fully account for the health implications of different governance regimes.

Business Combination Laws and CEO Tenure

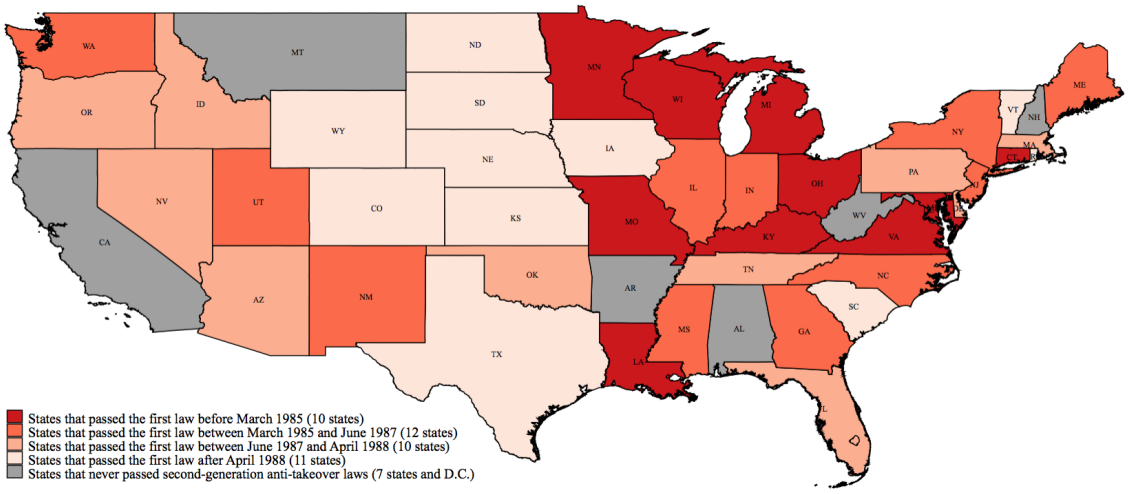
In addition to CEO pay, we can also explore how CEO tenure responds to the introduction of anti-takeover laws, which can also provide insight into how managers perceive this permanent change in corporate monitoring to affect job demands.

Similar to the case of CEO pay, theory does not provide a clear prediction as to how tenure should respond to the anti-takeover laws. On the one hand, CEOs may become entrenched and stay on the job longer. On the other hand, CEOs who reduce effort on the job might be fired more frequently. We estimate again hazard models, now analyzing CEO departure as the outcome variable. The results in columns (1) of Appendix-Table D.8 indicate that BC law treatment, $I(BC)$, decreases the separation hazard by around 22 percent, but the effect halves in magnitude and becomes statistically insignificant after controlling for year effects (column 2), with standard errors nearly doubling. In the specifications using the length of exposure variable BC (in columns 3 to 4), the estimated separation hazard falls by 4 to 10 percent. These results are suggestive of moderate increases in

tenure in response to BC law passage.

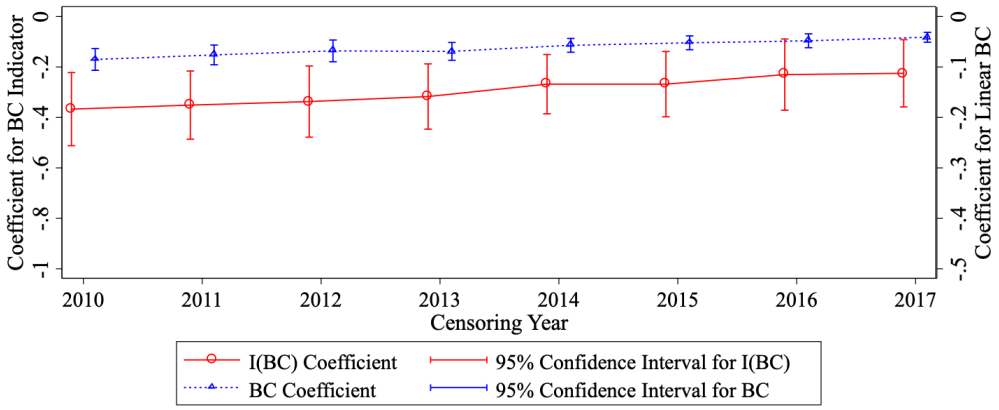
Further analyses of CEOs' age at the end of their tenure suggest that any increases in tenure would be driven by fewer CEOs stepping down in their 50s and early 60s. Appendix-Figure D.4 plots the retirement hazard separately for CEOs with and without BC law exposure. Exposure appears to lower the hazard before and increase it above age 65, including a long tail of tenures into the 80s and 90s. While the raw data is not as stark as for our longevity results, nor are the hazard estimates as robust, it is noteworthy for another reason: It helps rule out that the end of mandatory retirement through the amendment of the Age Discrimination in Employment Act (ADEA) in 1986 confounds our BC-law–longevity findings. Although there is a large spike in retirements at ages 64 and 65, there is no association between retirement at these ages and exposure to the BC laws.

An increase in tenure (or delayed retirement) as a result of anti-takeover insulation is also unlikely to be the channel for the estimated increase in longevity. To begin with, prior research has found small or even beneficial effects of retirement on health in the general population (Hernaes et al. (2013), Insler (2014), Fitzpatrick and Moore (2018)). In our population, a life expectancy advantage arising directly from tenure would run counter to the notion that the CEO job is demanding as the evidence in Bandiera et al. (2020) and Porter and Nohria (2018) on the intensity of CEO schedules and the constraints imposed by the CEO position imply. Moreover, the results in Section V.E on nonlinearities point to initial exposure effects, with prolonged exposure (from prolonged tenure) having no incremental impact on life expectancy.



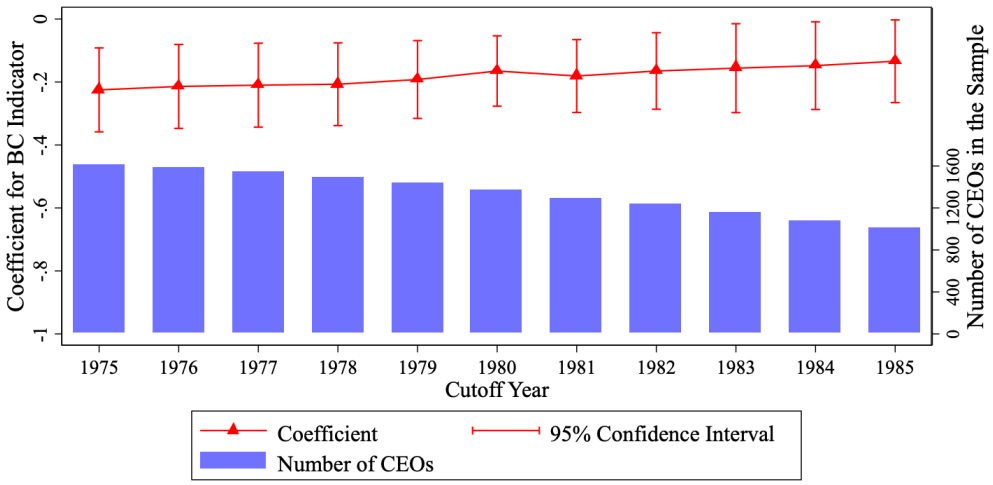
Appendix-Figure D.1.—First-time introduction of second-generation anti-takeover laws over time.

Notes: This figure visualizes the distribution of first-time enactments of any of the five most common second-generation anti-takeover laws over time, i. e., business combination (BC), fair price (FP), control share acquisition (CSA), poison pills (PP), and directors' duties (DD) laws. The graph omits the states of Alaska and Hawaii. Alaska did not adopt any second-generation anti-takeover laws. Hawaii adopted a CSA law on 4/23/1985, and DD and PP laws on 6/7/1988.



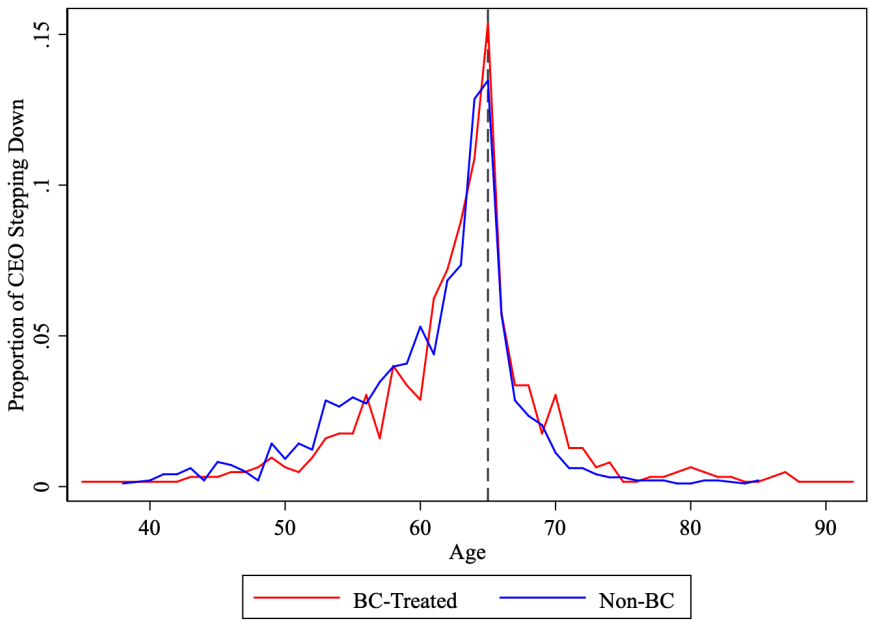
Appendix-Figure D.2.—Estimated effect of the BC law exposure when varying the censoring year.

Notes: This figure shows the estimated coefficients on the BC indicator variable $I(BC)$ and the cumulative BC variable BC when using the specifications from Table VI, columns 1 and 5, but varying the censoring date. In the main analysis, the cutoff date is Oct. 1, 2017, i. e., CEOs who did not pass away before this date are treated as censored. The alternative censoring dates are Dec. 31, 2016; Dec. 31, 2015; ...; up to Dec. 31, 2010. The number of CEOs in the sample remains unchanged when varying the cutoff, i.e. $N = 1,605$.



Appendix-Figure D.3.—Estimated effect of the BC law exposure when varying the sample cutoff year.

Notes: This figure shows the estimated coefficients on the BC indicator variable $I(BC)$ when using the specification from Table VI, column 1, but varying the sample. In the main sample, CEOs end their tenure in or later than 1975. We vary this cutoff year from 1975 to 1985, when the first BC law ever was passed. The blue (dark) bars are the number of CEOs in the sample. When the cutoff year is 1975 (our main sample), the number is 1,605 and the estimated coefficient is the same as shown in Table VI, column 4.



Appendix-Figure D.4. —Proportion of CEOs stepping down by age.

Notes: This figure depicts the proportion of CEOs stepping down at each age, split by whether or not a CEO was exposed to a business combination (BC) law. The vertical dashed line indicates age 65.

APPENDIX-TABLE D.1

BUSINESS COMBINATION LAWS AND MORTALITY –

ADDITIONAL CONTROLS AND STATE-OF-INCORPORATION FIXED EFFECTS

Dependent Variable: $Death_{i,t}$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Additional Controls				
I(BC)	-0.224***	-0.261***	-0.241***	-0.225***				
	[0.064]	[0.073]	[0.086]	[0.082]				
BC					-0.042***	-0.045***	-0.040***	-0.039***
					[0.007]	[0.007]	[0.007]	[0.007]
ln(Pay)	-0.019	-0.014	0.008	0.003	-0.023	-0.023	-0.012	-0.015
	[0.033]	[0.036]	[0.039]	[0.040]	[0.043]	[0.043]	[0.048]	[0.049]
ln(Assets)	0.020	0.015	-0.033	-0.041	0.028	0.023	-0.014	-0.021
	[0.019]	[0.025]	[0.043]	[0.041]	[0.018]	[0.023]	[0.036]	[0.035]
ln(Employees)	-0.016	-0.010	0.017	0.022	-0.018	-0.012	0.008	0.012
	[0.020]	[0.021]	[0.038]	[0.037]	[0.019]	[0.021]	[0.036]	[0.037]
Location FE (HQ)		Y	Y	Y		Y	Y	Y
Number of CEOs	1,503	1,503	1,503	1,503	1,503	1,503	1,503	1,503
Observations	49,052	49,052	49,052	49,052	49,502	49,052	49,052	49,052
				Panel B: State-of-Incorporation Fixed Effects				
I(BC)	-0.225***	-0.265***	-0.274***	-0.264***				
	[0.068]	[0.083]	[0.088]	[0.086]				
BC					-0.041***	-0.048***	-0.046***	-0.045***
					[0.006]	[0.007]	[0.007]	[0.007]
Location FE (Incorp.)		Y	Y	Y		Y	Y	Y
Number of CEOs	1,605	1,605	1,605	1,605	1,605	1,605	1,605	1,605
Observations	50,530	50,530	50,530	50,530	50,530	50,530	50,530	50,530
Age (Linear Control)	Y	Y	Y	Y	Y	Y	Y	Y
Year (Linear Control)		Y	Y			Y	Y	
FF49 FE			Y	Y			Y	Y
Year FE				Y				Y

NOTE. — Panel A reports hazard coefficients estimated as in Table VI, with additional controls for CEO pay, assets, and employees. Panel B reports hazard coefficients estimated as in Table VI, but including state-of-incorporation fixed effects instead of state-of-headquarters fixed effects. Controls and fixed effects (in addition to location fixed effects based on state-of-headquarters or state-of-incorporation) for both panels are indicated at the bottom of the table. All variables are defined in Appendix 2.6. Standard errors, clustered at the state-of-incorporation level, are shown in brackets. *, **, and *** denote significance at the 10, 5, and 1 percent level, respectively.

APPENDIX-TABLE D.2
FIRST-TIME SECOND-GENERATION ANTI-TAKEOVER LAWS AND MORTALITY

Dependent Variable: $Death_{i,t}$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Baseline Results				
I(FL)	-0.187***	-0.220***	-0.221***	-0.215***				
	[0.064]	[0.066]	[0.076]	[0.075]				
FL					-0.039***	-0.046***	-0.044***	-0.044***
					[0.006]	[0.006]	[0.006]	[0.006]
Number of CEOs	1,510	1,510	1,510	1,510	1,510	1,510	1,510	1,510
Observations	47,994	47,994	47,994	47,994	47,994	47,994	47,994	47,994
				Panel B: Additional Controls				
I(FL)	-0.173***	-0.190***	-0.170**	-0.157**				
	[0.060]	[0.062]	[0.070]	[0.068]				
FL					-0.040***	-0.044***	-0.040***	-0.039***
					[0.008]	[0.009]	[0.008]	[0.008]
ln(Pay)	-0.030	-0.023	0.005	-0.000	-0.020	-0.016	0.001	-0.002
	[0.032]	[0.037]	[0.037]	[0.037]	[0.041]	[0.043]	[0.045]	[0.045]
ln(Assets)	0.017	0.014	-0.058	-0.066*	0.027	1.026	-0.024	-0.031
	[0.019]	[0.026]	[0.037]	[0.036]	[0.018]	[0.024]	[0.033]	[0.032]
ln(Employees)	-0.012	-0.005	0.044	0.049	-0.020	-0.013	0.019	0.022
	[0.019]	[0.020]	[0.034]	[0.035]	[0.019]	[0.020]	[0.035]	[0.037]
Number of CEOs	1,464	1,464	1,464	1,464	1,464	1,464	1,464	1,464
Observations	46,660	46,660	46,660	46,660	46,660	46,660	46,660	46,660
Age (Linear Control)	Y	Y	Y	Y	Y	Y	Y	Y
Year (Linear Control)		Y	Y			Y	Y	
Location FE (HQ)		Y	Y	Y		Y	Y	Y
FF49 FE			Y	Y			Y	Y
Year FE				Y				Y

NOTE. — This table reports hazard coefficients estimated as in Table VI, but using the first-time introduction of any of the five most common second-generation anti-takeover laws as measure of lenient governance. The sample is restricted to CEOs appointed prior to the introduction of the anti-takeover law(s). Panel B adds additional controls for CEO pay, assets, and employees. Controls and fixed effects for both panels are indicated at the bottom of the table. All variables are defined in Appendix 2.6. Standard errors, clustered at the state-of-incorporation level, are shown in brackets. *, **, and *** denote significance at the 10, 5, and 1 percent level, respectively.

APPENDIX-TABLE D.3
EXCLUDING LOBBYING FIRMS, OPT-OUT FIRMS, AND FIRM-YEARS WITH FIRM-LEVEL DEFENSES

Dependent Variable: $Death_{i,t}$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Excluding Lobbying Firms				
I(BC)	-0.232***	-0.270***	-0.276***	-0.268***				
	[0.071]	[0.068]	[0.074]	[0.074]				
BC					-0.041***	-0.045***	-0.043***	-0.043***
					[0.006]	[0.007]	[0.008]	[0.008]
Number of CEOs	1,530	1,530	1,530	1,530	1,530	1,530	1,530	1,530
Observations	48,106	48,106	48,106	48,106	48,106	48,106	48,106	48,106
				Panel B: Excluding Opt-out Firms				
I(BC)	-0.218***	-0.244***	-0.226***	-0.216***				
	[0.065]	[0.081]	[0.081]	[0.080]				
BC					-0.042***	-0.045***	-0.041***	-0.040***
					[0.006]	[0.006]	[0.006]	[0.006]
Number of CEOs	1,532	1,532	1,532	1,532	1,532	1,532	1,532	1,532
Observations	48,180	48,180	48,180	48,180	48,180	48,180	48,180	48,180
				Panel C: Excluding Firm-level Defenses				
I(BC)	-0.232***	-0.271***	-0.269***	-0.256***				
	[0.068]	[0.078]	[0.087]	[0.087]				
BC					-0.043***	-0.047***	-0.044***	-0.044***
					[0.006]	[0.005]	[0.005]	[0.005]
Number of CEOs	1,599	1,599	1,599	1,599	1,599	1,599	1,599	1,599
Observations	43,417	43,417	43,417	43,417	43,417	43,417	43,417	43,417
Age (Linear Control)	Y	Y	Y	Y	Y	Y	Y	Y
Location FE (HQ)		Y	Y	Y		Y	Y	Y
Year (Linear Control)		Y	Y			Y	Y	
FF49 FE			Y	Y			Y	Y
Year FE				Y				Y

NOTE. — This table reports hazard coefficients estimated as in Table VI, but with additional sample restrictions. In Panel A, we exclude 46 firms that Karpoff and Wittry (2018) identify as firms that lobbied for the enactment of the second-generation anti-takeover laws. In Panel B, we exclude 61 firms that opted out of the second-generation anti-takeover laws, based on data from the Institutional Shareholder Services (ISS) Governance database. In Panel C, we exclude firm-years in which firms used firm-level defenses as identified from the the ISS data and data from Cremers and Ferrell (2014). Controls and fixed effects for all three panels are indicated at the bottom of the table. All variables are defined in Appendix 2.6. Standard errors, clustered at the state-of-incorporation level, are shown in brackets. *, **, and *** denote significance at the 10, 5, and 1 percent level, respectively.

APPENDIX-TABLE D.4

RESTRICTION TO YEARS AFTER THE END OF THE FIRST-GENERATION LAWS

Dependent Variable: $Death_{i,t}$						
	(1)	(2)	(3)	(4)	(5)	(6)
	Subsample A: Drop CEO-years pre-1982		Subsample B: Drop CEOs stepping down pre-1982		Subsample C: CEOs starting in or after 1982	
I(BC)	-0.267***		-0.227***		-0.417***	
	[0.083]		[0.062]		[0.089]	
BC		-0.044***		-0.042***		-0.036
		[0.006]		[0.008]		[0.028]
Age	0.117***	0.115***	0.116***	0.112***	0.124***	0.118***
	[0.005]	[0.005]	[0.005]	[0.006]	[0.013]	[0.018]
Location FE (HQ)	Y	Y	Y	Y	Y	Y
FF49 FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Number of CEOs	1,573	1,573	1,231	1,231	477	477
Observations	40,834	40,834	39,623	39,623	13,562	13,562

NOTE. – This table re-estimates columns (3) and (6) of Table VI with the sample restricted to the period when the first-generation anti-takeover laws lost their effect (in June 1982 after the *Edgar v. MITE* ruling). In subsample A, we drop all CEO-years prior to 1982, i. e., we restrict the sample to years from 1982 onward (albeit including the post-1982 years for CEOs who stepped down prior to 1982). In subsample B, we drop all CEOs who stepped down prior to 1982, i. e., we restrict the sample to CEOs who served during the "post-first-law period" (including CEO-years prior to 1982). Note that in terms of number of CEOs remaining, subsample B is more restrictive than subsample A. In subsample C, we restrict the sample to CEOs who began their tenure in or after 1982, i. e., subsample C is a subset of subsample B. All variables are defined in Appendix 2.6. Standard errors, clustered at the state-of-incorporation level, are shown in brackets. *, **, and *** denote significance at the 10, 5, and 1 percent level, respectively.

APPENDIX-TABLE D.5

EXCLUDING DE OR NY INCORPORATED, BANKING, OR UTILITY FIRMS

Dependent Variable: $Death_{i,t}$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Excluding DE/NY Firms				
I(BC)	-0.248***	-0.347***	-0.387***	-0.373***				
	[0.074]	[0.112]	[0.126]	[0.125]				
BC					-0.035**	-0.043***	-0.042**	-0.039**
					[0.015]	[0.017]	[0.020]	[0.020]
Number of CEOs	738	738	738	738	738	738	738	738
Observations	22,103	22,103	22,103	22,103	22,103	22,103	22,103	22,103
				Panel B: Excluding Banking Firms				
I(BC)	-0.286***	-0.319***	-0.332***	-0.320***				
	[0.072]	[0.077]	[0.084]	[0.082]				
BC					-0.056***	-0.060***	-0.057***	-0.057***
					[0.007]	[0.007]	[0.007]	[0.007]
Number of CEOs	1,328	1,328	1,328	1,328	1,328	1,328	1,328	1,328
Observations	42,322	42,322	42,322	42,322	42,322	42,322	42,322	42,322
				Panel C: Excluding Utility Firms				
I(BC)	-0.202***	-0.252***	-0.242***	-0.231***				
	[0.056]	[0.072]	[0.077]	[0.076]				
BC					-0.039***	-0.044***	-0.040***	-0.039***
					[0.005]	[0.005]	[0.005]	[0.005]
Number of CEOs	1,422	1,422	1,422	1,422	1,422	1,422	1,422	1,422
Observations	45,017	45,017	45,017	45,017	45,017	45,017	45,017	45,017
Age (Linear Control)	Y	Y	Y	Y	Y	Y	Y	Y
Location FE (HQ)		Y	Y	Y		Y	Y	Y
Year (Linear Control)		Y	Y			Y	Y	
FF49 FE			Y	Y			Y	Y
Year FE				Y				Y

NOTE. — This table reports hazard coefficients estimated as in Table VI with the sample restricted by states of incorporation or industries. In Panel A, we exclude firms that are incorporated in Delaware or New York (the two most common states of incorporation in our sample, see Table II). In Panel B, we exclude firms that are classified as "Banking" firms in the Fama-French 49 industry classification. In Panel C, we exclude firms that are classified as "Utilities" firms in the Fama-French 49 industry classification. Controls and fixed effects for all three panels are indicated at the bottom of the table. All variables are defined in Appendix 2.6. Standard errors, clustered at the state-of-incorporation level, are shown in brackets. *, **, and *** denote significance at the 10, 5, and 1 percent level, respectively.

APPENDIX-TABLE D.6
NONLINEAR EFFECTS AND PREDICTED EXPOSURE

Dependent Variable: $Death_{i,t}$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$BC^{(\min\text{-}p50)}$	-0.081***	-0.097***	-0.091***	-0.088***				
	[0.020]	[0.023]	[0.026]	[0.025]				
$BC^{(p51\text{-}\max)}$	-0.009	-0.008	-0.007	-0.008				
	[0.014]	[0.015]	[0.017]	[0.017]				
$\widehat{BC}$					-0.052***	-0.059***	-0.050**	-0.049**
					[0.020]	[0.018]	[0.024]	[0.024]
Age	0.105***	0.104***	0.113***	0.114***	0.108***	0.105***	0.115***	0.115***
	[0.006]	[0.006]	[0.004]	[0.004]	[0.008]	[0.009]	[0.008]	[0.008]
Year		0.007*	0.004			0.005	0.001	
		[0.004]	[0.004]			[0.006]	[0.004]	
Location FE (HQ)		Y	Y	Y		Y	Y	Y
FF49 FE			Y	Y			Y	Y
Year FE				Y				Y
Number of CEOs	1,605	1,605	1,605	1,605	1,605	1,605	1,605	1,605
Observations	50,530	50,530	50,530	50,530	50,530	50,530	50,530	50,530

NOTE. — This table shows hazard coefficients estimated from a Cox (1972) proportional hazards model. The dependent variable is an indicator that equals one if the CEO dies in a given year. The main independent variable in the left four columns is $\widehat{BC}$, a count variable of years of predicted cumulative exposure to a BC law. The main independent variables in the right four columns are $BC_{i,t}^{(\min\text{-}p50)}$ and $BC_{i,t}^{(p51\text{-}\max)}$, which capture BC law exposure up to the sample median and incremental exposure above the median, respectively. All variables are defined in Appendix 2.6. For the left four columns, we present standard errors clustered at the state-of-incorporation level, in brackets. For the right four columns, we present bootstrapped standard errors, using the block bootstrap method with 500 iterations, in brackets. *, **, and *** denote significance at the 10, 5, and 1 percent level, respectively.

APPENDIX-TABLE D.7

BUSINESS COMBINATION LAWS AND CEO PAY

Dependent Variable: ln($Pay_{i,t}$)			
	(1)	(2)	(3)
I(BC)	0.086	0.087*	0.041
	[0.058]	[0.047]	[0.051]
Age Controls	Y	Y	Y
Tenure Controls		Y	Y
Firm Characteristics		Y	Y
Location FE (HQ)	Y	Y	
FF49 FE	Y	Y	
Year FE	Y	Y	Y
Firm FE			Y
Number of CEOs	1,553	1,553	1,553
Observations	17,719	17,719	17,719

NOTE. — The table shows OLS estimates where the dependent variable is the logarithm of a CEO's total pay in a given year. In column (1), "Age Controls" includes linear age, and in columns (2) and (3) it includes linear and quadratic age. "Tenure Controls" includes linear and quadratic tenure. "Firm Characteristics" includes logarithms of asset size and the number of employees. Standard errors, clustered at the state-of-incorporation level, are shown in brackets. *, **, and *** denote significance at the 10, 5, and 1 percent level, respectively.

APPENDIX-TABLE D.8
BUSINESS COMBINATION LAWS AND TENURE

Dependent Variable: $Retirement_{i,t}$				
	(1)	(2)	(3)	(4)
I(BC)	-0.222***	-0.094		
	[0.068]	[0.106]		
BC			-0.098***	-0.044**
			[0.021]	[0.021]
Age	0.100***	0.099***	0.101***	0.099***
	[0.010]	[0.010]	[0.010]	[0.011]
Year	0.069***		0.095***	
	[0.007]		[0.014]	
Location FE (HQ)	Y	Y	Y	Y
FF49 FE	Y	Y	Y	Y
Year FE		Y		Y
Number of CEOs	1,605	1,605	1,605	1,605
Observations	17,864	17,864	17,864	17,864

NOTE. — This table shows hazard coefficients estimated from a Cox (1972) proportional hazards model. The dependent variable is an indicator that equals one if the CEO steps down in a given year. All variables are defined in Appendix 2.6. Standard errors are clustered at the state-of-incorporation level, in brackets. *, **, and *** denote significance at the 10, 5, and 1 percent level, respectively.

Chapter 3

Practical Finance: An Approximate Solution to Life-Cycle Portfolio Choice

James Choi and Canyao Liu

Abstract

We provide an economically interpretable and easy-to-calculate approximation to optimal portfolio choice over the life cycle. The standard literature that solves the numerical optimal portfolio policy requires complicated backward induction, making it hard to apply for providing financial advice. Real-world financial advisors, on the other hand, tend to neglect the risky nature of human capital and offer advice that is not truly optimal. We bridge the gap by first using a reduced-form regression to predict discount rates of future incomes over an agent's life. Our prediction method achieves an R^2 of more than 90% over a wide range of simulations. Furthermore, by plugging the discount rates we predict into Merton (1969) formula, we obtain an approximate solution that has an average difference within 2% when compared to optimal solution solved through backward induction.

1 Introduction

Popular personal investment advice offered by financial advisors or public media tend to neglect an important source of assets owned by working population – their labor income. In academic literature, theoretical results have been derived to answer the question of what would be the optimal portfolio allocation rule in terms of the ratio between risky assets like stocks and risk-free assets like Treasury bonds (Merton (1969)). Cocco et al. (2005) ("CGM (2005)") numerically solve the optimal portfolio policies under reasonable assumptions of life-cycle income process. Their method involves estimating the labor income process for the general population based on survey data and then solving the optimal solutions using backward dynamic calculation. Due to the complication in setting up the dynamic calculation process, the results of Cocco et al. (2005) are not popularized among financial advisors.

In this paper, we aim to bridge this gap by offering a practical approximation to the problem of optimal portfolio choice under risky labor income. Our target approximation function depend only on the structural parameters provided by each individual including their expected future income process and their attitude towards risk.

We start by carrying out extensive simulations using the Cocco et al. (2005) approach. We vary the parameters including relative risk aversion, permanent income shocks, temporary income shocks, riskfree rates, risk premiums, retirement incomes, and labor income paths. In total, we implement simulations under 2,187 parameter sets and achieve the optimal portfolio choice for each age and each cash-on-hand in each parameter set.

For each parameter set and each age, we estimate a discount rate that the agent uses to discount future cash flows. We estimate the discount rate by minimizing the squared distance between the optimal portfolio choice and the portfolio allocation we can calculate using Merton (1969)'s formula: $\alpha = \alpha^*(1 + \frac{H}{W})$. The key here is that the discount rate will affect the discounted human capital in the future. We observe that the discount rates

that minimize the squared distance have the following pattern over the life cycle: it first start high early in life, decreases approximately in a quadratic way to a bottom mid-life, rises back up in a quadratic way near retirement, and drops to nearly riskfree rate post retirement.

Observing these features of the life-cycle discount rates, we lay out two prediction models to predict the discount rates. For pre-retirement periods, we use a regression model that incorporates linear terms of the structural parameters like risk aversion and income shock variances, and linear and quadratic age terms, as well as the interaction terms between age and those structural parameters. The R^2 for the regression estimated across all the observations (more than 10 million) reaches 0.93. For post-retirement periods, we use a regression model that includes linear terms of structural parameters and linear age and their interaction, and the R^2 reaches 0.91 across all observations. The good prediction of discount rates leads to the good performance of our approximate solution. The average difference across all the parameter sets and across the life cycle is below 2% between our approximate solution and the solved optimal portfolio policy through backward induction. Overall, our approximate solution is close to the real optimal solution, easy to calculate, and straightforward to interpret.

Our work contributes to the literature on personal financial choice. Ever since the seminal work of Samuelson (1975), economists have developed a variety of theoretical frameworks in analyzing the optimal portfolio choice over the life cycle. Samuelson (1975) solve the life-time portfolio choice in the absence of labor income. Merton (1969) and Merton (1975) provide analytical solutions to portfolio choice in continuous-time and with the consideration of labor income path. In more recent work, Duffie et al. (1997), Koo (1998), and Viceira (2001) study how the portfolio choice behaves under infinite-horizon and uninsurable labor income risk. Insightful as those theoretical works are, they are hard to be used to offer direct quantitative guidance on real-life financial decisions. This need prompts the literature of numerical studies of optimal portfolio choices. Koo

(1999) analyzes the properties of optimal portfolio policies in a discrete-time model and provide numerical solutions under simplified assumptions of labor income. Cocco et al. (2005) solve the numerical optimal portfolio policies under estimated labor income paths and study its behavior under different scenarios. Huggett and Kaplan (2016) estimate the stock component in human capital and study its impact on portfolio choice. Even though those papers offer quantitative solutions for life-cycle portfolio choices, they have not yet been widely adopted in real-world financial advice due to numerical complexity and the challenge to interpret with economic intuition. Popular financial advice like Malkiel (1999), however, tend to neglect the risky nature of human capital. Our paper connects these two together by providing an easily calculated approximate solution that closely tracks the actual optimal solution from backward induction.

The rest of the paper is organized as follows. Section 2 illustrates the problem set-up and our simulation method. Section 3 demonstrates how we calculate the discount rates over the life cycle. Section 4 documents how we predict the discount rates through regression and use the predicted discount rates to achieve our approximate solution. Section 5 concludes and talks about future extensions.

2 Optimal Portfolio Choice: CGM (2005) Approach

1. Problem Set-up

We study the portfolio choice over the life cycle, where an agent needs to choose their consumption level as well as the allocation of his wealth between a risk-free asset and a risky asset. The mathematical formulation of the agent's problem is as follows.

Let t denote an adult's age[1]. The investor lives for a maximum of T periods and he works up until age K. In our main analysis, we assume no mortality risk[2]. Investor i 's

[1]The first period we look at in the life cycle is age 22.

[2]Mortality risk can be incorporated into our framework in the manner of Hubbard et al. (1995) by adding the probability of being alive in the dynamic choice formulation. In fact, as we show in the Appendix Figure A.1, mortality risk only affects the portfolio choices to a limited extent.

preferences are described by the time-separable power utility function:

$$E_1 \sum_{t=1}^{T} \delta^{t-1} \left(\frac{C_{it}^{1-\gamma}}{1-\gamma} \right)$$

where $\delta < 1$ is the discount factor, C_{it} is the level of date t consumption, $\gamma > 0$ is the coefficient of relative risk aversion.

Utility Function Even though we focus on power utility in our current simulations described below. Our simulation space can be easily generalized to use the time-separable utility function of Epstein and Zin (1991), where the recursive formulation of intertemporal utility disentangles risk aversion ($\gamma > 0$) and intertemporal substitution ($\psi > 0$) :

$$U_{it} = \left\{ (1-\delta) C_{it}^{1-(1/\psi)} + \delta \left(E_t U_{i,t+1}^{1-\gamma} \right)^{\frac{1-(1/\psi)}{1-\gamma}} \right\}^{\frac{1}{1-(1/\psi)}}.$$

Power utility situation is then nested as the special case where $\psi = 1/\gamma$. In the absence of a bequest motive, the terminal condition for the recursion is $U_{i,T+1} = 0$.

Labor Income Process During the working periods of life, we assume that log labor income is the sum of a deterministic component and two random components. Specifically, we assume that Y_{it} follows the process below:

$$log(Y_{it}) = a_0 + a_1 * t + a_2 * t^2 + a_3 * t^3 + v_{it} + \varepsilon_{it} \quad \text{for } t \leq K$$

where v_{it} is the permanent shock term and ε_{it} is the temporary shock term.
We assume that v_{it} follows a random walk process:

$$v_{it} = v_{it-1} + u_{it}$$

Here, u_{it} is i.i.d. each period, and $u_{it} \sim N(0, \sigma_u^2)$. We also have $\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$.

Cocco et al. (2005) further assume that the temporary shock ε_{it} is uncorrelated across households. On the other hand, they decompose the permanent shock u_{it} into an aggregate component ξ_t (distributed as $N\left(0, \sigma_\xi^2\right)$) and an idiosyncratic component ω_{it} (distributed as $N\left(0, \sigma_\omega^2\right)$): $u_{it} = \xi_t + \omega_{it}$.

For post-retirement periods, we assume that the post-retirement income Y_{it} is a constant fraction λ of permanent labor income in the last working-year:

$$\log\left(Y_{it}\right) = \log(\lambda) + \left(a_0 + a_1 * K + a_2 * K^2 + a_3 * K^3\right) + v_{iK} \quad \text{for } t > K.$$

Cocco et al. (2005) calibrate the model to match the real labor income data from Panel Study of Income Dynamics (PSID) for three different education groups: college graduates, high school graduates, and no high school. In our simulation studies, we borrow their estimates of labor income processes for those three education groups.

Financial Assets We assume that the single risky asset return follows a log normal distribution where:

$$log(R_t) \sim N(r, \sigma_r^2)$$

The riskfree asset yields a fixed return that satisfies $log(R_f) = r_f$, and the log risk premium of the risky asset is then $r - r_f$.

In addition, we assume that the risky asset return has no correlation with labor income shocks. In our simulation set-up, this can be generalized to assuming a fixed correlation between labor income shock and stock return with ease. We also assume that there is no borrowing and short-sale allowed.

Dynamic Optimization In each period t, the agent enters the period with wealth of W_{it}. Then he receives his labor income Y_{it}. We call cash-on-hand in period t as $W_{it} + Y_{it}$. Then the agent needs to decide his consumption this period, C_{it}, and how he should allocate the

remaining cash-on-hand excluding consumption between the risky and riskless asset. In period $t+1$, before earning new income, his wealth is then:

$$W_{i,t+1} = \left(\alpha_{it} R_{t+1} + (1-\alpha_{it}) R_f\right)\left(W_{it} + Y_{it} - C_{it}\right),$$

where α_{it} is the share he allocates to the risky asset.

The problem the investor faces is to maximize future sum of discounted utility while subject to the borrowing and short-sale constraints. The policy variables we want to solve for are the consumption choice C_{it} and the portfolio choice α_{it}, conditional of the state variable cash-on-hand. Given the set-up, the value function is homogeneous with respect to current permanent labor income. We exploit this scaleability to do the normalization. Then the Bellman equation is:

$$V_{it}\left(X_{it}\right) = \max_{C_{it}\geq 0, 0\leq\alpha_{it}\leq 1}\left[U\left(C_{it}\right) + \delta E_t V_{i,t+1}\left(X_{i,t+1}\right)\right] \quad \text{for } t < T,$$

where

$$X_{i,t} = W_{i,t} + Y_{i,t}$$

$$X_{i,t+1} = Y_{i,t+1} + \left(X_{it} - C_{it}\right)\left(\alpha_{it} R_{t+1} + (1-\alpha_{it}) R_f\right).$$

2. Solution Method

Parameterization We assume that people live up until age 100, i.e., $T = 79$ (from age 22 to age 100); and we assume that people retire at age 66, i.e., $K = 66$. We list out the parameters we use for labor income processes of different education groups in Table 1 below. As mentioned above, those numbers come from Cocco et al. (2005).

For discount factor, we assume that $\delta = 0.96$. For the standard deviation of the risky asset return, we assume that $\sigma_r = 0.157$. We vary the rest of the parameters in our simulations and details are illustrated in the next subsection.

Table 1
Labor Income Process Parameters

	a_0	a_1	a_2	a_3	σ_u^2	σ_ε^2	λ
College graduates	-4.3148	0.3194	-0.0577	0.0033	0.0169	0.0584	0.938873
High school graduates	-2.1700	0.1682	-0.0323	0.0020	0.0106	0.0738	0.68212
No high school	-2.1361	0.1684	-0.0353	0.0023	0.0105	0.1056	0.88983

Normalization Due to the fact that the value function is homogeneous with respect to current permanent income, we can normalize the problem to the extent that benefits our calculation. Here, for pre-retirement periods, our normalization unit is the current period's permanent income shock, i.e., we normalize by setting v_{it} to be 1. The normalized income in a pre-retirement period t is then:

$$\widetilde{Y_{it}} = \frac{exp(a_0 + a_1 * t + a_2 * t^2 + a_3 * t^3 + v_{it} + \varepsilon_{it})}{exp(v_{it} - 1)} = exp(a_0 + a_1 * t + a_2 * t^2 + a_3 * t^3 + 1 + \varepsilon_{it})$$

Define the growth rate of the deterministic part of the income between t and $t+1$ as g_{it}:

$$g_{it} = \frac{exp(a_0 + a_1 * (t+1) + a_2 * (t+1)^2 + a_3 * (t+1)^3)}{exp(a_0 + a_1 * t + a_2 * t^2 + a_3 * t^3)} - 1$$

Then the growth rate of permanent income between t and $t+1$ is approximately:

$$gp_{it} = exp(g_{it}) * exp(u_{i,t+1})$$

We can therefore get the formula for calculating the cash-on-hand at period $t+1$ using the savings from period t and the income at $t+1$:

$$Cash_{i,t+1} = Saving_{i,t} * Return_{t+1} / gp_{i,t+1} + exp(\varepsilon_{i,t+1})$$

This will be the key relationship we use over different time periods, and the derivation

of this formula is detailed in Appendix A.2.

Grid search for policies To solve the optimal portfolio choice and optimal consumption choice, we implement grid search method. In particular, we use the following search sets and search grids for our state and policy variables:

- Cash-on-hand range: in each period t before retirement, we have 201 grids ranging from 0.25 to 50 multiplier of last period $(t-1)$'s permanent labor income. In each period t after retirement, we have 201 grids ranging from 0.25 to 50 multiplier of the last period of working life's income (age 66's income). We solve for the consumption and portfolio policies for each age and each grid of cash-on-hand.

- Consumption policy: for each period t, we search the consumption policy starting from the lowest cash-on-hand value, and for each increase in cash-on-hand value, the search range assumes monotonically increasing consumption with respect to cash on hand. We have 101 grid points in total for consumption.

- Portfolio policy: for each period t, we simultaneously search for the consumption policy and portfolio policy. After fixing a consumption grid, we search for the portfolio choice by allocating the cash excluding consumption to risky asset and riskfree asset. Here, similar to consumption, we also assume that the allocation to risky asset monotonically increase as the cash on hand increases. We have 101 grids in total and each grid increase is 1%.

Value function interpolation and extrapolation At age 100, which is the last period of life, we calculate the value function at each cash-on-hand grid directly using the utility function formula.

For each age t before 100, for each consumption choice and portfolio choice and each realization of the risky asset return, we will end up with a cash-on-hand level for the next

period $t+1$. This cash-on-hand level may not end up on any given cash-on-hand grid. In this case, we use interpolation and extrapolation method to calculate the value function value for this off-the-grid cash-on-hand. The interpolation works as follows: if the cash-on-hand falls in between two grid points and is below 45, we use the standard cubic spline method to interpolate the value function value at that cash-on-hand; if the cash-on-hand value falls below 0.25, we use the cubic spline method to extrapolate the value function value[3]; if the cash-on-hand value goes above 45, we use a linear extrapolation based on the 196th and 197th grid point, leaving the 5 largest grid points' value to be extrapolated[4]. We verify that at age 99, this method yields almost identical portfolio choice compared to using exact value function at age 100, as shown in Appendix Figure A.6. Furthermore, we compare the portfolio choice solutions at different ages over the life cycle when we use 196th grid and 197th grid to extrapolate versus when we use 191st and 192nd grid to extrapolate and show in Appendix Figure A.7 that they are also nearly identical. Those provide proof of safety to our linear extrapolation method.

Approximation methods With the existence of labor income shocks and uncertain risky asset returns, we need to use approximation methods to calculate the expected value function. To reach this, we apply 9-grid Gaussian quadrature as documented in Tauchen and Hussey (1991).

3. Simulations

Our aim is to find a general reduced-form formula for approximating the optimal solution of portfolio choice over the life cycle. To reach this goal, we need to implement a large set of simulations to incorporate variations across all the important structural pa-

[3]In reality, this situation is not able to happen. The reason is that the cash-on-hand already incorporate this period's income and the lowest possible value is still above 0.25.

[4]The reason we leave 5 grid points out is that the largest grid points tend to have imprecise value function values due to sparsity of grids at the right end. Leaving those 5 points and using extrapolation each time helps improve stability.

rameters. Specifically, we carry out extensive simulations with varying income shocks, income paths, asset returns, etc. In total, we implement simulations across 2,187 sets of parameters. The choices of our parameters aim to reflect the scenarios most common faced in reality. Details of our parameter choices are listed below:

- Relative risk aversion (RRA): values to take are 4, 6, 10. The corresponding EIS under Power utility are then: 0.25, 1/6, 0.1.
- The standard deviation of the permanent income shocks: values to take are 0.13, 0.1029563, 0.1024695, corresponding to the three education levels, i.e., college, high school, and no high school.
- The standard deviation of the temporary income shocks: values to take are 0.24166, 0.27166, 0.32496, corresponding to the three education levels as well.
- The retirement income fraction as a percentage of the income at the last age before retirement: values to take are 0.938873, 0.68212, 0.88983, corresponding to the three education levels as well.
- The deterministic part of the labor income path over the life cycle: the three paths correspond to three income levels.
- Risk-free rates: values to take are 2%, 1%, and 0%.
- Risk premium of the risky asset: values to take are 4%, 3%, and 2%.

After achieving the solutions for each age and each cash-on-hand under each parameter set, we combine them together for further analysis.

4. Solutions

We obtain portfolio solutions under different scenarios. Different as they are, they share qualitative similarities. Here we show one set of life-cycle portfolio choices under RRA

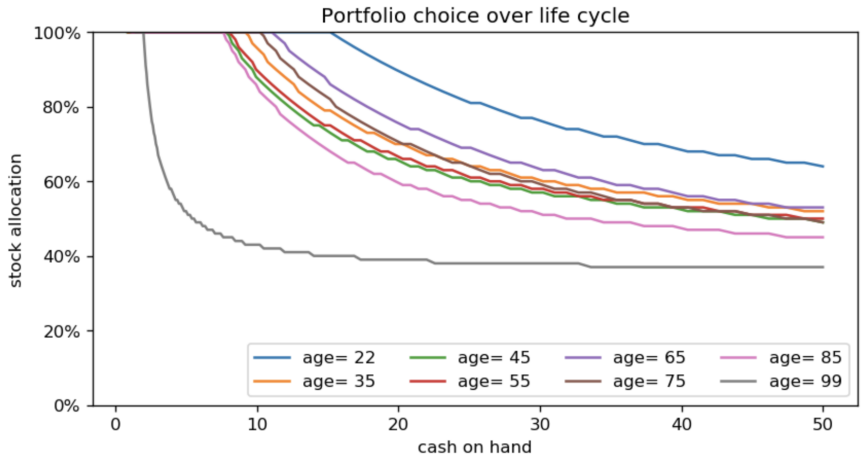
being 6, risk-free rate being 2%, risk premium being 4%, and assuming college education for labor income process.

Figure 1 shows portfolio choices from early in life (age 22) to the second to last living life (age 99), with an approximate 10-year gap between each age. For age 99, in theory we can solve the exact solution for portfolio choice. Without labor income, for risky asset with log-normal return, the portfolio choice will be:

$$\alpha^* = \frac{r - r_f + \frac{1}{2}\sigma_r^2}{\mu\sigma_r^2}$$

With the parameters in Figure 1, we can solve α^* to be 35.38%. When the cash-on-hand at age 99 increases to large values, the next-period labor income is nearly negligible. Therefore, the age-99 portfolio policy should approximate to α^*, which is the case as shown in our figure. For other retirement ages, there is no uncertain in future incomes and the discounted human capital is monotonically decreasing as the age increases. Therefore, for the same cash-on-hand, the investor tends to take more risk at younger age. Again, as the cash-on-hand increases, the portfolio choice should gradually approach α^*.

Figure 1. PORTFOLIO CHOICE OVER THE LIFE CYCLE



Notes. The figure shows the portfolio choice for age 22, 35, 45, 55, 65, 75, 85, and 99. The parameters for this scenario are: RRA = 6, R_f = 2%, $Risk\ premium$ = 4%, and labor income of college graduates.

For pre-retirement ages, the situation is more complicated as the labor income process is hump-shaped. Early in life, expecting future income to be very high, the agent tends to take on more risk. Therefore, we see that the portfolio share of the risky asset is high early in life and then gradually decrease. After mid-life, even though the agent's expected income is getting lower, but the risk of labor income is decreasing, too. The increasing certainty of future income makes the agent more comfortable to take on risk. So, after decreasing the share of risky asset up to certain age in mid-life, there is an increase of risky asset's share over the later periods before retirement. Then, when getting close to retirement, the share starts to decrease again. All of those patterns are captured in Figure 1.

3 Effective Discount Rates over Life Cycle

1. BMS Approximation

Merton (1969) and Merton and Samuelson (1992) derive under the continuous-time model that the portfolio allocation to the risky asset can be expressed as:

$$\alpha = \alpha^* * (1 + \frac{H}{W})$$

where W is the current period's wealth (cash-on-hand minus consumption) and H is the discounted human capital from future periods.

Our approximate solution for the optimal portfolio choice over the life cycle follows this heuristic. For each age and each cash-on-hand value, we conjecture that the optimal solution can be approximated by the formula above with proper choice of discount rates for human capital.

The key question is: how to find the proper discount rate at each age? We carry out this procedure dynamically through backward calculation. Specifically, we start from age

99. At age 99, the future income only comes from age 100, then we search for a discount rate $r_{d,99}$ such that the sum of squared distance between the formula implied portfolio policy and the solved portfolio policy from CGM (2005) method at each cash-on-hand is the smallest. In other words:

$$r_{d,99} = argmin_r \sum_{i=1}^{201} (\alpha^* * (1 + \frac{Income_{100}/(1+r)}{Cash_i - Consumption_{99,i}^{CGM}}) - \alpha_{99,i}^{CGM})^2$$

For the discount rate at age 98, there are two future incomes. For age 100's income, we discount it first by the solved age-99 discount rate and then by the age-98 discount rate. For age 99's income, we discount it only the age-98 discount rate. Again, we search for the discount rate at age 98 by minimizing the squared distance between formula implied policy and the CGM (2005) solved policy at age 98:

$$r_{d,98} = argmin_r \sum_{i=1}^{201} (\alpha^* * (1 + \frac{Income_{100}/((1+r_{d,99})(1+r)) + Income_{99}/(1+r)}{Cash_i - Consumption_{98,i}^{CGM}}) - \alpha_{98,i}^{CGM})^2$$

Similarly, we calculate the discount rate at each age before 98 dynamically, age by age. For age t, we have:

$$r_{d,t} = argmin_r \sum_{i=1}^{201} (\alpha^* * (1 + \frac{\sum_{j=t+1}^{100} \frac{Income_j}{(1+r)\prod_{m=t+1}^{99}(1+r_{d,m})}}{Cash_i - Consumption_{t,i}^{CGM}}) - \alpha_{t,i}^{CGM})^2$$

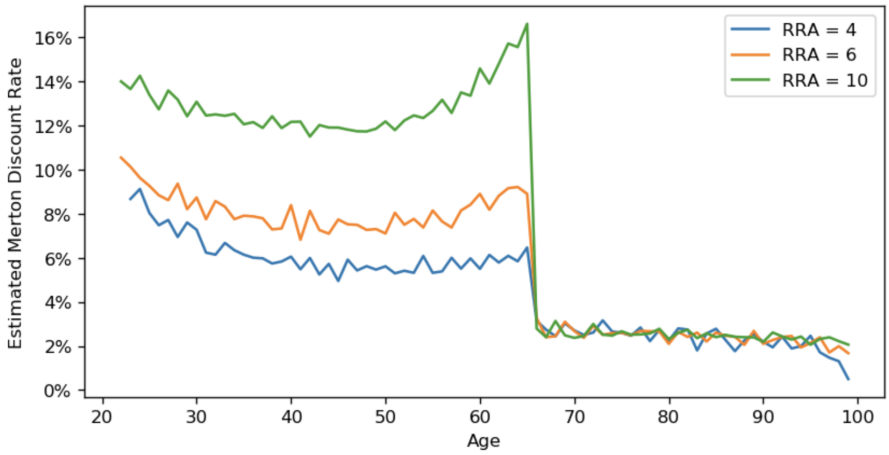
2. Discount Rate Pattern over Life Cycle

Based on the method we outline above, we can solve the discount rate over the life cycle. Figure 2 below plots the discount rates for three sets of parameters.

In all three scenarios, we assume that the agent is a college graduate, the riskfree rate is 2%, and the risk premium is 4%. The main differences among them is what we assume the relative risk aversion to be. The green line is for RRA being 4; the orange line is for RRA being 6, and the blue line is for RRA being 10. We can see that the three lines show

similar qualitative pattern over the agent's life. For post-retirement periods, we see that the effective discount rates are about 2%, which is the riskfree rate. This makes intuitive sense because post retirement, there is no uncertainty in terms of future incomes. The incentive behind the discount rate is mainly just driven by the riskfree rate and consumption-saving motives. Besides, in post-retirement periods, we also see a gradual slight decline of the discount rate as the agent ages. This gradual decline reflects the agent's early consumption motive: when the future incomes are higher, the agent tends to put higher discount rate and have more incentive to consume early.

Figure 2. DISCOUNT RATES ESTIMATED FOR DIFFERENT RRA



Notes. The figure shows the calculated discount rates over the life cycle. The parameters for this scenario are: $R_f = 2\%$, $Risk\ premium = 4\%$, and labor income of college graduates. We plot the scenarios for RRA being 4, 6, and 10.

In the pre-retirement periods, the discount rates first start high early in life. Expecting highly uncertain future income flows, the agent demands a high discount rate. As the agent ages, more permanent income shocks get settled and the future income flows are gradually getting more stable. This reduced uncertainty pushes the discount rate to reach a bottom in mid-life, around age 50. As the agent passes mid-life and faces retirement, the uncertainty bounces up again because we assume that the retirement income is a fixed proportion of the last-period permanent income and this assumption means that the realization of the last-

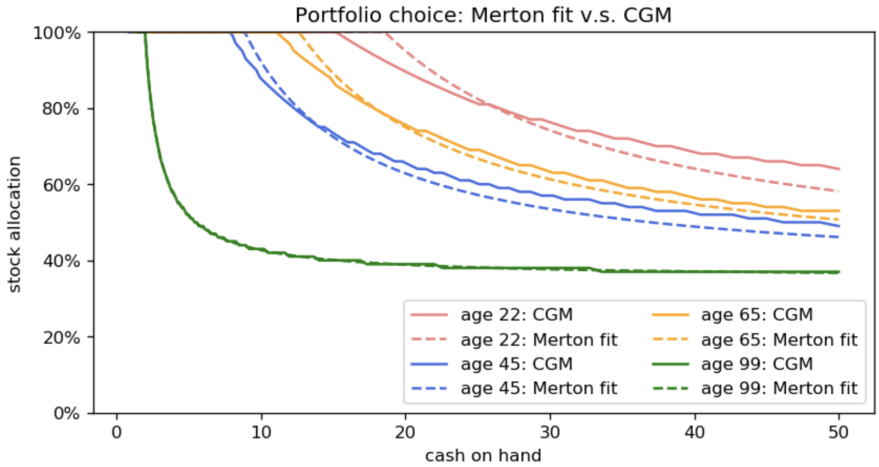
period labor permanent income will determine a large chunk of the agent's future wealth. Due to this concern, the agent demands a high discount rate again, which explains the rise in the periods toward retirement. Overall, we see a inverse hump-shaped curve before retirement and a linearly declining line post retirement. These motivate our approximation methods detailed below.

With the discount rates we solved, we can take them back into the BMS formula to calculate the "BMS approximation" of the portfolio choice over the life cycle. For the portfolio choice at age t, we have:

$$\alpha_{i,t}^{BMS} = \alpha^* * (1 + \frac{\sum_{j=t+1}^{100} \frac{Income_j}{\prod_{m=t}^{99}(1+r_{d,m})}}{Cash_i - Consumption_{t,i}^{CGM}})$$

Figure 3 depicts the comparison between the CGM (2005) solution and the BMS approximation for age 22, 45, 65, and 99 under the situation that $RRA = 6$, R_f = 2%, $Risk\ premium$ = 4%, and being a college graduate. We can see that the BMS approximation tracks the CGM (2005) solution well, especially for later periods in life.

Figure 3. MERTON FIT V.S. CGM SOLUTION



Notes. The figure shows the BMS approximation and the CGM (2005) solution for age 22, 45, 65, and 99. The parameters for this scenario are: $RRA = 6$, R_f = 2%, $Risk\ premium$ = 4%, and labor income of college graduates.

The closeness of the BMS approximation and the CGM (2005) solution gives us confidence in terms of providing an economically motivated and easily calculated approximation for portfolio choice. In the later part, we use the structural parameters we have to predict the discount rates over the life cycle and use the predicted discount rates to feed in the BMS formula, which we will denote as Choi and Liu approximation ("CL approximation").

3. Shape Sensitivity

In this subsection, we analyze how the shape of the discount rate pattern over the life cycle varies with respect to the parameters we set up.

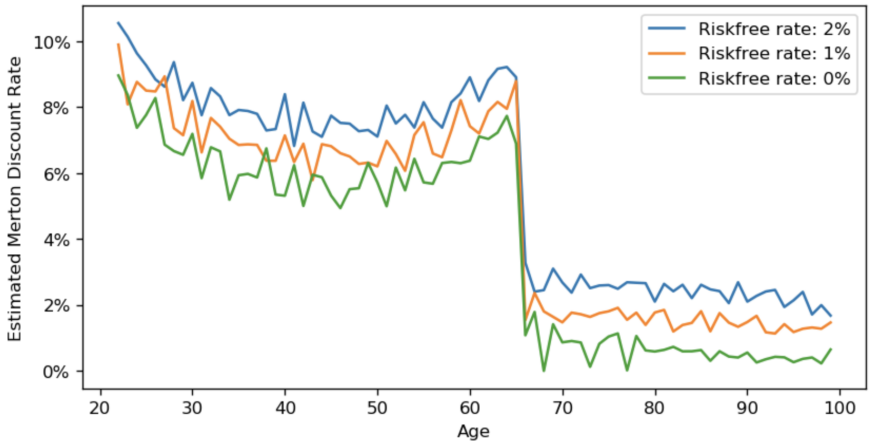
Figure 4 shows the discount rate patterns when we vary the riskfree rates among 0%, 1%, and 2%. We maintain the rest of the parameters the same. We can see that this shift in riskfree rates leads to almost an identical level shift of the entire discount rate curve. This makes intuitive sense as the riskfree rate serves as a benchmark for the agent to demand for discount. In Appendix Figure A.3 and Appendix Figure A.4, we further provide illustration on how the discount rates vary with respect to education level and risk premium. These variations provide support that by exploiting those structural parameters, we may be able to predict what the discount rate pattern should look like, and this is what we pursue in the following section.

4 An Approximate Solution to Life-Cycle Portfolio Choices

1. Discount Rate Prediction Method

In the aim of finding an easily interpretable and calculated portfolio choice over the life cycle, we build up a prediction model for the discount rates. After getting the predicted values of the discount rates, we use them in the BMS (1992) formula to calculate the discounted human capital at each age, and then calculate the predicted portfolio choice at

Figure 4. DISCOUNT RATES ESTIMATED FOR DIFFERENT RISKFREE RATES



Notes. The figure shows the discount rate patterns when we set the riskfree rate to 0%, 1%, and 2%, respectively. The rest of the parameters for the figure are: $RRA = 6$, $Risk\ premium$ = 4%, and labor income of college graduates.

each cash-on-hand value and in each age.

We exploit our simulated portfolio choice data. Recall that we have run 2,187 sets of simulations with varying parameters. For each simulation, we have solutions for each age and each cash-on-hand value. We also have our solved discount rates, which best approximate the CGM solutions through BMS formula, for each set of parameters and each age. We lay out our prediction model for discount rate in two parts: pre-retirement and post-retirement. This division is based on the difference in shape for the two. Specifically, for the pre-retirement, we estimate the following model:

$$\begin{aligned} Y_{it} =& \alpha + \beta_1 \mu_i + \beta_2 \rho_i + \beta_3 \sigma^2_{p,i} + \beta_4 \sigma^2_{t,i} + \beta_5 ret\_frac_i + \beta_6 R_{f,i} + \beta_7 age_{it} + \beta_8 age^2_{it} + \\ & \beta_9 age_{it}(\mu_i + \rho_i + \sigma^2_{p,i} + \sigma^2_{t,i} + ret\_frac_i + R_{f,i}) + \varepsilon_{it}, \end{aligned}$$

where i denotes the index for each parameter set and t denotes the period in the life cycle. Y_{it} represents the discount rate under parameter set i and time period t, μ_i represents the risk premium, ρ_i represents relative risk aversion, $R_{f,i}$ represents the riskfree rate, $\sigma^2_{p,i}$ represents permanent income shock, and $\sigma^2_{t,i}$ represents temporary income shock. The model

features quadratic age terms mimicking what is shown in Figure 2, as well as interaction terms between age and others to incorporate potential nonlinearity.

For post-retirement periods, we estimate the following regression model:

$$Y_{it} = \alpha + \beta_1 \mu_i + \beta_2 \rho_i + \beta_3 R_{f,i} + \beta_4 age_{it} + \beta_5 age_{it}(\mu_i + \rho_i + R_{f,i}) + \varepsilon_{it},$$

and μ_i, ρ_i, and $R_{f,i}$ carry the same meaning as above. For the post-retirement periods, we use this much simpler form as the income shocks and retirement income fraction should not affect how the agent look forward in the retirement periods. The linear form is also motivated by the observation from Figure 2. We add the interaction terms to capture potential nonlinearity.

2. Prediction Performance

Table 2 shows how our prediction model performs for the pre-retirement periods. Specifically, Column (1) in the table only includes the linear terms. With this simple form, we are already able to achieve 0.891 R^2 and adjusted R^2. Note that the regression is estimated over all parameter sets, all ages, and all cash-on-hands. So, achieving this high level of R^2 even with a simple linear model conveys the potential of our approximation. Column (2) adds in the squared term of age, and the R^2 and adjusted R^2 increase to 0.916. This is not surprising as we have already visibly observe the inverse hump-shape of the discount rates. Column (3) further adds in the interaction terms and this boosts the R^2 and adjusted R^2 further to 0.93. This indicates that there is some nonlinearity in the discount rate determination. Furthermore, due to the large number of observations we have, all of our estimated coefficients are at least 1% significant statistically. The sizes of the estimated coefficients also make intuitive sense: for instance, the coefficient for riskfree rate in column (3) is about 1, aligned with what Figure 4 implies.

Table 3 shows how our model performs for the post-retirement periods. Column (1)

Table 2
Merton Discount Rate Prediction for Pre-Retirement Periods

DEPENDENT VARIABLE: $DiscountRate_{i,j,t}$			
	(1)	(2)	(3)
$Constant$	-1.050***	-0.993***	-1.005***
	[0.000]	[0.000]	[0.001]
RRA_j	0.009***	0.009***	0.005***
	[0.000]	[0.000]	[0.011]
$Risk\ Premium_j$	-0.363***	-0.357***	-0.115***
	[0.000]	[0.000]	[0.001]
$Riskfree\ Rate_j$	1.000***	0.996***	1.066***
	[0.000]	[0.000]	[0.001]
$Retirement\ Fraction_j$	0.009***	0.009***	-0.016***
	[0.000]	[0.000]	[0.026]
$Permanent\ Shock_j$	3.885***	3.882***	2.651***
	[0.001]	[0.001]	[0.002]
$Temporary\ Shock_j$	0.020***	0.020***	0.022***
	[0.000]	[0.000]	[0.000]
Age_t	0.000***	-0.003***	-0.002***
	[0.000]	[0.000]	[0.000]
Age_t^2		0.000	0.000
		[0.000]	[0.000]
$Age\ Interactions$			Y
R-squared	0.891	0.916	0.930
Adjusted R-squared	0.891	0.916	0.930
Observations	14,474,911	14,474,911	14,474,911

Notes. This table shows OLS estimates of the pre-retirement regression model. Standard errors are shown in brackets. *, **, and *** denote significance at the 10, 5, and 1 percent level, respectively.

shows the result when we do not include the age related terms. Recall that we observe a slight declining trend over time for post-retirement discount rates. When we do not include the age term in the regression, we are ignoring this trend. We can see that this simple form already achieves good performance with 0.858 R^2 and adjusted R^2. In column (2), when we add in the age term, the R^2 and adjusted R^2 increase to 0.905. When we further

include the interaction terms in column (3), the R^2 and adjusted R^2 further increase to 0.914. Similar to pre-retirement regressions, all the coefficients are statistically significant at at least 1%.

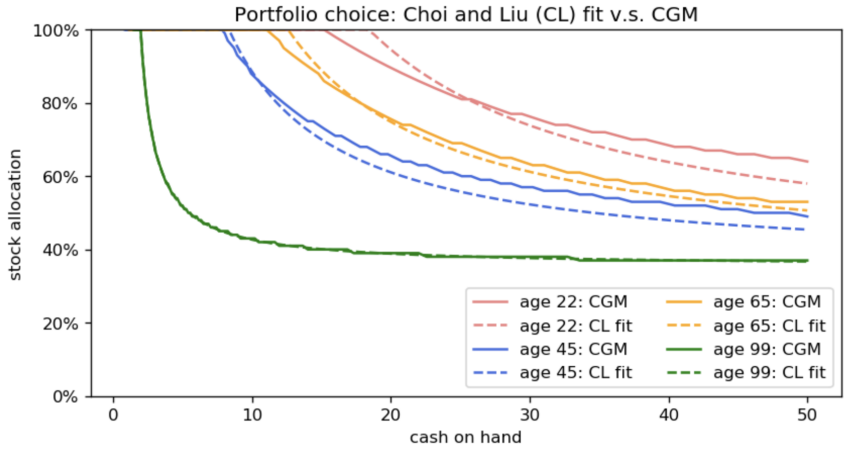
Table 3
Merton Discount Rate Prediction for Pre-Retirement Periods

DEPENDENT VARIABLE: $DiscountRate_{i,j,t}$			
	(1)	(2)	(3)
$Constant$	-0.889***	-0.875***	-1.034***
	[0.000]	[0.000]	[0.001]
RRA_j	0.000***	0.000***	-0.002***
	[0.000]	[0.000]	[0.011]
$Risk\ Premium_j$	0.007***	0.008***	0.388***
	[0.000]	[0.000]	[0.001]
$Riskfree\ Rate_j$	0.894***	0.894***	1.055***
	[0.000]	[0.000]	[0.001]
Age_t		-0.000***	0.002***
		[0.000]	[0.026]
$Age\ Interactions$			Y
R-squared	0.858	0.905	0.914
Adjusted R-squared	0.858	0.905	0.914
Observations	10,720,836	10,720,836	10,720,836

Notes. This table shows OLS estimates of the post-retirement regression model. Standard errors are shown in brackets. *, **, and *** denote significance at the 10, 5, and 1 percent level, respectively.

Besides using R^2 as an evaluation of how our prediction model and approximation method performs, we can also directly compare our CL approximation with the original CGM (2005) solutions. Figure 5 compares the two for a selected set of parameters. We can see that CL approximation tracks the original CGM (2005) solution very well. If we further compare this figure with Figure 3, we can find that the CL approximation is almost as good as the BMS approximation with numerically solved discount rates instead of predicted discount rates.

Figure 5. CHOI AND LIU ("CL") APPROXIMATION V.S. CGM SOLUTION



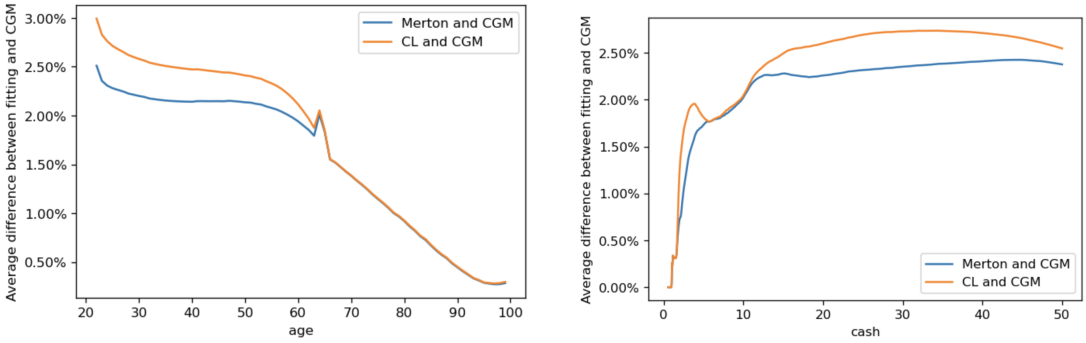
Notes. The figure compares the CL approximation and the CGM (2005) solution for age 22, 45, 65, and 99. The parameters for this scenario are: $RRA = 6$, $R_f = 2\%$, $Risk\ premium = 4\%$, and labor income of college graduates.

Figure 6 shows the compare of the MSE between CL approximation and the CGM (2005) solution and the MSE between BMS approximation and the CGM (2005) solution. Figure 6 (a) demonstrates how the MSEs vary over age. We can see that overall the MSE for CL approximation is not much larger than BMS approximation, i.e., is at most 30% larger. In addition, the gap between the two MSEs shrink as the age increases, and most of the gap concentrates on the pre-retirement periods. Figure 6 (b) illustrates how the MSEs vary over the wealth level. Again, we observe that the gap between the two MSEs is not large. Furthermore, the gap increases with respect to the wealth level and most of the gap happens at the large wealth levels: the wealth levels that are more than 10-time current period's income. In reality, those scenarios are not very common and the good performance at the normal range of wealth levels gives us confidence for our approximation to be useful in offering financial advice.

Figure 6. MSE COMPARISON: CL AND CGM V.S. BMS AND CGM

(a) MSE over Life Cycle

(b) MSE over Wealth Level



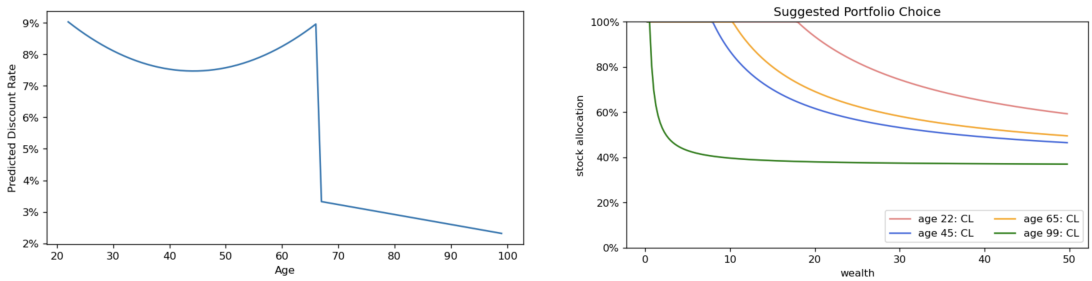
Notes. The figure compares the MSE between CL approximation and the CGM (2005) solution and the MSE between BMS approximation and the CGM (2005) solution.

3. An Illustrative Example

Here we give an illustrative example for how our approximate solution can be used in reality to offer people guidance on investing over the life cycle.

Suppose an individual has an estimated relative risk aversion of 5, and is a fresh college graduate at age 22. At the point of him asking for financial advice, we can estimate the historical average riskfree rate as an expected riskfree rate for the future. For this example, let's set it to 2.5%. For the risky asset return, we can approximate it using the historical return of the S&P 500 index. For instance, let's set its mean as 7% and standard deviation as 18.5%. With those information, we can use our estimated regression to get a fitted path of discount rate over his life cycle, as shown in the left graph of Figure 7. Once we have obtained the projected discount rates, we can put them into the Merton (1969) formula and estimate the life-cycle portfolio choice suited for this individual, as shown in the right graph of Figure 7. For example, we would suggest him to allocate 86.9% of wealth into stocks and 13.1% into Treasury bills at age 45 if he has a wealth of 10-time current income, and to allocate 66% into stocks and X34% into Treasury bills at age 90 with a wealth of 10-time age-66 income, etc.

Figure 7. Protfolio Advice from Choi and Liu Model



Notes: The figure shows how we can provide financial advice given an individual's profile. This left graph shows the predicted discount rate over the life time from our models. The right graph shows our portfolio investment advice for a selected set of ages.

4. Approximation Sensitivity

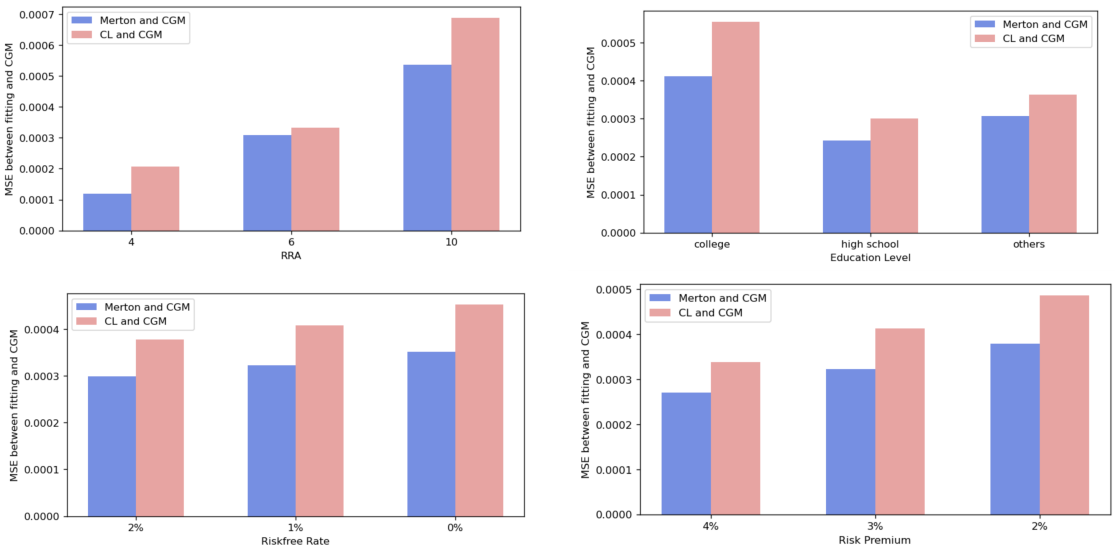
In this subsection, we discuss the sensitivity of our prediction of the discount rates and the corresponding approximation accuracy for portfolio choice with respect to the structural parameters.

Given that we currently set three choices for each parameter, this exercise is only shown as an illustration. Figure 8 shows how the MSEs change with respect to risk aversion, riskfree rate, risk premium, and education level. Qualitatively, it seems that the approximation works slightly differently for different levels of those parameters, indicating further space of improvements for our approximation model. We leave this for future exploration.

5 Conclusion and Discussion

To conclude, we provide an economically interpretable, mathematically straightforward, and easily-to-calculate formula for offering people financial investment advice. Our approximate solution avoids the complicated dynamic calculation in the academic literature but incorporates the essential features of optimal solution from it. We provide a bridge that better suits people's need in real life while at the same time maintains academic rigor.

Figure 8. MSE Variation with respect to Parameters



Notes: This figure shows how the MSEs vary with respect to different structural parameters. The upper left graph shows how the MSEs vary across different RRAs; the upper right graph shows how the MSEs vary with education level; the lower left graph shows how the MSEs change with respect to riskfree rates; and the lower right graph shows how the MSEs vary with the risk premiums.

Once collecting the risk aversion and projected income path from an individual, we are able to provide him suggested portfolio shares of risky asset over the life cycle for him to maximize his lifetime consumption utility.

Our paper opens a line of inquiries, and there are several directions we see for future exploration. For instance, future work can explore variations in the discount factor δ and in mortality rates, which we so far have not considered in our simulations. Next, we can explore results under more generalized utility function form: for example, under the general form of Epstein-Zin utility instead of Power utility. Through that, we can modify the current prediction model to separate the effect of relative risk aversion and elasticity of intertemporal substitution. Next, future work can also use finer grids and larger search ranges for the parameters to see how generalizable our results are, provided the improvement of computation power. Besides, it is worthwhile to use other assumptions of the retirement income: for instance, consider how unemployment can be brought into

the process and how the assumption of fixed amount of money in retirement instead of fixed fraction, which is also a common scheme, affects the results. Lastly, it will also be rewarding to explore how the simplified model can be used to guide the macroeconomic problem beyond the personal finance choice problem.

Appendix

A.1. Appendix Figures and Tables

This appendix presents the appendix figures and tables that are mentioned the main text. The illustrations of the figures and tables are shown in their notes.

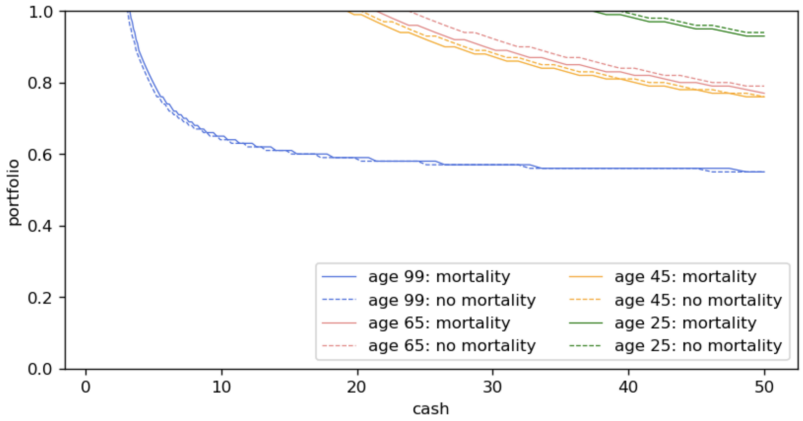
The first appendix figure compares the portfolio policies solved with and without the assumption of mortality over the life time. When we assume that there is mortality, the mortality risk estimates that we use come from Cocco et al. (2005) who cite the data from National Center for Health Statistics.

The second appendix figure depicts how the optimal portfolio policies at different ages solved through CGM (2005) method vary with respect to the structural parameters. Specifically, here, we show the variations across relative risk aversion, education level, riskfree rate, and risk premium. Not surprisingly, when risk aversion increases, the optimal portfolio policy tends to put higher weight on risky asset and the effect is strong across all ages when the relative risk aversion increases by 1. When riskfree rate increases, the optimal portfolio policy tends to put lower weight on risky asset. The effect of the riskfree increasing by 1% is stronger early in life and diminishes over time. When the risk premium increases, the optimal portfolio tends to overweight the risky asset. For the risk premium, the effect is strong across all ages.

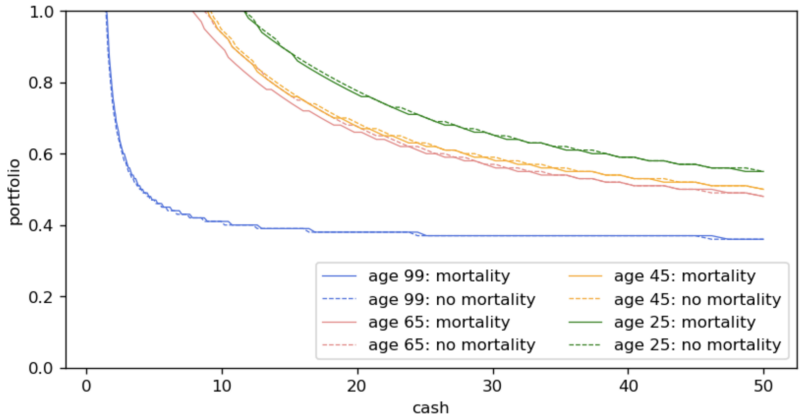
The next two appendix figures document how the estimated discount rates that minimize the squared distance between the Merton (1969) formula and the CGM (2005) solved policies change in response to the variation in education level and risk premium. For education, the discount rate pattern is pretty similar for high school graduates and non high school graduates, whereas the college graduates have significant higher discount rates for pre-retirement periods. The curvature for the college graduates during the pre-retirement periods is also stronger due to larger permanent income shocks in their labor income path.

Figure A.1. COMPARISON BETWEEN CGM-SOLVED PORTFOLIO POLICIES W./W.O. MORTALITY

(a) RRA = 4, College Graduate



(b) RRA = 6, High School Graduate

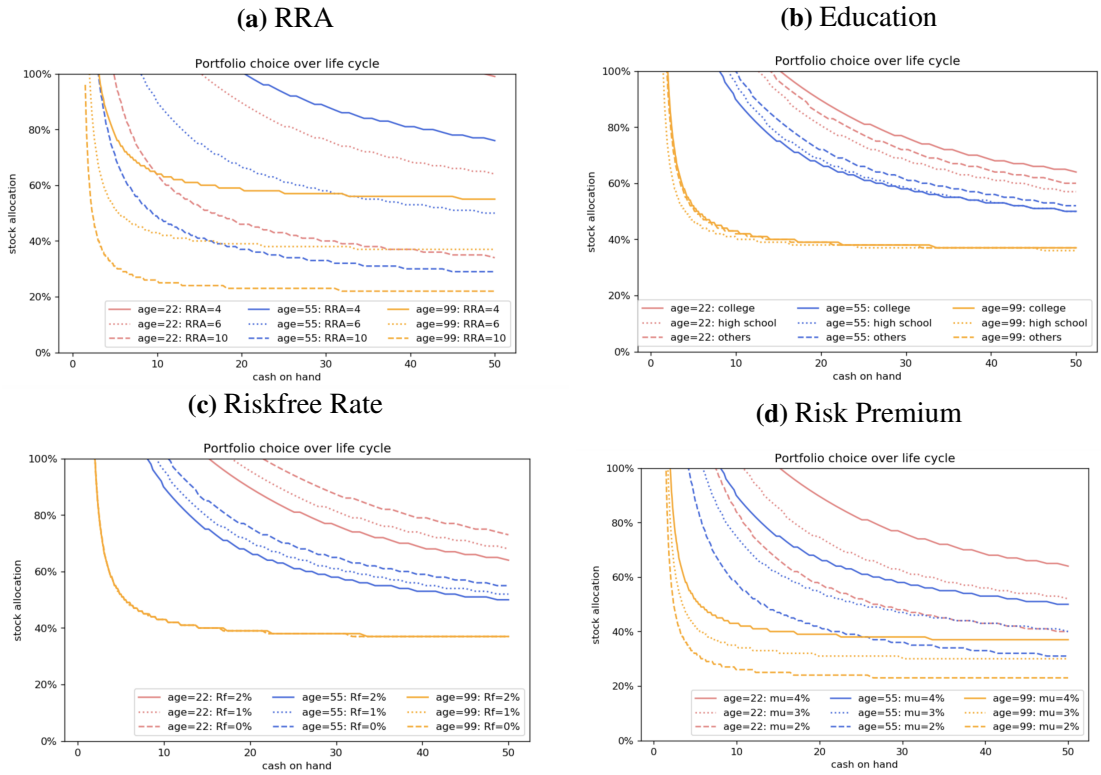


Notes: The figure compares the portfolio policies solved using the CGM (2005) methods with and without the assumption of mortality. For the upper graph, the parameters we use are RRA being 4, riskfree rate being 2%, risk premium being 4%, and the individual being a college graduate. For the lower graph, the parameters we use are RRA being 6, riskfree rate being 2%, risk premium being 4%, and the individual being a high school graduate.

For post-retirement periods, all three groups have almost identical discount rate pattern, which is close to the riskfree rate value.

For risk premium, an increase in risk premium pushes up the demanded discount rate in the pre-retirement periods. However, the size of the increase is not one-to-one like in the case of riskfree rate. This is also resonated in the regression table 2 as the estimated coefficient before risk premium is around 0.38 instead of 1. For post-retirement periods,

Figure A.2. PORTFOLIO CHOICE WITH RESPECT TO CHANGES IN PARAMETERS

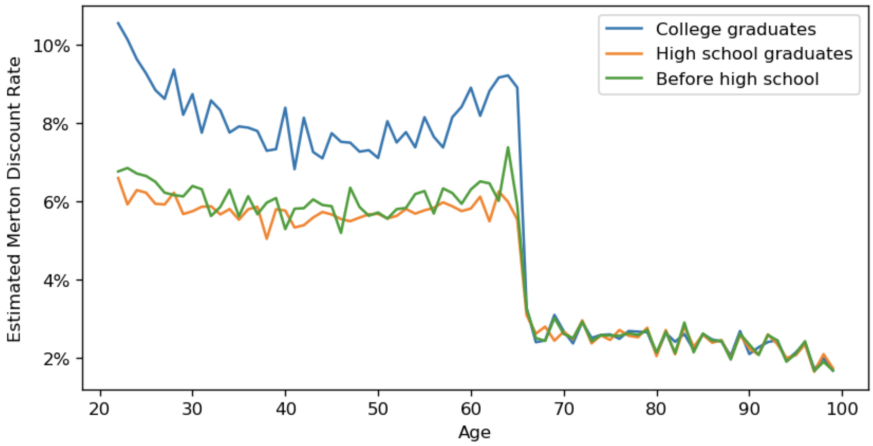


Notes: The figure shows how the optimal portfolio policies vary across different parameter settings. The default parameter set is relative risk aversion being 6, riskfree rate being 2%, risk premium being 4%, and being a college graduate. For each subfigure, we vary one single parameter as indicated on the subtitles.

the effect of varying risk premium is barely visible, which also align with what we see from table 3.

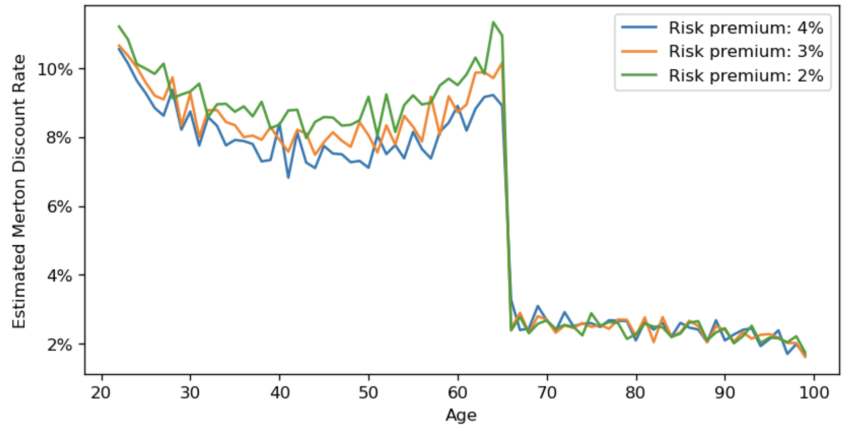
The last appendix figure compares the CL fit, the Merton fit, and the CGM optimal solution across different ages. In the main text, we only see the one-to-one comparison between CL fit and CGM solution, or between the Merton fit and the CGM solution. Here, we are putting them all together. In general, we can see that the CL fit closely tracks the Merton fit across all ages. For post-retirement periods, the three lines are almost overlapping with each other, indicating a great fit of using our approximation. For pre-retirement periods, there are some differences between CL fit and the CGM solution. However, we can see that most of the differences come from the difference between the Merton fit and

Figure A.3. DISCOUNT RATES ESTIMATED FOR DIFFERENT EDUCATION LEVELS



Notes: The figure shows how the calculated discount rates vary across education levels. The default parameter set is relative risk aversion being 6, riskfree rate being 2%, and risk premium being 4%.

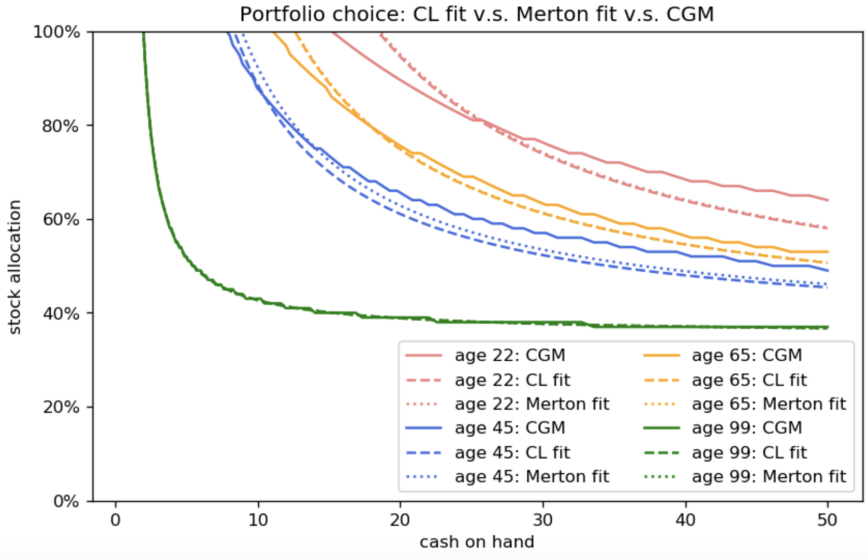
Figure A.4. DISCOUNT RATES ESTIMATED FOR DIFFERENT RISK PREMIUMS



Notes: The figure shows how the calculated discount rates vary across education levels. The default parameter set is relative risk aversion being 6, riskfree rate being 2%, and being a college graduate.

the CGM solution. Our precise prediction of the discount rates make our approximation already very close to the best the Merton formula can achieve. This indicates that if we hope to further improve our approximation accuracy, the main breakthrough would be coming from changing to a different formula than Merton (1969)'s.

Figure A.5. CL APPROXIMATION V.S. MERTON FIT V.S. CGM SOLUTION



Notes: The figure shows how the comparison among CL fit, Merton fit, and the CGM (2005) solutions. The parameter set for the figure is relative risk aversion being 6, riskfree rate being 2%, risk premium being 4%, and being a college graduate.

A.2. Mathematics of Normalization in Policy Function Iterations

In this appendix, I show the mathematical details on how the normalization works in the policy function iteration process.

Recall that we normalize by setting v_{it} to be 1. The normalized income in a pre-retirement period t is then:

$$\widetilde{Y}_{it} = \frac{exp(a_0 + a_1 * t + a_2 * t^2 + a_3 * t^3 + v_{it} + \varepsilon_{it})}{exp(v_{it} - 1)} = exp(a_0 + a_1 * t + a_2 * t^2 + a_3 * t^3 + 1 + \varepsilon_{it})$$

When we normalize the $t+1$ pre-retirement period income by the same normalization unit, we obtain:

$$\begin{aligned}\widetilde{Y}_{it+1} &= \frac{exp(a_0 + a_1 * (t+1) + a_2 * (t+1)^2 + a_3 * (t+1)^3 + v_{it+1} + \varepsilon_{it+1})}{exp(v_{it} - 1)} \\ &= exp(a_0 + a_1 * (t+1) + a_2 * (t+1)^2 + a_3 * (t+1)^3 + 1 + u_{it+1} + \varepsilon_{it+1})\end{aligned}$$

Without normalization, the cash-on-hand at $t+1$ follows:

$$Cash_{it+1} = Saving_{it} * Return_{t+1} + Income_{it+1}$$

With normalization, $Saving_{it}$ is measured in the unit of normalized period t's permanent income. In other words, we have:

$$Cash_{it+1} = Saving\ numeraire_{it} * exp(a_0 + a_1 * t + a_2 * t^2 + a_3 * t^3 + 1) * Return_{t+1} +$$
$$exp(a_0 + a_1 * (t+1) + a_2 * (t+1)^2 + a_3 * (t+1)^3 + 1 + u_{it+1} + \varepsilon_{it+1})$$

Now, we want to set the normalization unit to the permanent income of period $t+1$. Therefore, we should divide both sides by the permanent income of $t+1$:

$$Cash_{it+1}/exp(a_0 + a_1 * (t+1) + a_2 * (t+1)^2 + a_3 * (t+1)^3 + 1 + u_{it+1})$$
$$= Saving\ numeraire_{it} * exp(a_0 + a_1 * t + a_2 * t^2 + a_3 * t^3 + 1)/$$
$$exp(a_0 + a_1 * (t+1) + a_2 * (t+1)^2 + a_3 * (t+1)^3 + 1 + u_{it+1}) * Return_{t+1} + exp(\varepsilon_{it+1})$$

In other words, we have the following:

$$Cash\ numeraire_{it+1} = Saving\ numeraire_{it} * exp(a_0 + a_1 * t + a_2 * t^2 + a_3 * t^3 + 1)/$$
$$exp(a_0 + a_1 * (t+1) + a_2 * (t+1)^2 + a_3 * (t+1)^3 + 1 + u_{it+1})$$
$$* Return_{t+1} + exp(\varepsilon_{it+1})$$

Recall that we have defined the growth rate of the deterministic part of the income between t and $t+1$ as g_{it}:

$$g_{it} = \frac{exp(a_0 + a_1 * (t+1) + a_2 * (t+1)^2 + a_3 * (t+1)^3)}{exp(a_0 + a_1 * t + a_2 * t^2 + a_3 * t^3)} - 1$$

Then the growth rate of permanent income between t and $t+1$ is approximately (from the first-order expansion of exponential function):

$$gp_{it} = exp(g_{it}) * exp(u_{i,t+1})$$

We can therefore get the numerical formula for calculating the cash-on-hand at period $t+1$ using the savings from period t and the income at $t+1$:

$$Cash\ numeraire_{i,t+1} = Saving\ numeraire_{i,t} * Return_{t+1}/gp_{i,t+1} + exp(\varepsilon_{i,t+1})$$

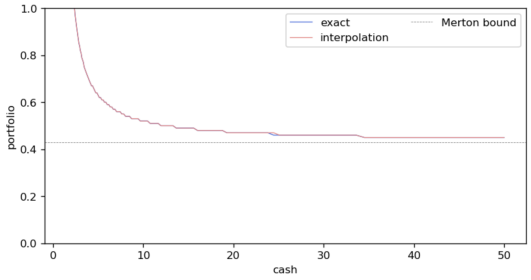
This is the essential formula we use in the coding of the dynamic backward calculation.

For post-retirement periods, the formula is much simpler as the growth rate of the permanent income for the post-retirement periods is always 1. So, in the post-retirement periods, the $gp_{i,t+1}$ component is just 1.

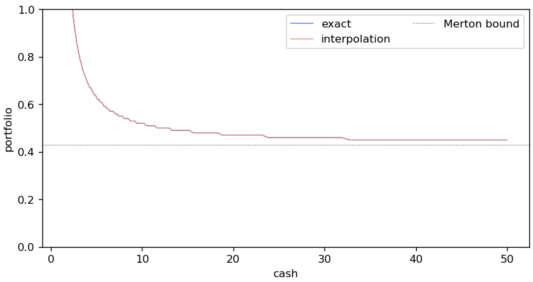
Comparison: Exact Solution v.s. Approximate Solution at Age 99 The figure below compares the age-99 solutions from the CGM (2005) method when we use extrapolation from 196th and 197th grid points and exact value function expression at age 100.

Figure A.6. COMPARISON BETWEEN 196-197 EXTRAPOLATION AND EXACT VALUE FUNCTION

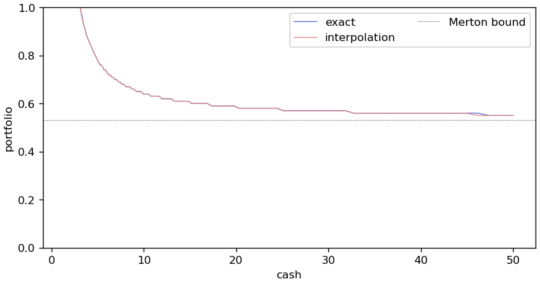
(a) RRA = 4, Riskfree Rate = 2%, Risk Premium = 4%



(b) RRA = 4, Riskfree Rate = 2%, Risk Premium = 3%



(c) RRA = 4, Riskfree Rate = 1%, Risk Premium = 3%

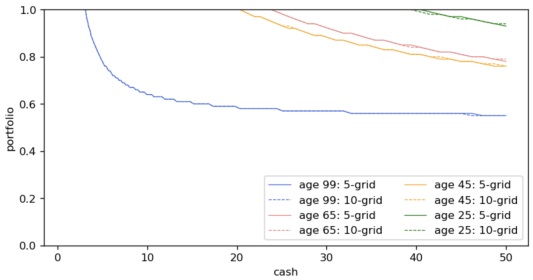


Notes: The figure compares the calculated age-99 portfolio between using exact solution and using 196th-to-197th-grid extrapolation. We assume college graduates for education.

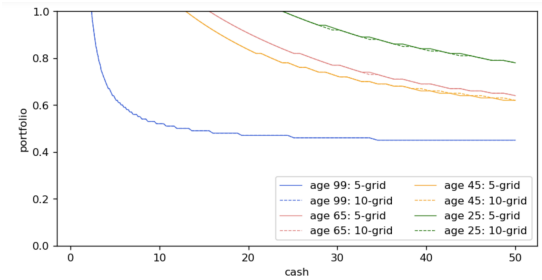
Comparison: Different Cutoff Points for Linear Extrapolation at Higher End The figure below compares the calculated optimal portfolio policies at different ages when we use 196th and 197th grid points to extrapolate and when we use 191st and 192nd grid points to extrapolate.

Figure A.7. COMPARISON BETWEEN 191-192 EXTRAPOLATION AND 196-197 EXTRAPOLATION

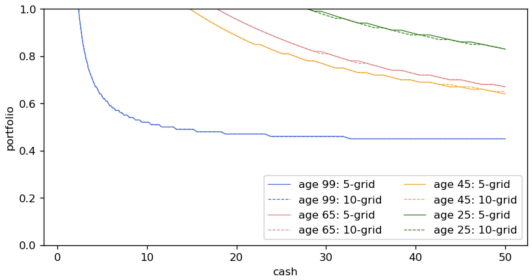
(a) RRA = 4, Riskfree Rate = 2%, Risk Premium = 4%



(b) RRA = 4, Riskfree Rate = 2%, Risk Premium = 3%



(c) RRA = 4, Riskfree Rate = 1%, Risk Premium = 3%



Notes: The figure compares the calculated portfolios between using 196th-to-197th grids to extrapolate and using 191st-to-192nd grids to extrapolate. We assume college graduates for education.

Bibliography

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey Wooldridge (2017), "When Should You Adjust Standard Errors for Clustering?," Technical report, No. w24003, National Bureau of Economic Research. 87, 93

Acharya, Viral V, Sreedhar T Bharath, and Anand Srinivasan (2007), "Does Industry-wide Distress Affect Defaulted Firms? Evidence from Creditor Recoveries," *Journal of Financial Economics*, 85, 787–821. 64, 74, 85

Agüera y Arcas, Blaise, Alexander Todorov, and Margaret Mitchell (2018), "Do Algorithms Reveal Sexual Orientation or Just Expose Our Stereotypes?," *Medium. Available at: https://link.medium.com/GO7FJgFgM1*. 65, 80

Aizer, Anna, Laura Stroud, and Stephen Buka (2016), "Maternal Stress and Child Outcomes: Evidence from Siblings," *Journal of Human Resources*, 51, 523–555. 69

Anderson, Michael and Michael Marmot (2012), "The Effects of Promotions on Heart Disease: Evidence from Whitehall," *The Economic Journal*, 122, 555–589. 62

Angrist, Joshua D and Jörn-Steffen Pischke (2008), Mostly Harmless Econometrics: An Empiricist's Companion Princeton University Press. 127

Antipov, Grigory, Moez Baccouche, Sid-Ahmed Berrani, and Jean-Luc Dugelay (2016), "Apparent Age Estimation from Face Images Combining General and Children-specialized Deep Learning Models," In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 96–104. 65, 76, 77, 109, 110, 113, 114

Atanassov, Julian (2013), "Do Hostile Takeovers Stifle Innovation? Evidence from Anti-takeover Legislation and Corporate Patenting," *The Journal of Finance*, 68, 1097–1131. 66

Babina, Tania (2020), "Destructive Creation at Work: How Financial Distress Spurs Entrepreneurship," *The Review of Financial Studies*, 33, 4061–4101. 64, 74, 85, 109

Baldauf, Markus, Christoph Frei, and Joshua Mollner (2021), "Principal Trading Arrangements: When Are Common Contracts Optimal?," *Management Science*. 7

Bandiera, Oriana, Renata Lemos, Andrea Prat, and Raffaella Sadun (2018), "Managing the Family Firm: Evidence from CEOs at Work," *The Review of Financial Studies*, 31, 1605–1653. 68

Bandiera, Oriana, Andrea Prat, Stephen Hansen, and Raffaella Sadun (2020), "CEO Behavior and Firm Performance," *Journal of Political Economy*, 128, 1325–1369. 63, 68, 131

Barbon, Andrea, Marco Di Maggio, Francesco Franzoni, and Augustin Landier (2019), "Brokers and order flow leakage: Evidence from fire sales," *The Journal of Finance*, 74, 2707–2749. 8

Ben-David, Itzhak, Jiacui Li, Andrea Rossi, and Yang Song (2021), "What do mutual fund investors really care about?," *The Review of Financial Studies*, forthcoming. 15

Bennedsen, Morten, Francisco Perez-Gonzalez, and Daniel Wolfenzon (2020), "Do CEOs Matter? Evidence from Hospitalization Events," *The Journal of Finance*, 75, 1877–1911. 68

Bertrand, Marianne (2004), "From the invisible handshake to the invisible hand? How import competition changes the employment relationship," *Journal of Labor Economics*, 22, 723–765. 69

Bertrand, Marianne and Sendhil Mullainathan (1998), "Corporate Governance and Executive Pay: Evidence from Takeover Legislation," Technical report, No. w6830, National Bureau of Economic Research. 67, 99, 128, 129

Bertrand, Marianne and Sendhil Mullainathan (2001), "Are CEOs Rewarded for Luck? The Ones Without Principals Are," *The Quarterly Journal of Economics*, 116, 901–932. 64

Bertrand, Marianne and Sendhil Mullainathan (2003), "Enjoying the Quiet Life? Corporate Governance and Managerial Preferences," *Journal of Political Economy*, 111, 1043–1075. 64, 66, 71, 124

Bhattacharya, Nilabhra, Ervin L. Black, Theodore E. Christensen, and Richard D. Mergenthaler (2007), "Who trades on pro forma earnings fnformation?," *The Accounting Review*, 82, 581–619. 16

Black, Sandra E, Paul J Devereux, and Kjell G Salvanes (2016), "Does Grief Transfer across Generations? Bereavements during Pregnancy and Child Outcomes," *American Economic Journal: Applied Economics*, 8, 193–223. 69

Boehmer, Ekkehart, Charlies Jones, Xiaoyan Zhang, and Xinran Zhang (2021), "Tracking retail investor activity," *The Journal of Finance*, 76, 2249–2305. 4, 16

Bogousslavsky, Vincent and Dmitriy Muravyev (2021), "Who trades at the close? Implications for price discovery and liquidity," *Available at SSRN*. 2, 7, 9

Borgschulte, Mark and Jacob Vogler (2019), "Run for Your Life? The Effect of Close Elections on the Life Expectancy of Politicians," *Journal of Economic Behavior & Organization*, 167, 18–32, URL http://www.sciencedirect.com/science/article/pii/S016726811930277X. 69

Borkan, Gary A and Arthur H Norris (1980), "Assessment of biological age using a profile of physical parameters," *Journal of Gerontology*, 35, 177–184. 63

Boyce, Christopher J and Andrew J Oswald (2012), "Do People Become Healthier after Being Promoted?," *Health Economics*, 21, 580–596. 62

Brondolo, E., K. Byer, P.J. Gianaros, C. Liu, A.A. Prather, K. Thomas, C.L. Woods-Giscombe, L. Beatty, P. DiSandro, and G.P. Keita (2017), "Stress and Health Disparities: Contexts, Mechanisms, and Interventions among Racial/Ethnic Minority and Low Socioeconomic Status Populations," *American Psychological Association (APA) Working Group Report*. 64

Buti, Sabrina, Barbara Rindi, and Ingrid M Werner (2017), "Dark pool trading strategies, market quality and welfare," *Journal of Financial Economics*, 124, 244–265. 8

Cain, Matthew D, Stephen B McKeon, and Steven Davidoff Solomon (2017), "Do Takeover Laws Matter? Evidence from Five Decades of Hostile Takeovers," *Journal of Financial Economics*, 124, 464–485. 66, 67, 74, 124

Camacho, Adriana (2008), "Stress and Birth Weight: Evidence from Terrorist Attacks," *American Economic Review*, 98, 511–515, URL https://www.aeaweb.org/articles?id=10.1257/aer.98.2.511. 69

Cesarini, David, Erik Lindqvist, Robert Östling, and Björn Wallace (2016), "Wealth, Health, and Child Development: Evidence from Administrative Data on Swedish Lottery Players," *The Quarterly Journal of Economics*, 131, 687–738. 69, 130

Chakravarty, Sugato and Kai Li (2003), "An examination of own account trading by dual traders in futures markets," *Journal of Financial Economics*, 69, 375–397. 8

Chen, Daniel and Darrell Duffie (2021), "Market fragmentation," *American Economic Review*, 111, 2247–74. 8

Cheng, Shijun, Venky Nagar, and Madhav V Rajan (2004), "Identifying Control Motives in Managerial Ownership: Evidence from Antitakeover Legislation," *The Review of Financial Studies*, 18, 637–672. 66, 74, 124

Chetty, Raj, Michael Stepner, Sarah Abraham, Shelby Lin, Benjamin Scuderi, Nicholas Turner, Augustin Bergeron, and David Cutler (2016), "The Association Between Income and Life Expectancy in the United States, 2001-2014," *JAMA*, 315, 1750–1766. 77

Choi, Sekyu, Bing Zhang, Sai Ma, Meryem Gonzalez-Celeiro, Daniel Stein, Xin Jin, Seung Tea Kim, Yuan-Lin Kang, Antoine Besnard, Amelie Rezza, et al. (2021), "Corticosterone inhibits GAS6 to govern hair follicle stem-cell quiescence," *Nature*, 592, 428–432. 76

Christensen, Kaare, Maria Iachina, Helle Rexbye, Cecilia Tomassini, Henrik Frederiksen, Matt McGue, and James W Vaupel (2004), ""Looking old for your age": genetics and mortality," *Epidemiology*, 15, 251–252. 63, 76

Christensen, Kaare, Mikael Thinggaard, Matt McGue, Helle Rexbye, Jacob v B Hjelmborg, Abraham Aviv, David Gunn, Frans van der Ouderaa, and James W Vaupel (2009), "Perceived age as clinically useful biomarker of ageing: cohort study," *BMJ*, 339. 63, 76

Cocco, Joao F, Francisco J Gomes, and Pascal J Maenhout (2005), "Consumption and portfolio choice over the life cycle," *The Review of Financial Studies*, 18, 491–533. 145, 147, 149, 150, 170

Cox, David R (1972), "Regression Models and Life Tables," *Journal of the Royal Statistical Society: Series B*, 34, 187–202. 86, 92, 105, 107, 141, 143

Credit Suisse (2019), "Off-exchange closing volumes rise," *Market commentary report*. 7

Cremers, Martijn and Allen Ferrell (2014), "Thirty Years of Shareholder Rights and Firm Value," *The Journal of Finance*, 69, 1167–1196. 125, 138

Cutler, David, Angus Deaton, and Adriana Lleras-Muney (2006), "The Determinants of Mortality," *Journal of Economic Perspectives*, 20, 97–120. 62

Dannhauser, C. D. and Jeffery Pontiff (2019), "Flow," *Working Paper, Boston College*. 16

Dotsch, Ron, Ran R Hassin, and Alexander Todorov (2016), "Statistical Learning Shapes Face Evaluation," *Nature Human Behaviour*, 1, 1–6. 65, 80

Duffie, Darrell, Wendell Fleming, H Mete Soner, and Thaleia Zariphopoulou (1997), "Hedging in incomplete markets with HARA utility," *Journal of Economic Dynamics and Control*, 21, 753–782. 146

East, Chloe N, Sarah Miller, Marianne Page, and Laura R Wherry (2017), "Multigenerational Impacts of Childhood Access to the Safety Net: Early Life Exposure to Medicaid and the Next Generation's Health," Technical report, No. w23810, National Bureau of Economic Research. 69

Easterbrook, Frank H and Daniel R Fischel (1981), "The Proper Role of a Target's Management in Responding to a Tender Offer," *Harvard Law Review*, 94, 1161–1204. 64

Edmans, Alex and Xavier Gabaix (2011), "The Effect of Risk on the CEO Market," *The Review of Financial Studies*, 24, 2822–2863. 67, 128

Eisfeldt, Andrea L and Camelia M Kuhnen (2013), "CEO Turnover in a Competitive Assignment Framework," *Journal of Financial Economics*, 109, 351–372. 72

Engelberg, Joseph and Christopher A Parsons (2016), "Worrying about the stock market: Evidence from hospital admissions," *The Journal of Finance*, 71, 1227–1250. 69

Epel, Elissa S, Elizabeth H Blackburn, Jue Lin, Firdaus S Dhabhar, Nancy E Adler, Jason D Morrow, and Richard M Cawthon (2004), "Accelerated Telomere Shortening in Response to Life Stress," *Proceedings of the National Academy of Sciences of the United States of America*, 101, 17312–17315. 62

Epstein, Larry G and Stanley E Zin (1991), "Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis," *Journal of political Economy*, 99, 263–286. 148

Ernst, Thomas, Jonathan Sokobin, and Chester Spatt (2021), "The value of off-exchange data," *Working Paper*. 8

Evans, William N and Craig L Garthwaite (2014), "Giving Mom a Break: The Impact of Higher EITC Payments on Maternal Health," *American Economic Journal: Economic Policy*, 6, 258–290. 69

Fama, Eugene F and Kenneth R French (1997), "Industry Costs of Equity," *Journal of Financial Economics*, 43, 153–193. 89

Farrell, Michael, T Clifton Green, Russell Jame, and Stanimir Markov (2021), "The democratization of investment research and the informativeness of retail investor trading," *Journal of Financial Economics*. 4

Fishman, Michael J and Francis A Longstaff (1992), "Dual trading in futures markets," *The Journal of Finance*, 47, 643–671. 7, 35

Fitzpatrick, Maria D and Timothy J Moore (2018), "The Mortality Effects of Retirement: Evidence from Social Security Eligibility at Age 62," *Journal of Public Economics*, 157, 121–137. 131

Franceschi, Claudio, Paolo Garagnani, Paolo Parini, Cristina Giuliani, and Aurelia Santoro (2018), "Inflammaging: A New Immune–Metabolic Viewpoint for Age-related Diseases," *Nature Reviews Endocrinology*, 14, 576–590. 64

Ganster, Daniel C and Christopher C Rosen (2013), "Work Stress and Employee Health: A Multidisciplinary Review," *Journal of Management*, 39, 1085–1122. 62

Garvey, Gerald T and Todd T Milbourn (2006), "Asymmetric Benchmarking in Compensation: Executives are Rewarded for Good Luck But Not Penalized for Bad," *Journal of Financial Economics*, 82, 197–225. 64

General, Surgeon (2014), "The Health Consequences of Smoking – 50 Years of Progress: A Report of the Surgeon General," In US Department of Health and Human Services. 66

Gibbons, Robert and Kevin J Murphy (1992), "Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence," *Journal of Political Economy*, 100, 468–505. 65, 71, 91, 109

Giroud, Xavier and Holger M Mueller (2010), "Does Corporate Governance Matter in Competitive Industries?," *Journal of Financial Economics*, 95, 312–331. 71, 127

Gompers, Paul, Joy Ishii, and Andrew Metrick (2003), "Corporate Governance and Equity Prices," *The Quarterly Journal of Economics*, 118, 107–156. 125

Gompertz, Benjamin (1825), "XXIV. On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies. In a Letter to Francis Baily, Esq. FRS &c," *Philosophical Transactions of the Royal Society of London*, 115, 513–583. 89

Gormley, Todd A and David A Matsa (2016), "Playing it Safe? Managerial Preferences, Risk, and Agency Conflicts," *Journal of Financial Economics*, 122, 431–455. 66, 71, 97, 124, 127

Heidt, Timo, Hendrik B Sager, Gabriel Courties, Partha Dutta, Yoshiko Iwamoto, Alex Zaltsman, Constantin von Zur Muhlen, Christoph Bode, Gregory L Fricchione, John Denninger, et al. (2014), "Chronic Variable Stress Activates Hematopoietic Stem Cells," *Nature Medicine*, 20, 754. 76

Hendershott, Terrence and Haim Mendelson (2000), "Crossing networks and dealer markets: Competition and performance," *The Journal of Finance*, 55, 2071–2115. 8, 24, 25

Hernaes, Erik, Simen Markussen, John Piggott, and Ola L Vestad (2013), "Does Retirement Age Impact Mortality?," *Journal of Health Economics*, 32, 586–598. 131

Hong, Harrison and Jiang Wang (2000), "Trading and returns under periodic market closures," *The Journal of Finance*, 55, 297–354. 7

Hubbard, R Glenn, Jonathan Skinner, and Stephen P Zeldes (1995), "Precautionary saving and social insurance," *Journal of political Economy*, 103, 360–399. 147

Huddart, Steven, John S Hughes, and Carolyn B Levine (2001), "Public disclosure and dissimulation of insider trades," *Econometrica*, 69, 665–681. 7

Huggett, Mark and Greg Kaplan (2016), "How large is the stock component of human capital?," *Review of Economic Dynamics*, 22, 21–51. 147

Human Mortality Database (2019), "USA, 1x1 Cohort Mortality Rates," *University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany)*, URL `mortality.org`. 90, 129

Hummels, David, Jakob Munch, and Chong Xiang (2016), "No Pain, No Gain: the Effects of Exports on Effort, Injury, and Illness," Technical report, No. w22365, National Bureau of Economic Research. 69

Insler, Michael (2014), "The Health Consequences of Retirement," *Journal of Human Resources*, 49, 195–233. 131

Jenter, Dirk and Katharina Lewellen (2015), "CEO Preferences and Acquisitions," *The Journal of Finance*, 70, 2813–2852. 64, 87

Jha, Prabhat, Chinthanie Ramasundarahettige, Victoria Landsman, Brian Rostron, Michael Thun, Robert N Anderson, Tim McAfee, and Richard Peto (2013), "21st-Century Hazards of Smoking and Benefits of Cessation in the United States," *New England Journal of Medicine*, 368, 341–350. 66, 90

Johnston, David W and Wang-Sheng Lee (2013), "Extra Status and Extra Stress: Are Promotions Good for Us?," *ILR Review*, 66, 32–54. 62

Kaniel, Ron, Shuming Liu, Gideon Saar, and Sheridan Titman (2012), "Individual investor trading and return patterns around earnings announcements," *The Journal of Finance*, 67, 639–680. 4

Kaplan, Greg and Sam Schulhofer-Wohl (2018), "The Changing (Dis-)Utility of Work," *Journal of Economic Perspectives*, 32, 239–258. 62

Karpoff, Jonathan M and Michael D Wittry (2018), "Institutional and Legal Context in Natural Experiments: The Case of State Antitakeover Laws," *The Journal of Finance*, 73, 657–714. 67, 98, 124, 125, 126, 138

Kelley, Eric K and Paul C Tetlock (2013), "How wise are crowds? Insights from retail orders and stock returns," *The Journal of Finance*, 68, 1229–1265. 4

Kelley, Eric K and Paul C Tetlock (2017), "Retail short selling and stock prices," *The Review of Financial Studies*, 30, 801–834. 4

Keloharju, Matti, Samuli Knüpfer, and Joacim Tåg (2020), "CEO Health and Corporate Governance," Technical report, No. 1326, Research Institute of Industrial Economics. 68

Kennedy, Brian K, Shelley L Berger, Anne Brunet, Judith Campisi, Ana Maria Cuervo, Elissa S Epel, Claudio Franceschi, Gordon J Lithgow, Richard I Morimoto, Jeffrey E Pessin, et al. (2014), "Geroscience: Linking Aging to Chronic Disease," *Cell*, 159, 709–713. 64

Kim, So Ra, Yu Ri Jung, Hye Jin An, Dae Hyun Kim, Eun Ji Jang, Yeon Ja Choi, Kyoung Mi Moon, Min Hi Park, Chan Hum Park, Ki Wung Chung, et al. (2013), "Anti-Wrinkle and Anti-Inflammatory Effects of Active Garlic Components and the Inhibition of MMPs via NF-κB Signaling," *PloS One*, 8, e73877. 76

Koijen, Ralph S. J. and Stijn Van Nieuwerburgh (2020), "Combining Life and Health Insurance," *Quarterly Journal of Economics*, 135, 913–958. 69

Koo, Hyeng Keun (1998), "Consumption and portfolio selection with labor income: a continuous time approach," *Mathematical Finance*, 8, 49–65. 146

Koo, Hyeng Keun (1999), "Consumption and portfolio selection with labor income: A discrete-time approach," *Mathematical Methods of Operations Research*, 50, 219–243. 146

Kuka, Elira (2020), "Quantifying the Benefits of Social Insurance: Unemployment Insurance and Health," *Review of Economics and Statistics*, 102, 490–505. 69

Kyle, Albert S (1985), "Continuous auctions and insider trading," *Econometrica: Journal of the Econometric Society*, 1315–1335. 3, 23, 30, 36, 37

Lavetti, Kurt (2020), "The Estimation of Compensating Wage Differentials: Lessons from the Deadliest Catch," *Journal of Business & Economic Statistics*, 38, 165–182. 67

Lazarus, Richard S and Susan Folkman (1984), Stress, Appraisal, and Coping Springer Publishing Company. 64

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton (2015), "Deep Learning," *Nature*, 521, 436–444. 111

Lee, Charles M.C. and Balkrishna Radhakrishna (2000), "Inferring investor behavior: Evidence from TORQ data," *Journal of Financial Markets*, 3, 83–111. 16

Malkiel, Burton Gordon (1999), A random walk down Wall Street: including a life-cycle guide to personal investing WW Norton & Company. 147

Marmot, Michael (2005), Status Syndrome: How Your Social Standing Directly Affects Your Health A&C Black. 62

Mas, Alexandre and Amanda Pallais (2017), "Valuing Alternative Work Arrangements," *American Economic Review*, 107, 3722–3759. 67

McEwen, Bruce S (1998), "Protective and Damaging Effects of Stress Mediators," *New England Journal of Medicine*, 338, 171–179. 62

Merton, Robert C (1969), "Lifetime portfolio selection under uncertainty: The continuous-time case," *The review of Economics and Statistics*, 247–257. 145, 146, 156

Merton, Robert C (1975), "Optimum consumption and portfolio rules in a continuous-time model," In Stochastic optimization models in finance, 621–661, Elsevier. 146

Merton, Robert C and Paul Anthony Samuelson (1992), "Continuous-time finance,". 156

Milligan, Kevin and Mark Stabile (2011), "Do child tax benefits affect the well-being of children? Evidence from Canadian child benefit expansions," *American Economic Journal: Economic Policy*, 3, 175–205. 62

Olshansky, S Jay and Bruce A Carnes (1997), "Ever Since Gompertz," *Demography*, 34, 1–15. 89

Opler, Tim C and Sheridan Titman (1994), "Financial Distress and Corporate Performance," *The Journal of Finance*, 49, 1015–1040. 64, 74, 85

Parkhi, Omkar M, Andrea Vedaldi, and Andrew Zisserman (2015), "Deep Face Recognition,". 111

Persson, Petra and Maya Rossin-Slater (2018), "Family Ruptures, Stress, and the Mental Health of the Next Generation," *American Economic Review*, 108, 1214–1252. 69

Pickett, Kate E and Richard G Wilkinson (2015), "Income inequality and health: a causal review," *Social Science & Medicine*, 128, 316–326. 62

Porter, Michael E. and Nitin Nohria (2018), "How CEOs Manage Time," *Harvard Business Review*, 96, 42–51. 63, 68, 131

Puterman, Eli, Alison Gemmill, Deborah Karasek, David Weir, Nancy E Adler, Aric A Prather, and Elissa S Epel (2016), "Lifespan Adversity and Later Adulthood Telomere Length in the Nationally Representative US Health and Retirement Study," *Proceedings of the National Academy of Sciences*, 113, E6335–E6342. 62

Rablen, Matthew D and Andrew J Oswald (2008), "Mortality and Immortality: The Nobel Prize as an Experiment into the Effect of Status upon Longevity," *Journal of Health Economics*, 27, 1462–1471. 69

Röell, Ailsa (1990), "Dual-capacity trading and the quality of the market," *Journal of Financial Intermediation*, 1, 105–124. 7, 31, 35

Ruhm, Christopher J (2000), "Are recessions good for your health?," *The Quarterly Journal of Economics*, 115, 617–650. 62

Samuelson, Paul A (1975), "Lifetime portfolio selection by dynamic stochastic programming," *Stochastic optimization models in finance*, 517–524. 146

Sapolsky, Robert M (2005), "The Influence of Social Hierarchy on Primate Health," *Science*, 308, 648–652. 62

Sarkar, Asani (1995), "Dual trading: Winners, losers, and market impact," *Journal of Financial Intermediation*, 4, 77–93. 3, 4, 7, 23, 35, 36

Scharfstein, David (1988), "The Disciplinary Role of Takeovers," *The Review of Economic Studies*, 55, 185–199. 64

Schwandt, Hannes (2018), "Wealth Shocks and Health Outcomes: Evidence from Stock Market Fluctuations," *American Economic Journal: Applied Economics*, 10, 349–377. 130

SEC DERA (2017), "Bats market close: Off-exchange closing volume and price discovery," *SEC memorandum*. 2, 5, 7, 15

Shanthikumar, Devin (2003), "Small and large trades around earnings announcements: Does trading behavior explain post-earnings-announcement drift?," *Working Paper*. 16

Simonyan, Karen and Andrew Zisserman (2014), "Very Deep Convolutional Networks for Large-scale Image Recognition," *arXiv preprint arXiv:1409.1556*. 111

Smith, James P (1999), "Healthy Bodies and Thick Wallets: the Dual Relation between Health and Economic Status," *Journal of Economic perspectives*, 13, 145–166. 62

Snyder-Mackler, Noah, Joseph Robert Burger, Lauren Gaydosh, Daniel W Belsky, Grace A Noppert, Fernando A Campos, Alessandro Bartolomucci, Yang Claire Yang, Allison E Aiello, Angela O'Rand, et al. (2020), "Social Determinants of Health and Survival in Humans and Other Animals," *Science*, 368. 62

Sullivan, Daniel and Till Von Wachter (2009), "Job Displacement and Mortality: An Analysis Using Administrative Data," *The Quarterly Journal of Economics*, 124, 1265–1306. 69, 90

Tauchen, George and Robert Hussey (1991), "Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models," *Econometrica: Journal of the Econometric Society*, 371–396. 153

Viceira, Luis M (2001), "Optimal portfolio choice for long-horizon investors with nontradable labor income," *The Journal of Finance*, 56, 433–470. 146

Viscusi, W Kip and Joseph E Aldy (2003), "The Value of a Statistical Life: A Critical Review of Market Estimates throughout the World," *Journal of Risk and Uncertainty*, 27, 5–76. 99, 129

Wang, Yilun and Michal Kosinski (2018), "Deep Neural Networks Are More Accurate Than Humans at Detecting Sexual Orientation from Facial Images," *Journal of Personality and Social Psychology*, 114, 246. 65, 80

WSJ (2018), "Goldman cashes in on passive-investing boom with big 4 P.M. trade," *The Wall Street Journal (August 26, 2018)*. 2, 5, 12

Yang, Liyan and Haoxiang Zhu (2020), "Back-running: Seeking and hiding fundamental information in order flows," *The Review of Financial Studies*, 33, 1484–1533. 7

Ye, Mao (2010), "Non-execution and market share of crossing networks," *Available at SSRN 1719016*. 8

Yen, Gili and Lee Benham (1986), "The Best of All Monopoly Profits is a Quiet Life," *Journal of Health Economics*, 5, 347–353. 68

Zhang, Bing, Sai Ma, Inbal Rachmin, Megan He, Pankaj Baral, Sekyu Choi, William A. Goncalves, Yulia Shwartz, Eva M. Fast, Yiqun Su, Leonard I. Zon, Aviv Regev, Jason D. Buenrostro, Thiago M. Cunha, Isaac M. Chiu, David E. Fisher, and Ya-Chieh Hsu (2020), "Hyperactivation of Sympathetic Nerves Drives Depletion of Delanocyte Stem Cells," *Nature*, 577, 676–681. 76

Zhu, Haoxiang (2014), "Do dark pools harm price discovery?," *The Review of Financial Studies*, 27, 747–789. 3, 8, 12, 24, 25, 36