The China-U.S. Equity Valuation Gap

Geert Bekaert Shuojia Ke Xue Wang Xiaoyan Zhang

December 2022

Abstract

The Chinese earnings yield differential relative to the U.S. switches from negative to positive around 2009, with the aggregate variation masking substantial cross-sector variation. Changes in sectoral composition and (changing) growth expectations are not important determinants of the variation in China-U.S. valuation differentials. Instead, changes in ownership structure, and most importantly cross-sectional and temporal variation in financial openness, are the key contributors. In addition, we show that IPOs in the banking sector and its internationalization played a critical role in the (relative) valuation change.

JEL Classification: F36, G15, G18
Keywords: Chinese stock valuation, market integration, financial development, earnings yields, speculative trading, growth expectations, emerging market discount.

* Bekaert is at Columbia Business School and CEPR (email: gb241@gsb.columbia.edu). Ke is at the People's Bank of China Guangzhou Branch (email: keshj.15@pbcsf.tsinghua.edu.cn). Wang is at the School of Finance, Nankai University (email: xuewang@nankai.edu.cn). Zhang is at PBC School of Finance of Tsinghua University (email: zhangxiaoyan@pbcsf.tsinghua.edu.cn). We are grateful for the comments of Kalok Chan, Jennifer Huang, Clark Liu, Christian Lundblad, Chang Ma, Jun Qian, Chenyu Shan and participants of the China Finance Research Conference, the ABFER Annual Conference and the CICF Annual Conference. We thank Pengfei Li and Zhaojing Chen for early research assistance on this project.

The China-U.S. Equity Valuation Gap

Abstract

The Chinese earnings yield differential relative to the U.S. switches from negative to positive around 2009, with the aggregate variation masking substantial cross-sector variation. Changes in sectoral composition and (changing) growth expectations are not important determinants of the variation in China-U.S. valuation differentials. Instead, changes in ownership structure, and most importantly cross-sectional and temporal variation in financial openness, are the key contributors. In addition, we show that IPOs in the banking sector and its internationalization played a critical role in the (relative) valuation change.

JEL Classification: F36, G15, G18
Keywords: Chinese stock valuation, market integration, financial development, earnings yields, speculative trading, growth expectations, emerging market discount.

I. Introduction

China is a unique emerging market, featuring the second largest economy and the second largest equity market in the world. While most emerging markets liberalized their capital markets in the late 80s or early 90s (Bekaert, Harvey and Lundblad, 2003), China liberalized inward and outward capital flows in a controlled, gradual manner, making it even more unique. China only has a tiny B-share market intended for foreign investors, and (limited) foreign investment in the A-share market was only allowed with the launch of the Qualified Foreign Institutional Investor Program in 2002. The past decade has witnessed China's transformative steps to open up further to global investors, for example, through the introduction of the Renminbi Qualified Foreign Institutional Investor Program launched in 2011, and the launch of Shanghai (Shenzhen)-Hong Kong Stock Connect in 2014 (2016). In addition, a large number of A-share firms cross-listed shares in the B-share market, the Hong Kong market (H-shares) or other international markets.

Given that the stock market sets the cost of capital and capitalizes growth opportunities for public firms, understanding valuations in the world's second largest economy is of paramount importance. In this paper, we therefore analyze the valuation differentials between China and the U.S. during China's unusual path towards integration. Because the U.S. market is by far the largest equity market in the developed world, it is a reasonable benchmark for fair valuation.[1] With the popular price-to-earnings (PE) ratio highly skewed, we use the reciprocal of the PE ratio, the earnings yield (EY), which has better distributional properties, as the main valuation measure.

Figure 1 presents some initial evidence by showing the time series of aggregate EYs for China and comparing it with the EYs for emerging markets and the U.S. The EY ratio of emerging markets (the dotted line) is on average 1.95% higher than the EY ratio of the U.S. (the dashed line),

[1] Doidge, Karolyi and Stulz (2020) show that a large valuation gap opened up between the U.S. and other developed markets after the global financial crisis. However, the gap with emerging markets stayed stable.

reflecting the well-known emerging market valuation discount. Intriguingly, before 2009, the China A-share stocks have a low average EY ratio (the solid line), 1.19% lower than that of the U.S., and 4.07% lower than that of the emerging market index. The valuation ratio patterns change after 2009. Chinese EYs quickly and significantly increase and become in line with those of emerging markets, but higher than those of U.S. firms. The aggregate change hides large cross-sector variation in valuation changes, with some sectors witnessing large increases, but others witnessing decreases in EY differentials (see Section IV for more detail). We exploit this cross-sector variation empirically.

We propose three hypotheses to explain (the changes in) the valuation gap between China and the U.S. First, a changing sector decomposition might explain the stylized facts in Figure 1 if the relative importance of high (low) EY industries has increased (decreased) over time. Second, we hypothesize that a gradual liberalization and financial development process effectively increased aggregate discount rates, resulting in the observed EY variation. Specifically, a unique feature of the Chinese market is that local (especially retail) investors traditionally have valued stocks richer than foreign investors (see e.g. the Mei, Scheinkman and Xiong (2009) study on the A-share premium relative to B-shares). Gradual increased foreign ownership may then lower local valuations directly and through spillover effects, as are also observed in local emerging markets in response to ADR listings (Fernandes, 2009). However, in China the international valuation effect is the opposite to what is observed for other emerging markets. Finally, changing growth prospects may explain the variation. The capitalization of strong growth opportunities in China, not present in a mature economy such as the U.S., may explain its lower EY before 2009. After all, in the decade before 2009, China enjoyed double digit annual GDP growth rates, while the U.S., as well as other developed markets, typically feature annual GDP growth rates of around 2%. Perhaps, later in the sample, China's growth prospects diminished and market valuations began to be more reflective of the usual emerging market risks.

Our main tests follow Bekaert, Harvey, Lundblad and Siegel (2011, BHLS henceforth) and focus on explaining earnings yield differentials between Chinese and U.S. stock portfolios. We reject the first hypothesis by showing that the earnings yield differential between China and U.S. is dominated by valuation differentials in the same sectors rather than industry structure.

For the second and third hypotheses, we use a multivariate regression framework with portfolio fixed effects and a plethora of valuation fundamentals, including proxies for financial openness, domestic ownership structure, financial market development, and growth expectations. To maximize our econometric power and exploit cross-sectional variation in valuations, we not only use industry portfolios but also characteristic portfolios, organized across various value-relevant characteristics, most importantly, ownership structure. We find that the above channels jointly explain 42.8% of the total variation in the valuation differentials between China and the U.S. between 1995 and 2018. Of the total explanatory power, the financial openness channel explains more than 70%, but the growth expectation channel accounts for only around 4%. For a shorter 2003-2018 sample, the domestic ownership structure channel becomes more important, accounting for over 35% of the explained variation, while the contribution of financial openness decreases commensurately to 39%, remaining dominant. Growth expectations also become more important, explaining more than 10% of the variation.

Our paper is related to the general literature on how market segmentation affects equity valuation. The lower EY of China relative to the U.S., or "China valuation premium", between 1995 and 2009, violates the predictions of standard market segmentation/integration theory, as in Errunza and Losq (1985), and Bekaert and Harvey (2003). This unique premium pattern has attracted academic interest before. For example, Bailey, Chung and Kang (1999) document that among 11 markets that both issue domestic (like A shares) and foreign securities (like B shares), the Chinese stock market is the only market where domestic prices are higher than foreign prices. Fernald and Rogers (2002) also document the discount received by foreigners relative to the domestic A-share market. Mei, Scheinkman and Xiong (2009) attribute the A-B premium to speculative trading in A-shares, and Chan, Menkveld and Yang (2008) attribute it to information asymmetry. Chan, Menkveld and Yang (2007) show that the price discovery mostly happens in the A-share market, justifying our focus on A-shares. While previous research uses the same companies in the A- versus B-share markets to study Chinese valuation patterns, which represent a very small fraction of the overall Chinese equity market, we focus on the general time-series and cross-sectional valuation gaps between Chinese and U.S. firms. Focusing more on economy-wide forces, our paper connects to the literature examining the factors that affect valuation differentials across (mostly) emerging markets. The previous literature finds factors such as financial openness restrictions and poor stock market development (BHLS, 2011), political risk (Bekaert, 1995; Erb, Harvey and Viskanta, 1996), and poor liquidity (Lesmond, 2005; Bekaert, Harvey and Lundblad, 2007) to effectively segment markets. Finally, there is a rapidly growing literature on the Chinese equity market, studying the cross-section of expected returns (Liu, Stambaugh and Yuan, 2019), the return and performance gap between domestically listed and externally listed firms (Allen, Qian, Shan and Zhu, 2022), the return effects of specific episodes of liberalization, such as the Hong Kong Connect programs (Chan and Kwok, 2018; Liu, Wang and Wei, 2021) and the efficiency and informativeness of the Chinese stock market (Carpenter, Lu and Whitelaw, 2021). We come back to some of these articles when discussing our actual results.

Compared to the extant literature, we provide two contributions. First, our study is one of the first to detail the times-series and cross-sector dynamics of stock market valuation in China, the second largest equity market in the world. Second, our method clearly attributes the valuation gap to the financial openness and domestic ownership structure channels. State (retail) ownership is associated with low (high) valuations, but increased international accessibility is most strongly associated, both in the time series and the cross section, with lower valuations, in contrast to the conventional financial openness effects.

The paper is organized as follows. In Section II, we describe the data and valuation variables, and document the stylized facts on the China-U.S. equity valuation gap. In Section III, we introduce our empirical framework. Section IV examines three hypotheses to explain the time-series and cross-sectional variation in valuation differences. Section V analyzes an alternative model focusing on Chinese earnings yields. Section VI provides further discussion and Section VII concludes.

II. Data and Stylized Facts

We introduce the data in Section II.A, and describe the construction of valuation variables at the portfolio level in Section II.B. We document the stylized facts about the China-U.S. equity valuation gap at both market and portfolio levels in Section II.C.

II.A Data Sources

Our main sample period is from 1995 to 2018. The Chinese stock exchanges were opened in 1990, but with a limited number of listed firms, explaining our sample start deferral to 1995. Given that the valuation data is available at the quarterly frequency, we construct all variables in this study at that frequency.

We obtain Chinese firm level data from Datastream, WIND, CSMAR, and Suntime. Datastream provides Level 4 sector classifications for each firm. From WIND, we collect basic firm level accounting, trading, and institutional ownership data. From CSMAR, we obtain share structure data, such as information on state ownership, and analyst data. Suntime provides additional data on analysts' sales and earnings forecasts. Analysts-based growth expectations and institutional/retail ownership variables are mostly available after 2003, which restricts our sample length when using these variables. Following Liu, Stambaugh, Yuan (2019), we apply the following data filters to the Chinese data: 1) exclude stocks that have become public within the past two quarters; 2) drop stocks that have less than 45 daily return observations during the most recent quarter; and 3) drop stocks that have less than 120 daily return observations during the most recent year. We adopt the first filter (the "IPO filter") for Chinese firms because Chinese IPO pricing (and hence valuation) is heavily influenced by the Chinese Security Regulatory Commission (CSRC), which may not necessarily reflect market consensus. Assuming the IPO's valuation only gradually incorporates market forces, we omit the first two quarters post IPO.[2]

Firm level data for the U.S. are obtained from Datastream, CRSP, Compustat, and I/B/E/S. Again, we obtain Level 4 sector classifications for each firm from Datastream. CRSP and Compustat provide basic firm level information regarding stock trading and accounting variables, and I/B/E/S provides analyst data. We first restrict our sample to common equities with share codes of 10 or 11. Then we adopt filters 2) and 3) described above to the U.S. data. Because the IPO pricing mechanism in the U.S. is market-based, we skip the IPO filter for U.S. firms.

II.B Valuation Variables

Our key valuation ratio is the earnings yield (EY), the reciprocal of the price-to-earnings (PE) ratio. Following BHLS (2011), we prefer the earnings yield to the price-to-earnings ratio for three reasons: the PE is highly positively skewed, while the EY has better distributional properties;

[2] Allen, Qian, Shan and Zhu (2022) argue that the IPO process in China may contribute to the poorer operating performance of A-listed shares relative to externally listed shares.

the PE is not defined when earnings are zero, while the EY is not affected; and the EY is in units that can be easily interpreted in terms of discount rates and expected cash flow growth rates. To capture cross-sectional valuation differences at the firm level in a parsimonious way and reduce firm level noise, we group firms into portfolios. We define the earnings yield for portfolio *j* in quarter *t* as:

$$EY_{j,t} = \frac{\sum_{i=1}^{N_j} Total\ annualized\ net\ income_{i,j,t}}{\sum_{i=1}^{N_j} Price_{i,j,t} \times Number\ of\ common\ equity_{i,j,t}}, \tag{1}$$

where N_j is the number of stocks in portfolio j, $Price_{i,j,t}$ is the price of stock *i* in portfolio *j* at the end of quarter *t*, $Number\ of\ common\ equity_{i,j,t}$ is the latest reported number of common equity shares in a firm's quarterly or annual report, and $Total\ annualized\ net\ income_{i,j,t}$ is the sum of quarterly net income from quarter *t*-4 to quarter *t-1*. Following the previous literature, if $Total\ annualized\ net\ income_{i,j,t}$ is negative, we set it to zero.[3] The aggregate earnings yield at the market level, EY_t, is obtained by replacing N_j in Equation (1) with $N = \sum_{j=1}^{J} N_J$, where J is the number of portfolios, and N is the total number of firms in the market.

To combine firms with similar systematic risk, we first group firms into 38 industry sectors, using Level 4 sector classifications from Datastream. We drop 5 sectors, because China has no listed firms in "Nonlife Insurance", "Tobacco", "Real Estate Investment Trust", "Equity Investment Instrument", and the "Non-Equity Investment Instrument" sectors. Given that valuation ratios can be significantly affected by firm level characteristics, we also construct 21 characteristic portfolios in addition to the 33 industry portfolios. Specifically, we separate firms by state ownership, institutional ownership, retail ownership, international accessibility, liquidity,

[3] The results are robust when we first aggregate firm level annualized net income into a portfolio level annualized net income and then set negative portfolio level annualized net incomes to zero.

and size. For most characteristics, we construct two portfolios, each including firms with the highest or lowest 30% values on the characteristic. With technology related stocks sometimes receiving very high valuations, we construct a "Tech portfolio", which includes firms in the TMT sectors ("Fixed and Mobile Telecom," "Media", "Software and Computer Services" and "Technology, Hardware and Equipment"), and the corresponding non-tech portfolio. Our last set of portfolios differentiates stocks according to where they list. There are three listing boards in China as of 2018 with large differences in listing requirements: the Main board, SME board and the ChiNext board. Firms listed on the main board are usually large companies, while the SME board mainly includes small firms and ChiNext is a board which aims to attract innovative firms with lower listing requirements. More details on the construction of these portfolios are provided in Appendix A.

After we construct the China portfolios, we compute the benchmark valuation ratios for matching U.S. firms. We first compute the matched U.S. portfolios sorted on sector classification, institutional and retail ownership, international accessibility, liquidity, size, and listing boards. When the sorting variable is not readily available to the U.S. firms (which is true for state ownership, international accessibility and listing board), we form the U.S. counterpart as a weighted average of U.S. sector yields to reflect the sectoral composition of the Chinese portfolios. For instance, state ownership is not observable for U.S. firms, so we construct the U.S. counterpart by matching the sector composition of the Chinese state ownership portfolio.

II.C Stylized Facts

The top row of Table 1 presents the time-series average of the number of stocks, market value, market value in percentage of total market value, and EY at the market level for China and the U.S.[4] From 1995 to 2018, the average aggregate EY for China is 4.94% (PE ratio of 25.9), compared to the U.S. aggregate EY at 5.13% (PE ratio of 20.3). The average valuation gap over 24 years might not seem substantial, but Figure 1, plotting aggregate EYs over time for China and the U.S., tells a different story. Before 2009, China's EY is mostly below that of the U.S. The yield difference peaks in 2001Q1, with China's EY less than half the U.S.'s. After 2009, China's EY increases significantly, reaching 8.86% at the end of the sample (2018Q4), which is 3.29% above that of the U.S.

While the time-series variation of these market valuations is quite dramatic, the persistent sign switch of the valuation gap between China and U.S. around 2009 is most notable. We therefore examine whether there is a significant structural break in the EY differential, using the methodology in Bai, Lumsdaine and Stock (1998). The test result suggests the break in the market EY differentials is significant at the 5% significance level (sup-Wald statistic=10.03). The estimated break date is 2009Q3, with a 90% confidence interval of 2007Q2 to 2011Q4. To save space, we relegate a detailed discussion of this test to Appendix B.

Starting from the second row of Table 1, we report summary statistics on the average cross-sectional differences at the sector portfolio level. The sector structure in China and the U.S. is quite different. In China, the sector with the largest market cap is "Banks and Life Insurance", representing 14% of the total market capitalization, with the "Real Estate Investment and Services" sector a distinct second, representing about 8% of the market on average. In the U.S., the sector weights are more balanced. The largest two sectors are "Technology Hardware & Equipment", and "Software and Computing Services," with both representing around 9%. Out of 33 industry sectors, 29 Chinese sectors have lower average EYs than their U.S. counterparts. The remaining 4

[4] The summary statistics for PEs can be found in the Online Appendix Table OA1.

sectors, “Fixed and Mobile Telecom”, “Real Estate Investment & Services”, “Banks & Life Insurance”, and “Mining”, have higher EYs than their U.S. counterparts by 0.03%, 0.55%, 0.67%, and 1.63%, respectively.

We report the time-series average of EYs for characteristic portfolios in Online Appendix Table OA2. We indeed observe significant differences in EYs across characteristic portfolios. For example, in China, higher state ownership (SO) is associated with higher EYs and lower valuations, with a non-negligible difference of 2.46% between the highest and lowest SO groups, possibly because state-owned firms are less profitable. Firms with international accessibility have higher EYs by on average 2.26% than firms not accessible to global investors, which is in sharp contrast to the conventional wisdom that foreign investors increase stock valuations.

III. Valuation Framework

The classic valuation model, the Gordon model, holds that with constant expected cash flow growth rates and discount rates, and full payout of earnings, the earnings yield reflects the difference between the discount rate and the expected cash flow growth rate. When discount rates and expected cash flows are time-varying, valuation remains tractable when these variables follow AR(1) processes with Gaussian innovations (see e.g. Bekaert and Harvey, 2000; BHLS, 2011). We derive a simple dynamic model under these assumptions, with derivation details in Appendix C, giving rise to the following price-to-earnings (PE) ratio expression:

$$PE_{j,t}^{c} = \sum_{k=1}^{\infty} \exp\left(a_{j,k}^{c} + b_{j,k}^{c}\delta_{j,t}^{c} + g_{j,k}^{c}cf_{j,t}^{c}\right), \tag{2}$$

where c indexes the country, j the portfolio, $\delta_{j,t}^{c}$ is the discount rate for portfolio j in country c, $cf_{j,t}^{c}$ represents expected cash flows for portfolio j in country c, thus $\exp\left(a_{j,k}^{c} + b_{j,k}^{c}\delta_{j,t}^{c} + g_{j,k}^{c}cf_{j,t}^{c}\right)$ in the infinite sum reflects the current value of the dividend k periods in the future divided by the current dividend. We obtain an expression for the earnings yield (EY) ratio, by taking the reciprocal and linearizing Equation (2):

$$EY_{j,t}^{c} \approx \tilde{a}_j^c + \tilde{b}_j^c \delta_{j,t}^c + \tilde{g}_j^c cf_{j,t}^c \,. \tag{3}$$

To save space, we provide the definitions of $\tilde{a}_j^c, \tilde{b}_j^c\ and\ \tilde{g}_j^c$ in Appendix C.

To make this valuation equation operational for empirical estimation, we make two different sets of simplifying assumptions. For the first approach, we follow BHLS (2011), who assume economic and financial integration, where systematic risk is driven by CAPM betas that are industry but not country specific, and persistent growth opportunities are also global in nature. In that case, the time variation in discount rates and expected cash flows is entirely driven by global variables and the b_j^c and g_j^c coefficients are independent of country, indexed by *c*. Under economic and financial integration, BHLS show that the earnings yield differentials between comparable portfolios from China and the U.S. should be close to zero. Empirically, the explanatory variables must measure deviations of Chinese cash flow and discount rate variables from the global benchmark. In principle, the explanatory variables should also accommodate portfolio specific betas with respect to global discount rates, and we test for this separately.

The second approach completely focuses on Chinese EYs, without considering earnings yields from the U.S. Formally, we assume country-specific structural parameters and linearize around country averages, which renders the b_j^c and g_j^c coefficients independent of specific portfolios indexed by *j*. Under the second approach, variation in the Chinese EY is explained by proxies for portfolio specific cash flow expectations and discount rates. Our main assumption for discount rates is that they depend linearly on the (time-varying) ownership structure for the portfolio, but we also accommodate cross-portfolio and time-series liquidity differences.

The resulting empirical specification is a portfolio-quarter panel model:

$$Y_{j,t} = ak + bk'\overline{X_{j,t}} + ck'\overline{Control_{j,t}} + \mu_j + ek_{j,t}. \quad (4)$$

where $k = 1, 2$. For the first specification (the "Valuation Difference Model"), the dependent variable, $Y_{j,t} = DIFEY_{j,t} = EY^{CN}_{j,t} - EY^{US}_{j,t}$, is the EY differential between China and the U.S. for portfolio j in quarter t; the independent variables, $\overline{X_{j,t}} = DIFX_{j,t} = X^{CN}_{j,t} - X^{US}_{j,t}$, represent the cross-country differences in explanatory variables.[5] For the second model (the "China EY Model"), we have $Y_{j,t} = EY^{CN}_{j,t}$, and $\overline{X_{j,t}} = X^{CN}_{j,t}$, the latter representing variables capturing time-series and cross-sectional variation in growth prospects and discount rates. We defer a detailed explanation of all explanatory variables to Section IV.

The control variables, $\overline{Control_{j,t}}$, are second-order to the analysis; they reflect variation outside the valuation framework, or cross-sectional or time-series variation of earnings yields not directly reflecting discount rate or cash flow variation**.** For the Valuation Difference Model, we follow BHLS and include the leverage differential, earnings growth volatility differential, and the minimum number of firms (see Appendix D for exact definitions), as controls. Higher leverage means higher financial risk and thus higher discount rates, whatever the model of systematic risk. The earnings growth volatility differential partially determines $\tilde{a}^c_j$ in Equation (3). A portfolio with higher (idiosyncratic) earnings volatility may, all else equal, be more valuable than a portfolio with less volatility (see also Pástor and Veronesi, 2006). Finally, we control for the number of firms, which potentially affects the accuracy of the portfolio level measure. We include the minimum number of firms between the two portfolios in the computation as our third control variable. For the "China EY Model", we use leverage and earnings growth volatility as controls.

[5] For variables that are only available for the Chinese market, we assume their U.S. counterpart values are time invariant, which will then affect the intercept, but not the slope coefficients.

The parameters μ_j represent portfolio fixed effects to take into account unobserved time-invariant factors that are specific to each portfolio. We cannot possibly accommodate all differences between Chinese and U.S. industries, and including portfolio fixed effects allows the regression to focus on the important temporal and cross-sectional changes we document in Figures 1 and 2. We double cluster the standard errors by portfolio and quarter, as in Petersen (2009) and Thompson (2011).

The Valuation Difference Model has the advantage of generating direct predictions for the valuation gap, and is the main specification we consider in Section IV. The "China EY Model" is more flexible and is discussed together with a few extensions in Section V.

Since there is strong evidence of a structural break, we also modify our empirical specification by adding a break dummy. For instance, for the Valuation Difference Model, we estimate:

$$DIFEY_{j,t} = a3 + \gamma Break_t + b3'DIFX_{j,t} + c3'Control_{j,t} + \mu_j + e3_{j,t}. \tag{5}$$

The break dummy, $Break_t$, is set to be one after 2009Q3, and zero otherwise. The break dummy is important in testing our various hypotheses in later sections, because a successful model that can fully explain the time variation in the earnings yield differential, should also account for the break, and render the break dummy coefficient insignificant.

IV. Explaining the Valuation Gap: The Valuation Difference Model

In this section, we consider three hypotheses to explain the time-series and cross-sectional variation in valuation differences between China and U.S., relying on the valuation difference model with various explanatory variables. We start with the industry structure hypothesis in Section IV.A. Section IV.B. examines discount rate factors, focusing on financial openness, domestic ownership structure, and market development. In Section IV.C., we investigate the role of growth prospects. In Section IV.D, we compare the explanatory power of all variables.

IV.A Hypothesis I: Changes in Industry Structure

It is conceivable that the sign switch of the China-U.S. EY differential is driven by an increase in the market shares of high EY industries in the Chinese stock market. In this section, we focus on the industry structure, and thus only consider the 33 industry portfolios, not the 21 characteristic portfolios.

Figure 2 shows the EY differentials between Chinese and U.S. industry portfolios both in the first 5 years (X-axis) and in the last 5 years (Y-axis) of the sample. The figure shows large cross-industry variation in EY differentials but also very varied time-series variation across portfolios. Perhaps surprisingly, there are still many sectors below the horizontal line, indicating lower EYs than their U.S. counterparts. However, 20 out of 33 industries are above the diagonal line, indicating that the increase in EY differentials is not driven by a few industries but prevalent for the majority of industries. Among them, the "Banks & Life Insurance" portfolio experiences the largest increase in EY differential, from a negative -2.5% in 1995-1999 to a large positive 7.1% in 2014-2018. We discuss the role of the banking industry in detail in Section VI.

To formally examine the role of industry structure for the valuation gap, we consider the following decomposition of the EY differential between China and the U.S.:

$$\begin{aligned} DIFEY_t = EY_t^{CN} - EY_t^{US} &= \sum_{j=1}^{33} w_{j,t}^{CN}\left(EY_{j,t}^{CN} - EY_{j,t}^{US}\right) + \sum_{j=1}^{33}\left(w_{j,t}^{CN} - w_{j,t}^{US}\right)EY_{j,t}^{US} \\ &= DIF\_VAL_t + DIF\_STR_t. \end{aligned} \tag{6}$$

Here $w_{j,t}^{CN}$ and $w_{j,t}^{US}$ are the weights of industry j in terms of market capitalization in China and the U.S. respectively. The first component, DIF_VAL, represents the EY differential within the same sector between China and the U.S., and thus it constitutes a pure valuation differential. The second component, DIF_STR, captures sectoral weight differences between China and the U.S. and represents the valuation effect of a different industry structure. This decomposition exercise is conducted each quarter.

As shown in the first row of Table 1, the time-series average of the EY at the market level for China (the U.S.) is 4.94% (5.13%). Following Equation (6), we further decompose the EY differential of -0.19% into DIF_VAL and DIF_STR, and present the decomposition results in Table 2, Panel A. We find that the first component has a time-series average of -0.69% and the second an average of 0.50%. That is, the portfolio composition component, DIF_STR, partially mitigates the negative pure valuation differential reflected in DIF_VAL. We then compute how much each component contributes to the total variance of the overall differential. The valuation component, DIF_VAL, accounts for 99% of the variation of total EY differentials, while the structure component, DIF_STR, contributes only 1%. Thus, variation in the valuation gap between China and the U.S. is dominated by valuation changes within sectors rather than changes in sector structure.

Moreover, relying on the decomposition, we further investigate which component drives the previously recognized break. We perform the break point test for both the valuation and the sector composition components and report the results in Table 2, Panel B. When focusing on the valuation component, DIF_VAL, the break is significant at the 5% level with a break date of 2009Q2. However, when focusing on the composition component, DIF_STR, we fail to detect a structural break. We conclude that time variation in the China-U.S. valuation gap is mostly driven by valuation changes in the same sectors rather than by changes in industry structure.

IV.B Hypothesis II: Changes in Discount Rates

The valuation model in Section III suggests discount rate changes may account for the variation in earnings yield. In this section, we examine the explanatory power of various discount rate variables and highlight three sets of variables related to financial openness, domestic ownership structure, and market development.

Financial Openness

We postulate that compared to Chinese domestic investors, foreign investors tend to require higher discount rates for Chinese stocks. The difference in the required discount rates between domestic and foreign investors can arise from the lack of investment alternatives for Chinese investors, especially retail investors, forcing them to effectively accept lower discount rates for domestic stocks than less constrained foreign investors (e.g. Fernald and Rogers, 2002).[6] Mei, Scheinkman and Xiong (2009) attribute higher A-share stock prices to speculative components induced by retail investors. Foreign investors may also require higher discount rates because of their information disadvantage regarding Chinese stocks compared to domestic investors (see e.g. Chan, Menkveld and Yang, 2008). The evidence on the valuation of cross-listed shares and the B-share market is largely consistent with these hypotheses.

The past decades have witnessed the gradual liberalization process of the Chinese stock market. Therefore, a switch from purely domestic to foreign investors in the A-share market (with potential spillover effects) may cause valuations to drop and earnings yields to increase. Thus, the gradual liberalization process and the cross-sector differences in this process can potentially explain the dynamics of the valuation gap.

The relative ownership of foreign investors is clearly affected by regulations on financial openness and international accessibility. Our first financial openness variable therefore is REGOPEN, a discrete variable measuring China's regulatory process towards more financial

[6] Modeling heterogeneity in beliefs and views through discount rate differences across different investors, is consistent with, for example, the demand-based asset pricing in Koijen and Yogo (2019) or the different beliefs of fund managers in Shumway, Szefler and Yuan (2009).

openness. The REGOPEN variable is set to zero at the beginning of the sample, and increases with the announcement and implementation of major regulations on financial openness. It ranges between 0 and 7, and we list the major Chinese regulatory changes in Appendix E.[7] The REGOPEN variable is intended to capture gradual market-wide change in discount rates brought about by increased foreign ownership but it does not differentiate across Chinese firms with different degrees of accessibility to foreign investors.

To measure international accessibility in the cross section in each quarter, we further construct three international accessibility (IA) variables using firm level information. The first variable, IA1, is a discrete variable, adding up four firm level dummy variables (the presence of B shares, H shares, an ADR, and membership of the Mainland-Hong Kong Stock Connect). The second variable, IA2, is the ratio of the market capitalization of B shares, H shares, and ADRs to the firm's total market capitalization. To construct portfolio level IA1/IA2, we value weight firm level IA1s/IA2s within the portfolio, using the firm's last quarter market capitalization as weight. The third variable, IA3, measures the market share of firms with positive firm level IA1 within the portfolio, which is particularly relevant if there are strong sectoral spillover effects in terms of international pricing. These three variables, not surprisingly, are highly correlated, showing an average correlation coefficient of 0.78.

We plot the time-series of market level international accessibility measures at the market

[7] REGOPEN is set to one after B-shares become investable for Chinese investors in 2000Q1. For the next six events, which include allowing "Qualified Foreign Institutional Investors" to invest in A-shares, allowing "Qualified Domestic Institutional Investors" to invest in foreign markets, allowing "Renminbi Qualified Institutional Investors" to invest in A-shares, setting up the Shanghai-Hong Kong Connect, setting up the Shenzhen-Hong Kong Connect and incorporating A-shares into the MSCI Index, we add one to the REGOPEN variable after each event. For those events that are announced but implemented in different quarters, we separately incorporate announcement and actual implementation effects when we define the REGOPEN variable. For instance, the REGOPEN variable increases by 0.5 when the QFII (QDII) is announced and another 0.5 when the QFII (QDII) is implemented. We give higher weight (0.67) to the announcement effects of Shanghai-Hong Kong Connect, Shenzhen-Hong Kong Connect and MSCI incorporation because it is likely that these events had more impactful announcement effects (see Liu, Wang and Wei, 2021), and we assign a 0.33 weight to the implementation date of the above three events.

level in Figure 3, Panel A. They generally increase over time, except for the early sample where many domestic firms went public. Different from our IA2 measure, IA1 and IA3 also take the Shanghai-Hong Kong and Shenzhen-Hong Kong Stock Connect programs into account. In 2014Q4 and 2016Q3, the launching of "Connects" greatly increased the international accessibility of the A share market. According to our market level IA3 measure, by the end of 2018, the market share of A share stocks accessible to international investors through at least one channel (cross-listing or "connects"), reached 83.6%. This also implies that after 2014, most portfolios in our analysis have some level of international accessibility, with the most open portfolios having 100% international accessibility.

Apart from these financial openness variables, we consider two less direct variables. Following Frankel (1992) and BHLS (2011), we employ the "real interest differential", calculated as the difference between the real interest rates in China and the U.S. The real interest rate is one component of the discount rate, thus a higher real interest rate should be associated with higher earnings yields. BHLS (2011) also suggest that high levels of political risk can effectively segment an emerging market from international investment. To measure political stability, we obtain the "overall rating" from ICRG, and calculate the ratio of the Chinese value over the U.S. value. A higher number represents less political risk (higher stability), and should be associated with higher valuations and lower earnings yields.

In column (1) of Table 3, we report the estimation results for the Valuation Difference Model using the financial openness measures. The REGOPEN measure receives a positive and statistically significant coefficient, consistent with financial openness being positively associated with the EY differential between Chinese stock portfolios and their U.S. counterparts. Among the portfolio-specific international accessibility variables, only IA2 receives a statistically significant positive coefficient. In other emerging markets, this would constitute a surprising result as being priced internationally results typically in lower discount rates and lower earnings yields (higher valuations). The coefficient on the real interest rate differential is statistically insignificant. The political rating measure has a negative coefficient that is statistically significant at the 10% level, consistent with previous literature. The adjusted R^2 of the regression reaches 39.2%.

When we further include the break dummy in column (2), REGOPEN and IA2 are the only variables that maintain statistical significance. The coefficient on the break dummy is economically very small in magnitude and not statistically significant. Thus, variation in financial openness helps remove the Chinese-U.S. valuation gap, by driving down, not up as in other emerging markets, Chinese equity market valuations.

Domestic Ownership Structure

One important difference between the Chinese and U.S. stock markets is the role of the Chinese government as a shareholder, especially before the Split-Share Structure Reform in 2005. State-owned enterprises are often deemed less efficient than privately-owned ones because they may pursue political objectives and have weaker corporate governance (see e.g. Shleifer and Vishny, 1997; Megginson, Nash and Van Randenborgh, 1994). Alternatively, the government may serve as an implicit guarantor and provide protection in financial distress under certain circumstances (e.g. Faccio, Masulis, and McConnell, 2006). Such features of state ownership can be reflected in discount rates (e.g. Ben-Nasr, Boubakri, and Cosset, 2012; Boubakri, El Ghoul, Guedhami and Megginson, 2018) and thus affect earnings yields.

Moreover, institutional ownership shows substantial time-series and cross-portfolio variation in China.[8] Institutional ownership is typically associated with higher valuations and thus lower earnings yields. Finally, given few outside investment options, the speculative behavior of domestic retail investors may potentially lead to unrealistically high market valuations and low earnings yields (see e.g. Mei, Scheinkman and Xiong, 2009).

We use the available data to construct proximate measures of domestic state, institutional and retail ownership. We measure state ownership as the fraction of total shares that are owned by the state, while institutional (retail) ownership is measured as the fraction of tradable shares that are owned by institutions (retail) investors. In China, the number of total and the number of tradable shares can differ greatly because of the split-share structure. We measure institutional (retail) ownership as a faction of tradable shares because these investors generally do not have access to non-tradable shares and they are more likely to affect firm valuations through their trading instead of active monitoring (see Jiang and Kim, 2015).

We estimate state ownership from information on the ten largest shareholders from CSMAR and more precise information regarding state-owned shares among the non-tradable shares.[9] For institutional ownership, we use institutional holdings data from WIND regarding the firms' top ten holders of tradable shares and more precise information for some categories of institutional investors (including mutual funds). Finally, retail investor ownership is defined as (1 - institutional ownership - tradable state ownership - insider ownership). Insiders are defined as

[8] Institutional ownership here includes QFII holdings, but the share of such holdings is negligible due to the limited QFII quota allowed by the Chinese government. Therefore, the institutional and retail ownership introduced in this section can proxy for the "domestic" part of the ownership structure which is different from the previous subsection focusing on the "foreign" part.

[9] Firms listed in the A-share market report their number of non-tradable shares by categories in their financial statements in detail. One of the categories is non-tradable shares owned by the state. It is a precise measure, but only accounts for non-tradable state-owned shares. The top ten shareholders data shows the ten largest shareholders for these companies, taking both tradable and non-tradable shares into account. We can use it to calculate firms' state ownership, but it is not as precise because it only takes the holdings of the top ten largest shareholders into account. To incorporate the information in both variables, we take the larger value of these two variables as our state ownership proxy (see Appendix D for variable construction details).

directors, supervisors or managers in a company, or large individual shareholders who show up in the firms' ten largest shareholders profile, with data obtained from CSMAR. Because institutional and retail ownership are measured in terms of the percentage of tradable shares, the subtracted state ownership here is the fraction of tradable shares that are owned by the state (i.e. tradable state-owned shares over total tradable shares), which is different from the previous state ownership measure that we use as an explanatory variable. Given our incomplete picture of institutional holdings, our institutional ownership measure for China is a lower bound to the true estimate, making our retail ownership an upper bound to true retail ownership.[10]

As an alternative proxy for retail ownership in China, we also compute the "Standardized Number of Shareholders (SNS)", which is the number of shareholders divided by the number of total tradable shares, multiplied by 1000. The data are obtained from WIND. The rationale here is that retail ownership should be positively correlated with the number of shareholders, especially in the A share market where individual investors own 99.78% of stock trading accounts according to the 2019 Shanghai Stock Exchange Statistical Yearbook.

For the U.S., state ownership data is not available, and the government typically does not own private equity. For institutional ownership, we follow Ferreira and Matos (2008), and use Factset LionShares data to calculate institutional holdings. Insider owner's information for U.S. firms is from Thomson Reuters. U.S. retail investor ownership is then defined as (1 - institutional ownership - insider ownership).

All the domestic ownership variables are only available from 2003 onwards except for state

[10] To partially verify the impact of data issues, we compare our institutional and retail ownership data with precise numbers available for the Shanghai Stock Exchange. We find that the correlation between the exchange's measures and our proxy measures is quite high, being 97.7% for the institutional ownership series, and 93.9% for the retail ownership series. Our CSMAR data under-estimate institutional ownership on average by 4.3% and over-estimate retail ownership on average by 8.3%. As long as the bias does not show strong cross-sectional or temporal variation, our panel regressions should still provide useful information.

ownership, which is available from 1995. We present the market average ownership variables over time in Figure 3, Panel B. Average state ownership in China is 45.6%, and this share has not changed much over time. While institutional ownership in the U.S. is on average 80.3%, in China it is on average only around 15.7%. Summary statistics for these variables can be found in the Online Appendix Table OA3. Chinese institutional ownership increases sharply over time up to about 2008, but decreases afterwards. The decrease may be caused by higher state ownership in tradable shares after the Split-Share Structure Reform. Retail ownership appears to decrease over time.

Columns (3)-(6) of Table 3 present the panel regression results for the ownership variables. Column (3) shows a positive but insignificant relationship between EY differentials and state ownership over the full sample period. In column (4), we add the break dummy. The positive impact of state ownership on EY differentials becomes stronger, with the coefficient being 0.056 and statistically significant at the 1% level, consistent with Ben-Nasr, Boubakri, and Cosset (2012). The break dummy coefficient is statistically significant, which suggests that state ownership alone cannot explain away the difference in EY over the full sample period. In columns (5)-(6), we use all the ownership variables jointly for the shorter sample period of 2003-2018. All ownership variables are significantly different from zero except retail ownership. Jointly, these variables explain 45% of the variation in China-U.S. EY differentials. When we add the break dummy in column (6), it is no longer statistically significant and much smaller in magnitude, suggesting that variation in ownership structure helps explain the valuation gap change. Importantly, both higher institutional and retail ownership (the latter proxied by the SNS variable), are associated with higher valuations.

Market Development and Liquidity

We also consider variables that measure general stock market development in China and the U.S. because they can affect valuation multiples through improved allocative efficiency (see e.g. Wurgler, 2000), relaxing financial constraints (see e.g. Love, 2003) and/or improving market liquidity (see BHLS, 2011). To measure market development at the regulatory level, we create a discrete variable, REGDEV, which captures the stage-by-stage market modernization process and ranges between 0 and 2.5.[11] As for direct development indicators, we use the log of the number of public firms and a modified market capitalization to GDP indicator. While market capitalization to GDP is often used as a development measure, its numerator is affected by stock market fluctuations which obviously also affect our dependent variable. We therefore create a relative development indicator that controls for recent stock market returns. Specifically, we first calculate the ratio of Chinese market capitalization to GDP over U.S. market capitalization to GDP. We also calculate the ratio of the one year past cumulative market return in China over the one-year cumulative market return in the U.S. We then calculate the z-scores of these two variables by subtracting their means and dividing by their standard deviations. Our "Adjusted market development" measure is the difference of the z-scores of these two variables at each point in time.

We also consider two direct (il)liquidity indicators, Zeros (the proportion of daily zero returns per quarter, see Lesmond, Ogden and Trzcinka (1999)), and Turnover (quarterly value traded over market capitalization). Liquidity has been shown to capture cross-sectional variation in discount rates in both the U.S. and China (see Amihud, 2002; Amihud, Mendelson and Pedersen,

[11] It starts from zero and is set to 1 after the Split-Share Structure Reform in 2005Q1. It then increases by 0.5 with the following three events, the announcement of the Margin Trading and Short-Selling Program, its official start, and the start of a registration-based IPO system. We choose to assign the Split-Share Structure Reform a value of 1, and the other three regulatory changes all 0.5, because the former is widely considered more impactful, see Liao, Liu and Wang (2014).

2006; Carpenter, Lu and Whitelaw, 2021) and the latter indicator has also been used in the development literature as a stock market development indictor (see e.g. Levine and Zervos, 1998). We plot the market level turnover rate in Figure 3, Panel C, and we observe that the Chinese stock market turnover rate is quite high, especially during stock market booms. The average quarterly market turnover rate is 1.00, but during market booms it can easily exceed 2.00. In contrast, the average quarterly market turnover rate in U.S. is only 0.47.[12]

Columns (7) and (8) of Table 3 reports the estimation results of Equations (3) and (5), using the market development variables. All the variables are statistically significant. REGDEV has a positive coefficient, suggesting that better development leads to lower valuation and higher EY differentials. While this may, at first glance, be counterintuitive, a key reform in this regard was the 2010 permission for short-selling, which could contribute to suppressing excessive valuations. The number of public firms and the adjusted market development measure are associated with lower EY differentials, which is in line with expectations if these variables measure improved liquidity and efficiency. The coefficient on the (il)liquidity differential has the expected sign, consistent with the positive relationship between illiquidity and discount rates. We also find a statistically significant negative coefficient on turnover. While this may be consistent with higher turnover being associated with higher liquidity, which, in turn, increases valuations, Mei, Scheinkman and Xiong (2009) suggest that turnover is an indirect proxy for retail ownership, so that this effect may well reflect variation in retail ownership. The multivariate regression in column

[12] We report summary statistics of these development and liquidity variables in Online Appendix Table OA4, and they are overall consistent with China being an emerging and the U.S. a developed market. The Chinese stock market also has a lower number of public firms and a smaller stock market size in relation to the size of its economy. In terms of Zeros, the Chinese stock market is significantly less liquid than the U.S., but this is partly due to some Chinese A-share listed firms frequently suspending their trading because of a merger or acquisitions, seasoned equity offerings etc.

(7) features an R^2 of 35.7%.[13] When we add the break dummy in column (8), it remains highly statistically significant, suggesting that changing development and liquidity conditions did not play a meaningful role in explaining time-series variation in the China-U.S. valuation gap.

IV.C Hypothesis III: Changes in Growth Prospects

As shown in Equation (2), changes in China's expected growth differential relative to the U.S. may explain the valuation dynamics we document. We consider one market level and two portfolio level growth prospect measures: the GDP growth rate, expected sales growth, and expected earnings (net income) growth.

We collect GDP data and firm level analyst data from the China National Bureau of Statistics, Bureau of Economic Analysis, CSMAR, Suntime, and I/B/E/S. For the two portfolio level measures, we first aggregate analysts' firm level median sales and earnings forecast estimates into portfolio level measures, and then calculate the portfolio level expected annualized sales growth and earnings growth over a three-year horizon (see Appendix D for construction details). Notice that the analyst data is only available after 2003; thus, the statistics and estimation results involving these measures all start from 2003. Figure 3, Panel D shows time-series of GDP growth rates and market level earnings/sales growth expectations. All three measures show that China's growth prospects are slowing down over time, especially in the short sample of 2003-2018. As expected, on average, China's GDP growth is 8.8%, much higher than average growth in the U.S., which is 2.5% over the sample period.

Table 4 reports the estimation results for the Valuation Difference Model. The first two regressions include only GDP growth, and pertain to the sample of 1995 to 2018. In column (1), a 1% increase in the GDP growth differential, significantly (at the 1% level) decreases the average

[13] In a robustness check, we split REGDEV into two separate variables, one focusing on regulation regarding short-selling, the other tracking the other regulatory changes. Both coefficients are positive and statistically significant.

EY differential by 18.4 basis points. The adjusted R^2 is 25.0%. In column (2), we add the break dummy, and the break dummy coefficient is statistically significant. In columns (3) to (4), we further include analyst expectations for sales and earnings growth, with the sample starting in 2003. The coefficient on the GDP growth differential is more negative than in column (1) and remains highly statistically significant. In column (3), both earnings and sales growth expectations show negative coefficients, but only the sales growth coefficient is significant (at the 1% level). Thus, the increase in EY differentials between China and U.S. portfolios is significantly related to a decrease in expected sales growth differentials. Column (4) shows that the break dummy is substantially reduced in value but remains statistically significant. The adjusted R^2 increases to 43.2%.[14]

We expect that the explanatory power of the analyst forecasts may depend on their quality, in the sense that forecast quality should increase with the number of analysts and decrease with forecast dispersion, and thus the dependence of earnings yields on earnings growth expectations should increase in absolute magnitude with the number of analysts and decrease with forecast dispersion. Therefore, we include these two variables and their interactions with earnings growth expectations in column (5), where forecast dispersion is computed as the standard deviation of analyst earnings forecasts, standardized by the absolute value of the average forecast. Column (5) shows the direct effect of growth expectations to be insignificant, but the interaction effect with the number of analysts is indeed significantly negative and the interaction effect with forecast dispersion is significantly positive. These interaction effects increase the adjusted R^2 to 49.3%.

[14] In unreported results, we consider two robustness checks. First, we relax the implicit assumption that all industries have the same "GDP growth beta". Because the cross-sectional variation in growth betas is co-linear with the fixed effects, we only add GDP growth multiplied with the corresponding GDP beta. Second, because current GDP growth may not be the best forecast of future GDP growth, we use forecasts from a regression model with current GDP growth and the yield spread as predictors. The results are robust.

Note that the direct effects of the number of analysts and forecast dispersion are also significant. The positive coefficient on the analyst variable is consistent with the fact that large firms with better analyst coverage are less dominated by retail investors, which tends to lead to lower prices. Higher forecast dispersion may effectively function as an indicator of optimism and cause current over-pricing (see Diether, Malloy and Scherbina, 2002) explaining its negative sign.

To summarize, we find that growth prospects are part of the explanation for the time variation in the China-U.S. earnings yield gap, especially when we focus on the 2003 to 2018 sample, but they cannot explain away the structural break. Our finding that growth prospects are priced more strongly into Chinese stock prices since 2003 is consistent with the results in Carpenter, Lu and Whitelaw (2021) regarding the price informativeness of the Chinese stock market.

IV.D Horse Race Between Discount Rates and Growth Prospects Channels

Our analysis of various determinants of the Chinese-U.S. EY differentials suggests that financial openness and domestic ownership structure not only explain a non-trivial fraction of their overall variation, but may also have contributed to the change in valuation differentials from negative to positive over time. Stock market development and growth expectations variables explain variation in EY differentials as well, but do not fully explain the change in the valuation gap. We now run a horse race among the various potential explanations. In order to obtain a parsimonious set of factors, we employ the general-to-specific search algorithm of Hendry (1995) and Hendry and Krolzig (2001), implemented in PcGets. This algorithm eliminates insignificant variables through an intricate "testing-down" process and generates a final set of variables with significant coefficients. A detailed discussion is provided in Appendix F.[15]

In Table 5, we report the results of applying the PcGets procedure to the long sample of

[15] We also examine robustness using an alternative, simpler model selection procedure which selects variables using a two-step procedure based only on univariate t-stats. The results are largely robust (see Online Appendix Table OA7).

1995-2018 and the short sample of 2003-2018. There are a total of 14 (19) variables for the long (short) sample, respectively. For each specification, we report the selected variables, the final coefficients and t-statistics in the first column, and a variance decomposition in the second column for both samples. The variance decomposition reports the covariance between the product of each coefficient and the corresponding independent variable with the fitted value in the regression, divided by the variance of the fitted value. The numbers therefore add up to 100%.

For the 1995-2018 sample, 6 out of 14 variables are selected, and the adjusted R^2 is 43.8%. Four groups of variables are represented. The selected “Financial Openness” variables are unmistakably the main driver of the explained variation. REGOPEN accounts for 18% of the explained variation. Most important by far is the IA2 variable, measuring the market capitalization represented by B, H and ADR shares, which accounts for 45.2% of the explained variation. State ownership survives the model selection procedure and accounts for 7.4% of the explained variation. The “Market Development” category accounts for 16.7% of the explained variation, accounted for by the Zeros and Turnover variables. The explanatory power of GDP growth rate is relatively low and accounts for 4.9% of the explained variation.

We conclude that the most important driver of the cross-sectional and time-series variation in the China-U.S. earnings yield differentials is the cross-sectional and time-series variation in international accessibility. Note that the break dummy is not selected.

For the more recent 2003-2018 sample, the variables selected by the PcGets procedure are similar to those for the long sample period. The main category remains “Financial Openness”, with the IA2 variable now accounting for 32.8% of the total explained variation. REGOPEN also survives the model selection procedure and contributes 6.4% to the explained variation. The second most important variable is state ownership, accounting for 23.8% of the explained variation.

For the market development variables, both Zeros and Turnover remain important, jointly accounting for 15.5% of the explained variation. GDP growth now accounts for 11.1% of the explained variation but no analyst variables survive the selection procedure. The overall adjusted R^2 is now 52.2%.

Our results show that both discount rates and growth prospects channels contribute to the cross-sector and time-series variation of the valuation differentials, but financial openness is the most important contributor. Our selected variables account for a significant part of the observed EY differential variations (43.8% for the long sample and 52.2% for the short sample). Figure 4, Panel A further plots the actual EY differential at the market level and the EY differential predicted from our PcGets model. In both the long and short samples, the time series of the predicted values closely match the data time series. If we regress the actual market level valuation gap on the model predicted values, the adjusted R^2 is 61% for the long sample and 58% for the short sample.

V. The China EY Model

In this section, we discuss results for the China EY Model and also consider some extensions accommodating cross-portfolio and time variation in betas.

V. A. Results for the China EY Model

Table 6 contains the estimation results for the China EY Model, using the PcGets exercise to fit the China valuation ratios with the four groups of variables introduced in Section IV. Three interesting patterns are apparent. First, the adjusted R^2 is 61.5% for the 1995-2018 sample, and 65.1% for the shorter 2003-2018 sample. These R^2's are considerably higher than the regressions for the differenced earnings yield, where it is 43.8% for the longer, 52.2% for the shorter sample.

Second, for the long sample, REGDEV now survives and accounts for 24.1% of the explained variation. Its sign is positive so that regulatory financial development, including allowing short selling, leads to lower valuations and higher EYs. This increases the relative contribution of the market development variables to above 30%, which comes mostly at the cost of the relative contribution of the financial openness variables which decreases from over 60% to 38%. While a number of additional variables are selected among the financial openness variables, they jointly have a negative variance contribution and the IA2 variable remains the key variable. State ownership accounts for about 7% and growth expectations for about 4% of the explained variation, similar to the results for the Valuation Difference Model.[16]

Third, for the shorter sample, starting in 2003, the contribution of financial openness decreases from 38.0% to 31.8% with the slack picked up by ownership structure and growth expectations. The total contribution of domestic ownership structure variables is over 30%, with nearly 60% accounted for by state ownership and the remainder by the retail ownership variable "SNS". Among the growth expectations variables, the earnings growth expectations variable now gets selected and is more important than GDP growth. Financial openness and ownership structure remain the dominant variables explaining the temporal and cross-portfolio variation in Chinese earnings yields. Also, note that for the short sample, the contribution of turnover, which could be an indirect measure of retail ownership, is 10.9%.

In Figure 4, Panel B, we plot the fitted value for the Chinese earnings yield at the market level, presented in dashed lines (long sample) and dotted lines (short sample). The time series presented are aggregated from portfolio level earnings yields. If we regress the market level Chinese earnings yield on the predicted values using the China model, the adjusted R^2's are 83% for the long sample and 77% for the short sample. Clearly, our Chinese earnings yield model

[16] Splitting up REGDEV into a variable associated with short-sell regulations and one tracking the remaining regulatory changes, leaves the results largely unchanged. For the long sample, both variables are selected and their combined variance contribution is very close to that of REGDEV.

captures the bulk of the variation in the Chinese earnings yield.

V. B. Model Extensions

In this section, we consider several extensions of our base model, accommodating cross-sectional and/or time series variation in the exposure to aggregate cash flow expectations and discount rates.

We first consider including cash flow betas. In terms of growth prospects, the use of GDP growth implicitly assumes that different portfolios have the same exposure with respect to aggregate growth expectations. To accommodate different exposures to GDP growth, we measure the sensitivity of EY for a particular portfolio to changes in GDP. Specifically, we regress portfolio level earnings yields on GDP growth over the full sample using quarterly data, and then define the coefficient as the GDP beta, as follows,

$$EY_{j,t} = c_j + \beta_j^{GDP} GDP_t + \varepsilon_{j,t} \tag{7}$$

where $EY_{j,t}$ is portfolio level earnings yield, GDP_t is the GDP growth rate, and β_j^{GDP} represents the GDP beta. To understand whether GDP betas help to explain Chinese EY, we include GDP betas and their interaction with the GDP growth rate in Equation (4) for the PcGets procedure. This model generates higher adjusted R^2 for both the long and short sample periods. However, the relative contributions of the financial openness, domestic ownership and financial development variables to the explanatory power are virtually unchanged.

As to discount rates, our model can be viewed as assuming a portfolio-specific discount rate function linear in the explanatory variables. A slightly different approach would assume that the aggregate Chinese discount rate is portfolio-specific, because different investors (e.g. retail versus foreign investors) might use different discount rates. Meanwhile, different industries may also exhibit different exposures to these aggregate discount rates. In addition, Chinese stocks may also have exposure to global factors, which we proxy for by the U.S. market (see Appendix G for model details). We empirically consider several such full “segmentation” and “partial segmentation” models, and estimate either constant or time-varying betas for our various portfolios using weekly return data. Our panel estimations for the China EY model are then redone accommodating our initial variables and these variables multiplied with the portfolio betas. We invariably find that the additional variables simply “noise up” the regressions and adjusted R^2s are smaller than in our original specification. We therefore do not further consider these specifications and relegate the results to the Online Appendix Table OA5. These results are not surprising, because while showing some time variation, the betas of Chinese portfolios with respect to the Chinese stock market return are invariably close to 1.0. This indicates that beta variation is a second-order effect in our analysis.

VI. Further Discussions

We argue that major changes in valuations in China and their differences relative to the U.S. benchmark are most closely associated with a gradual financial openness process, best measured by our IA2 variable, measuring the market capitalization of international shares relative to the total market capitalization in that particular industry. The direct and indirect shift of ownership from locals to foreigners caused valuations to drop and the curious China valuation premium to disappear over time. Section VI.A provides more color on the mechanism behind the panel regression result through an analysis of the industry specific effects and the associated spillover effects. We also show the impact of the IPO wave that took place in from the third quarter in 2006 and throughout 2007. Section VI.B then, even more specifically, discusses the role of the banks and life insurance sector in the documented valuation gap changes. Section VI.C summarizes some further robustness checks.

VI. A Cross-Industry Analysis of Valuation Gap Changes

It is informative to sketch the economic background behind our findings. First, we already know that foreigners price Chinese stocks lower than do Chinese locals from the extensive literature on the premium of A-shares relative to B-shares (see e.g. Mei, Scheinkman and Xiong, 2009). In 2001, when Chinese citizens were allowed to invest in B-shares, the A-B share premium decreased substantially, but because of the small size of the B-share market, this did not generate a large effect on market wide valuations in the A-share market.

Second, the gradual liberalization process then involved multiple cross-listings of A-shares in the form of B and H shares and other international listings. As Figure 3 shows, this process was not smooth and the IA2 variable does not show a uniform trend upwards. The presence of international listings at lower valuations put relative downward pressure on the A-share prices themselves and also likely generated spillover effects within and outside the industry of the cross-listing firm. Such local spillover effects have been documented in the international finance literature on ADRs (see e.g. Fernandes, 2009).

We now illustrate this process with industry-specific data in China. Figure 5, Panel A shows on the vertical axis the slope coefficient of a portfolio-specific time-series regression of the EY of the international firms within an industry on a constant and the industry IA2 variable. On the horizontal axis we show the same slope coefficient, but for the EY of the domestic portion of the industry portfolio. When the coefficients are positive and statistically significant at the 10% level for both, the dots are solid. It is apparent that the bulk of the coefficients are positive, consistent with our panel result. Moreover, while our panel result used the full industry portfolio, we now see that these results are not solely the result of direct pricing effects on the international firms but also reflect spillover effects. Moreover, the spillover effect can extend across industries.

When we run a panel regression of industry EYs on industry IA2 and aggregate IA2, where the latter is the average IA2 across industries, we find that both are highly statistically significant, and the second coefficient (at 0.329) is in fact much larger in magnitude than the first (0.123), suggesting we may have even under-estimated the valuation effects of increased openness.

Furthermore, our findings cannot be decoupled from the rapid growth in the Chinese stock market through a large IPO wave occurring between 2006Q3 and 2007Q4. It was the IPO wave that not only made China in the second largest equity market in the world but it also helped substantially increase the internationalization of the Chinese stock market. Figure 5, Panel B shows the evolution of the IA2 variable at the market level with and without the firms that conducted their IPOs during the IPO wave. When the IPO firms are removed, the IA2 variable actually peaks around 2008 but then declines. However, including the IPO wave causes a sharp increase of the IA2 variable. In terms of market capitalization, more than 70% of the IPOs were in the Oil Equipment, Services & Oil and Gas Producers, Banks & Life Insurance, and Mining industries. By the end of our sample, these industries were in the top 5 in terms of internationalization.[17]

Moreover, the Construction & Materials, Chemicals, and Industrial Engineering industries were among the top industries in terms of number of IPOs and were in the top 5 of IA2 levels early in the sample. Thus, the IPO wave facilitated the growth and internationalization of the Chinese equity market, which led to lower valuations a few years later as firms internationalized further. Still, the banking sector played a particularly important role in this process, to which we now turn.

[17] Note that these industries are also of somewhat strategic importance and in fact had relatively high state ownership, especially the Banks and Life insurance and Oil Equipment, Services & Oil and Gas Producers industries. At the end of the sample, state ownership in those two industries had actually increased relatively to the early part of our sample. Thus, the positive effect of the state ownership variable on earnings yield differentials may be partially driven by these industries as well.

VI.B. The Role of the Banks & Life Insurance Sector

Recall that Figure 2 shows an outsized role played by the "Banks & Life Insurance" portfolio in the evolution of valuation gap. Its EY differential moves from a negative -2.5% in 1995-1999 to a large positive 7.2% in 2014-2018. Could the banking industry be the main driver of the market-wide changes in the valuation gap? This can only be true if it comprises a large portion of the market. Figure 6, Panel A presents the market share of the banking sector. Before 2007, the Banks and Life insurance sector constituted a small fraction (lower than 10%) of the total market capitalization. Then, its market share increased substantially after several important IPOs of state-owned banks in the mid-2000s, exceeding 30% at one point. This relative increase in the importance of the banking sector is even more dramatic when market shares are computed in terms of earnings (which drive the sector weights in PE ratios). Given its small market share, the banking sector cannot account for the negative EY differences observed in the early part of the sample, but it can contribute to the rise of the EY differential in the later part of the sample.

To examine this conjecture, we present the EY differentials with and without the banking portfolio in Figure 6, Panel B. Up until 2007, the two lines essentially coincide, reflecting that the banking portfolio constituted a negligible part of the market. However, after 2007, and especially after 2009, the two lines diverge with the increase in the EY differential more pronounced for the overall statistic than for the one without the banking portfolio. We also conduct the structural break test without the banking sector. When we exclude the banking portfolio from the market yields, the sup-Wald test is 3.64, which is not significant at the 10% level. Clearly, this indicates that valuation changes in the banking portfolio are an important contributor to the structural break.

While this is true, it is entirely consistent with our finding that the financial openness channel is the main driver of valuation gap changes. This is because the banking sector, on average, features the second highest international accessibility out of all 33 sector portfolios. Moreover, its international accessibility increases dramatically over time due to the dual listings of big state-owned banks in the A-share market and the Hong Kong Stock Exchange. Therefore, the valuation change in the banking industry is largely explained by the financial openness channel. The lower international valuations may be due to foreign investors being more skeptical about the well-document non-performing loan problems in Chinese banks and releasing this information to the A-share market (see Zhang, Cai, Dickinson and Kutan, 2016).

VI. C. Robustness Checks

It is possible that ownership reacts to valuations rather than the other way around. We therefore also re-run our regressions lagging the independent variables by one quarter. The Online Appendix Table OA6 shows that all the key results are nearly unchanged.

Our construction of the REGOPEN/REGDEV variables makes implicit assumptions on the relative importance of the various regulatory reforms. We have therefore also considered regressions where the various reforms are added as separate dummy variables. Among the regulatory reforms, the introduction of RMB Qualified Foreign Institutional Investors (RQFII) in 2011Q4, which allows RQFII quota-holders to invest directly into Chinese A shares with their offshore RMB, is the most important. It accounts for 24.6% (9.8%) of the total explained variation in EY differentials in the long (short) sample. The first transaction of Qualified Foreign Institutional Investors (QFII) and the announcement of Shanghai/Shenzhen-Hong Kong Connect contribute 9.5% and 5.3%, respectively, to the explained variation in the long sample.

VII. Conclusion

We study valuation differentials in China and the U.S. over the past 24 years at the portfolio level. We first document a curious valuation gap, with Chinese price earnings ratios being substantially higher than those of the U.S. in the first half of our sample (which starts in 1995), in contrast to the usual discount observed for emerging markets. The valuation gap disappears in the second half of the sample. There is also a cross-portfolio dimension to these valuation gaps, both in terms of magnitude and sign. With valuations linked to costs of capital and signals about future growth opportunities, it is important to understand what drives these valuation differentials and their evolution over time.

Focusing on earnings yield differentials, we examine a number of potential explanations. First, we examine differences in industry structure across China and the U.S. We find that sector differences generally play a minor role in driving valuation differentials across the two countries. Second, we then focus on valuation differentials across industries and consider four groups of potential explanations: financial openness, domestic ownership structure, market development, and growth prospects. For the longer sample, the most important variable group by far is financial openness, with international accessibility the dominant variable, accounting for more than 70% of the explained variation, followed by market development (15%). We have insufficient information on growth prospects, and growth expectations account for less than 4% of the explained variation. For the shorter sample, financial openness still accounts for 39.2% of the explained variation. With more data available on domestic ownership structure, its explanatory power increases to 35.7%, whereas growth prospects account for 10.8% of the explained variation. This result is rather robust across different specifications. The banking sector's increased prominence going hand in hand with its increased earnings yields did play a non-trivial role in engineering higher earnings yields for the Chinese market as a whole. However, the role of the banking sector is also driven by its increased foreign ownership over time, being the portfolio with the second highest international accessibility out of all 33 industry portfolios.

China witnessed a gradual opening of its shares to foreign investors, and foreign investors price Chinese stocks at lower valuations than do domestic investors, especially retail investors. This gradual integration of Chinese into global capital markets helped eliminate the Chinese valuation gap. However, the valuation gap has not disappeared completely for all industries. If foreigners' value Chinese stocks at more realistic valuations than do domestic investors, the increased foreign ownership may in fact make stock market valuations in China more informative to economic policy makers, e.g. in predicting economic activity. We defer testing this conjecture to future research, but it is noteworthy that the explanatory power of growth-related variables has increased over time. In addition, with Chinese stocks now priced more like regular emerging market stocks, further liberalizations may lead to price revaluations, rather than devaluations, as is already apparent in the effects of the recent Shanghai-Hong Kong Stock connect program (see Chan and Kwok, 2018; Liu, Wang and Wei, 2021; Ma, Rogers and Zhou, 2021).

References

Allen, F., Qian, J., Shan, C., & Zhu, J. (2022). Dissecting the long-term performance of the Chinese stock market. *Working Paper*.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets*, 5(1), 31-56.

Amihud, Y., Mendelson, H., & Pedersen, L. H. (2006). Liquidity and asset prices. *Foundations and Trends® in Finance*, 1(4), 269-364.

Bai, J., Lumsdaine, R. L., & Stock, J. H. (1998). Testing for and dating common breaks in multivariate time series. *Review of Economic Studies*, 65(3), 395-432.

Bailey, W., Chung, Y. P., & Kang, J. K. (1999). Foreign ownership restrictions and equity price premiums: what drives the demand for cross-border investments?. *Journal of Financial and Quantitative Analysis*, 34(4), 489-511.

Bekaert, G. (1995). Market integration and investment barriers in emerging equity markets. *World Bank Economic Review*, 9*(1)*, 75-107.

Bekaert, G., & Harvey, C. R. (2000). Foreign speculators and emerging equity markets. *Journal of Finance*, 55(2), 565-613.

Bekaert, G., & Harvey, C. R. (2003). Emerging markets finance. *Journal of Empirical Finance*, 10(1-2), 3-55.

Bekaert, G., Harvey, C. R., & Lumsdaine, R. L. (2002). Dating the integration of world equity markets. *Journal of Financial Economics*, 65(2), 203-247.

Bekaert, G., Harvey, C. R., & Lundblad, C. T. (2003). Equity market liberalization in emerging markets. *Journal of Financial Research*, 26(3), 275-299.

Bekaert, G., Harvey, C. R., & Lundblad, C. (2007). Liquidity and expected returns: Lessons from emerging markets. *Review of Financial Studies*, 20(6), 1783-1831.

Bekaert, G., Harvey, C. R., Lundblad, C. T., & Siegel, S. (2011). What segments equity markets?. *Review of Financial Studies*, *24(*12), 3841-3890.

Ben-Nasr, H., Boubakri, N., & Cosset, J. C. (2012). The political determinants of the cost of equity: Evidence from newly privatized firms. *Journal of Accounting Research*, 50(3), 605-646.

Boubakri, N., El Ghoul, S., Guedhami, O., & Megginson, W. L. (2018). The market value of government ownership. *Journal of Corporate Finance*, 50, 44-65.

Carpenter, J. N., Lu, F., & Whitelaw, R. F. (2021). The real value of China's stock market. *Journal of Financial Economics*, 139(3), 679-696.

Chan, M. K., & Kwok, S. (2018). Connecting the markets? Recent evidence on China's capital account liberalization. *Economic Modelling*, 70, 417-428.

Chan, K., Menkveld, A. J., & Yang, Z. (2007). The informativeness of domestic and foreign investors' stock trades: Evidence from the perfectly segmented Chinese market. *Journal of Financial Markets*, 10(4), 391-415.

Chan, K., Menkveld, A. J., & Yang, Z. (2008). Information asymmetry and asset prices: Evidence from the China foreign share discount. *Journal of Finance*, 63(1), 159-196.

Diether, K. B., Malloy, C. J., & Scherbina, A. (2002). Differences of opinion and the cross section of stock returns. *Journal of Finance*, *57*(5), 2113-2141.

Doidge, C., Karolyi, G. A., & Stulz, R. M. (2020). Is financial globalization in reverse after the 2008 global financial crisis? Evidence from corporate valuations (No. w27022). *National Bureau of Economic Research*.

Erb, C. B., Harvey, C. R., & Viskanta, T. E. (1996). Political risk, economic risk, and financial risk. *Financial Analysts Journal*, 52(6), 29-46.

Errunza, V., & Losq, E. (1985). International asset pricing under mild segmentation: Theory and test. *Journal of Finance*, 40(1), 105-124.

Faccio, M., Masulis, R. W., & McConnell, J. J. (2006). Political connections and corporate bailouts. *Journal of Finance*, 61(6), 2597-2635.

Fernald, J., & Rogers, J. H. (2002). Puzzles in the Chinese stock market. *Review of Economics and Statistics*, 84(3), 416-432.

Fernandes, N. (2009). Market liberalizations at the firm level: Spillovers from ADRs and implications for local markets. *Journal of International Money and Finance*, 28(2), 293-321.

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

Ferreira, M. A., & Santa-Clara, P. (2011). Forecasting stock market returns: The sum of the parts is more than the whole. *Journal of Financial Economics*, 100(3), 514-537.

Frankel, J. A. (1992). Measuring international capital mobility: a review. *American Economic Review*, 82(2), 197-202.

Hendry, D. F. (1995). Dynamic Econometrics. *Oxford University Press.*

Hendry, D. F. & Krolzig, H. M. (2001). Computer automation of general-to-specific model selection procedures. *Journal of Economic Dynamics and Control*, 25(6-7), 831-866.

Jiang, F., & Kim, K. A. (2015). Corporate governance in China: A modern perspective. *Journal of Corporate Finance*, *32*, 190-216.

Koijen, R. S., & Yogo, M. (2019). A demand system approach to asset pricing. *Journal of Political Economy*, 127(4), 1475-1515.

Lesmond, D. A. (2005). Liquidity of emerging markets. *Journal of Financial Economics*, *77*(2), 411-452.

Lesmond, D. A., Ogden, J. P., & Trzcinka, C. A. (1999). A new estimate of transaction costs. *The Review of Financial Studies*, 12(5), 1113-1141.

Levine, R., & Zervos, S. (1998). Stock markets, banks, and economic growth. *American Economic Review*, 88(3), 537-558.

Liao, L., Liu, B., & Wang, H. (2014). China′ s secondary privatization: Perspectives from the split-share structure reform. *Journal of Financial Economics*, 113(3), 500-518.

Liu, J., Stambaugh, R. F., & Yuan, Y. (2019). Size and value in China. *Journal of Financial Economics*, *134*(1), 48-69.

Liu, C., Wang, S., & Wei, K. J. (2021). Demand shock, speculative beta, and asset prices: Evidence from the Shanghai-Hong Kong stock connect program. *Journal of Banking & Finance*, *126*, 106102.

Love, I. (2003). Financial development and financing constraints: International evidence from the structural investment model. *Review of Financial Studies*, 16(3), 765-791.

Ma, C., Rogers, J. H., & Zhou, S. (2021). The effect of the China Connect. *SSRN Working Paper*

Martin, I. (2017). What is the Expected Return on the Market?. *Quarterly Journal of Economics*, 132(1), 367-433.

Megginson, W. L., Nash, R. C., & Van Randenborgh, M. (1994). The financial and operating performance of newly privatized firms: An international empirical analysis. *Journal of Finance*, 49(2), 403-452.

Mei, J., Scheinkman, J. A., & Xiong, W. (2009). Speculative Trading and Stock Prices: Evidence from Chinese AB Share Premia. *Annals of Economics & Finance*, *10*(2), 225-255.

Pástor, Ľ., & Veronesi, P. (2006). Was there a Nasdaq bubble in the late 1990s?. *Journal of Financial Economics*, 81(1), 61-100.

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

Shleifer, A., & Vishny, R. W. (1997). A survey of corporate governance. *Journal of Finance*, 52(2), 737-783.

Shumway, T., Szefler, M., & Yuan, K. (2009). The information content of revealed beliefs in portfolio holdings. *University of Michigan Working Paper*.

Thompson, S. B. (2011). Simple formulas for standard errors that cluster by both firm and time. *Journal of Financial Economics*, 99(1), 1-10.

Wurgler, J. (2000). Financial markets and the allocation of capital. *Journal of Financial Economics*, 58(1-2), 187-214.

Zhang, D., Cai, J., Dickinson, D. G., & Kutan, A. M. (2016). Non-performing loans, moral hazard and regulation of the Chinese commercial banking system. *Journal of Banking & Finance*, 63, 48-60.

Table 1. Summary Statistics

This table reports the time-series average of number of stocks, market value (MV) in billion U.S. dollars, market value in percentage of the total market value, and earnings yield (EY) for the market and each industry portfolio in China and the U.S., from 1995Q1 to 2018Q4. The Chinese sample covers all firms that listed in the A share market. The U.S. sample includes all common stocks listed in New York Stock Exchange, NASDAQ and American Stock Exchange. All variables are constructed on a quarterly basis. We calculated earnings at quarter t as the annualized net income by summing up net income from quarter t-4 to quarter t-1. MV is calculated as the sum of all stocks' price multiplied by common shares outstanding, converted to U.S. dollars using the quarter-end exchange rate. EY is total earnings divided by market value for common equity. Negative values of firm earnings are set to zero before being aggregated into sector level. The detailed description of the portfolio formation is in Appendix A.

	China				U.S.			
	n(stocks)	MV ($ billion)	MV (%)	EY (%)	n(stocks)	MV ($ billion)	MV (%)	EY (%)
Market	1,400	2,578	100.00	4.94	3,976	13,833	100.00	5.13
Industry Portfolios								
Aerospace & Defense	9	14	0.44	1.83	62	302	2.10	5.54
Alternative Energy	6	10	0.24	2.23	12	8	0.06	3.10
Automobiles & Parts	59	86	3.62	4.90	41	145	1.16	7.45
Banks & Life Insurance	10	589	13.77	7.82	536	1,223	9.13	7.15
Beverages	24	65	2.61	3.50	23	329	2.47	4.29
Chemicals	127	110	6.35	3.46	83	265	2.01	5.85
Construction & Materials	79	110	3.84	4.58	79	108	0.75	5.21
Electricity	44	79	5.05	5.44	57	351	2.64	6.42
Electronic & Electrical Equipment	115	119	4.33	2.52	190	181	1.38	4.30
Financial Services	14	83	2.01	3.39	142	718	5.01	6.37
Fixed and Mobile Telecom	3	12	0.64	5.23	58	516	4.23	5.20
Food & Drug Retailers	7	5	0.22	2.85	37	201	1.46	4.92
Food Producers	57	61	2.40	2.84	84	284	2.10	5.59
Forestry & Paper	17	10	0.51	3.82	16	31	0.27	5.21
Gas, Water and Multiutilities	14	14	0.78	3.07	51	137	1.02	5.72
General Industrials	19	17	1.11	4.04	48	489	3.71	4.91
General Retailers	65	51	3.14	3.18	205	881	6.15	4.52
Health Care Equipment & Services	10	11	0.23	2.16	261	523	3.65	4.53
Household Goods & Home Construction	26	40	1.49	5.16	101	287	2.10	5.51
Industrial Engineering	114	130	4.83	3.37	129	218	1.55	5.98
Industrial Metals & Mining	68	114	5.88	4.21	35	77	0.57	6.99
Industrial Transportation	43	78	3.59	4.35	67	212	1.49	5.98
Leisure Goods	19	19	1.59	3.30	49	72	0.51	4.05
Media	18	26	0.74	2.15	125	600	4.32	4.36
Mining	28	93	2.83	4.84	32	40	0.34	3.21
Oil Equipment, Services & Oil and Gas Producers	12	200	6.82	5.44	178	1,076	7.84	6.21
Personal Goods	55	42	1.90	3.45	80	194	1.44	4.94
Pharmaceuticals & Biotechnology	92	109	4.03	3.00	232	1,202	8.78	4.08
Real Estate Investment & Services	101	109	7.58	4.51	30	21	0.13	3.96
Software & Computer Services	39	44	1.30	1.88	279	1,281	8.46	3.92
Support Services	31	25	1.53	3.06	212	293	2.11	4.07
Technology Hardware & Equipment	47	60	2.49	2.57	280	1,208	8.64	4.39
Travel & Leisure	31	49	2.21	3.38	162	360	2.43	5.10

Table 2. Explaining the Valuation Gap: Changes in Industry Structure

This table shows the decomposition of the market level earnings yield differential into two components using the following formula: $DIF\_EY_t = EY_t^{CN} - EY_t^{US} = \sum_{j=1}^{N} w_{j,t}^{CN}\left(EY_{j,t}^{CN} - EY_{j,t}^{US}\right) + \sum_{j=1}^{N}\left(w_{j,t}^{CN} - w_{j,t}^{US}\right)EY_{j,t}^{US} = DIF\_VAL_t + DIF\_STR_t$. DIF_VAL represents the EY differential within the same sector between China and the U.S., and DIF_STR represents the valuation effect of the different industry structure between China and the U.S. Panel A presents the time-series average of the two components and variance decomposition results. Panel B shows the Bai, Lumsdaine and Stock (1998) break test results for the two components. ***, ** and * indicate significances at the 1%, 5% and 10% levels using two-tailed tests.

Panel A. Decomposition

				Variance Decomposition	
DIFEY (%) Mean	DIF_VAL (%) Mean	DIF_STR (%) Mean	DIFEY (%) Variance	Cov(DIF_VAL,DIFEY) /Var(DIFEY)	Cov(DIF_STR, DIFEY) /Var(DIFEY)
-0.19	-0.69	0.50	3.549	0.99	0.01

Panel B. Break Point Test

Variable	Sup-Wald Statistic	Estimated Break Point	90% Confidence Interval
DIF_VAL	11.68**	2009:02	2007:03-2011:01
DIF_STR	4.27	2000:04	1996:03-2006:01

Table 3. Explaining the Valuation Gap: Changes in Discount Rates

This table reports the results of portfolio-quarter panel regressions of valuation differentials on discount rate variables from 1995Q1 to 2018Q4. The dependent variable is the portfolio level earning yield differential between China and the U.S., DIFEY. All independent variables are differences between China and the U.S. except for "Dummy: 2009Q3" which is a dummy variable equal to 1 (0) during period after (before) 2009Q3. Financial openness variables include regulatory financial openness (REGOPEN) and three international accessibility measures (IA1, IA2 and IA3). Domestic ownership structure variables include state ownership, institutional ownership, retail ownership, and standardized number of shareholders (SNS). Market development variables include regulatory financial development (REGDEV), number of public firms, adjusted market development, zeros, and turnover. Control variables include leverage differentials, earnings growth differential and minimum number of stocks. Definitions of all these variables are described in detail in Appendix C. The regressions include portfolio fixed effects and the standard errors are double clustered by portfolio and time. T-statistics are in parentheses. ***, ** and * indicate significances at the 1%, 5% and 10% levels using two-tailed tests.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Break Dummy		0.007		0.025***		0.003		0.020***
		(1.294)		(5.347)		(0.605)		(3.962)
REGOPEN	0.004***	0.003**						
	(3.690)	(2.522)						
IA1	0.007	0.009						
	(0.896)	(1.016)						
IA2	0.141***	0.129***						
	(5.256)	(4.680)						
IA3	-0.016	-0.015						
	(-1.503)	(-1.418)						
Real interest rate	0.004	0.001						
	(0.069)	(0.027)						
Overall political rating	-0.039*	-0.022						
	(-1.899)	(-1.104)						
Chinese state ownership			0.023	0.056***	0.086***	0.089***		
			(1.551)	(3.757)	(3.906)	(4.059)		
Institutional ownership					-0.074***	-0.069***		
					(-4.744)	(-4.555)		
Retail ownership					-0.016	-0.013		
					(-1.359)	(-1.062)		
SNS					-0.060**	-0.054**		
					(-2.615)	(-2.236)		
REGDEV							0.015***	0.007**
							(4.349)	(2.252)
Number of public firms							-0.012**	-0.015**
							(-2.149)	(-2.623)
Adj. market development							-0.003***	-0.003***
							(-2.690)	(-2.682)
Zeros							0.127***	0.161***
							(4.085)	(5.255)
Turnover							-0.006***	-0.005***
							(-4.205)	(-4.082)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Portfolio fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample years	1995-2018	1995-2018	1995-2018	1995-2018	2003-2018	2003-2018	1995-2018	1995-2018
Number of observations	4,873	4,873	4,873	4,873	3,408	3,408	4,873	4,873
Adjusted R-square	0.392	0.396	0.235	0.334	0.449	0.450	0.357	0.380

Table 4. Explaining the Valuation Gap: Changes in Growth Prospects

This table reports the results of portfolio-quarter panel regressions of valuation differentials on growth prospects variables. The dependent variable is the portfolio level earning yield differential between China and the U.S., DIFEY. All independent variables are differences between China and the U.S. except for "Dummy: 2009Q3" which is a dummy variable equal to 1 (0) during period after (before) 2009Q3. The growth prospects variables include GDP growth rate, earnings growth expectation and sales growth expectation. GDP growth rate starts from 1995Q1. Due to availability of analyst forecast data, analyst-related variables including earnings growth expectation and sales growth expectation are only available after 2003Q1. Control variables include leverage differentials, earnings growth differential and minimum number of stocks. Definitions of all the variables are described in detail in Appendix C. The regressions include portfolio fixed effects and the standard errors are double clustered by sector and time. T-statistics are in parentheses. ***, ** and * indicate significances at the 1%, 5% and 10% levels using two-tailed tests.

	(1)	(2)	(3)	(4)	(5)
Break Dummy		0.019***		0.010**	0.009**
		(4.288)		(2.406)	(2.028)
GDP growth rate	-0.184***	-0.138***	-0.328***	-0.283***	-0.314***
	(-3.036)	(-2.790)	(-4.642)	(-4.992)	(-5.359)
Earnings growth expectation			-0.002	-0.002	0.003
			(-1.401)	(-1.272)	(0.740)
Sales growth expectation			-0.025**	-0.021**	-0.017
			(-2.296)	(-2.069)	(-1.576)
Earnings growth expectation×number of analysts					-0.011***
					(-3.461)
Earnings growth expectation×forecast dispersion					0.037***
					(4.065)
Number of analysts					0.004*
					(1.679)
Forecast dispersion					-0.025***
					(-2.947)
Controls	Yes	Yes	Yes	Yes	Yes
Portfolio fixed effects	Yes	Yes	Yes	Yes	Yes
Sample years	1995-2018	1995-2018	2003-2018	2003-2018	2003-2018
Number of observations	4,873	4,873	3,317	3,317	3,204
Adjusted R-square	0.250	0.314	0.418	0.432	0.493

Table 5. PcGets Model Selection (Valuation Difference Model)

This table reports the PcGets model selection results based on the valuation difference model using all explanatory variables. The left panel shows results for variables available from 1995 to 2018 while the right panel shows results for variables from 2003 to 2018. The dependent variable is the portfolio level earning yield differential between China and the U.S., DIFEY. All independent variables are differences between China and the U.S. except for "Dummy: 2009Q3" which is a dummy variable equal to 1 (0) during period after (before) 2009Q3 and overall political rating which is constructed by taking the ratio of Chinese over U.S. variables. The regressions include portfolio fixed effects and the standard errors are double clustered by sector and time. We apply the PcGets procedure to pick up the most important independent variables. The overall variance contribution of each selected variable is reported in columns (2) and (4). A detailed description of the PcGets procedure is provided in Appendix E. T-statistics are in parentheses. ***, ** and * indicate significances at the 1%, 5% and 10% levels.

	1995-2018		2003-2018	
	(1)	(2)	(3)	(4)
Financial Openness				
REGOPEN	0.004***	18.0%	0.003***	6.4%
	(5.175)		(4.288)	
IA2	0.199***	45.2%	0.154***	32.8%
	(7.343)		(5.625)	
Domestic Ownership Structure				
Chinese state ownership	0.031***	7.4%	0.065***	23.8%
	(2.967)		(4.250)	
Market Development				
Zeros	0.104***	10.6%	0.186***	9.3%
	(4.150)		(4.536)	
Turnover	-0.003***	6.1%	-0.004**	6.2%
	(-2.948)		(-2.535)	
Growth Prospects				
GDP growth rate	-0.247***	4.9%	-0.257***	11.1%
	(-4.303)		(-3.633)	
Control Variables				
Minimum number of firms	-0.010***	-22.1%		
	(-4.450)			
Portfolio Fixed Effects		29.9%		10.5%
Total Variance Contribution		100%		100%
Number of observations	4,873		3,317	
Adjusted R-square	0.438		0.522	

Table 6. PcGets Model Selection (China EY Model)

This table reports the PcGets model selection results based on the China EY model using all explanatory variables. The dependent variable is the Chinese earnings yield (EY^{CN}). The left panel shows the result for variables available from 1995 to 2018 while the right panel shows the result for variable from 2003 to 2018. The regressions include portfolio fixed effects and the standard errors are double clustered by sector and time. We apply the PcGets procedure to pick out the most important independent variables. The overall variance contribution of each selected variable is reported in columns (2) and (4). Detailed description of the PcGets procedure is provided in Appendix E. T-statistics are in parentheses. ***, ** and * indicate significances at the 1%, 5% and 10% levels.

	1995-2018		2003-2018	
	(1)	(2)	(3)	(4)
Financial Openness				
IA2	0.175***	38.5%	0.146***	31.8%
	(8.121)		(6.034)	
Real interest rate	-0.139***	4.8%		
	(-5.051)			
Overall political rating	0.001***	-5.3%		
	(2.727)			
Domestic Ownership Structure				
Chinese state ownership	0.026***	7.2%	0.050***	20.0%
	(3.342)		(4.557)	
SNS			-0.076***	14.2%
			(-6.283)	
Market Development				
REGDEV	0.011***	24.1%		
	(6.506)			
Adjusted market development	-0.003***	-0.2%	-0.002***	1.5%
	(-4.039)		(-2.951)	
Zeros	0.177***	10.4%	0.140***	7.1%
	(4.187)		(2.783)	
Turnover	-0.003***	5.2%	-0.006***	10.9%
	(-3.592)		(-3.718)	
Growth Prospects				
GDP growth rate	-0.320***	4.3%	-0.154**	3.5%
	(-4.670)		(-2.266)	
Earnings growth expectation			-0.016***	9.4%
			(-3.829)	
Portfolio Fixed Effects		11.0%		1.5%
Total Variance Contribution		100%		100%
Number of observations	4,873		3,337	
Adjusted R-square	0.615		0.651	

Figure 1. Valuation Gap between China and U.S.

This figure plots the time-series of earnings yield (EY) for China, the U.S., and the Datastream Emerging Market Index, during the period of 1995Q1 - 2018Q4. For China and the U.S., we construct the market level EY ratio from individual firm data following the Datastream method which is also applied in BLHS (2011). We first calculated firm level earnings at quarter t as the trailing annualized net income by summing up net income from quarter t-4 to quarter t-1. Negative values of firm earnings are set as zero before being aggregated into market level. Total market value is calculated as the summation of all the stocks' price multiplied by common shares outstanding at the end of each quarter. EY is total earnings divided by total market value. The data for the Emerging Market Index is obtained from Datastream using the data series of "TOTMKEK".

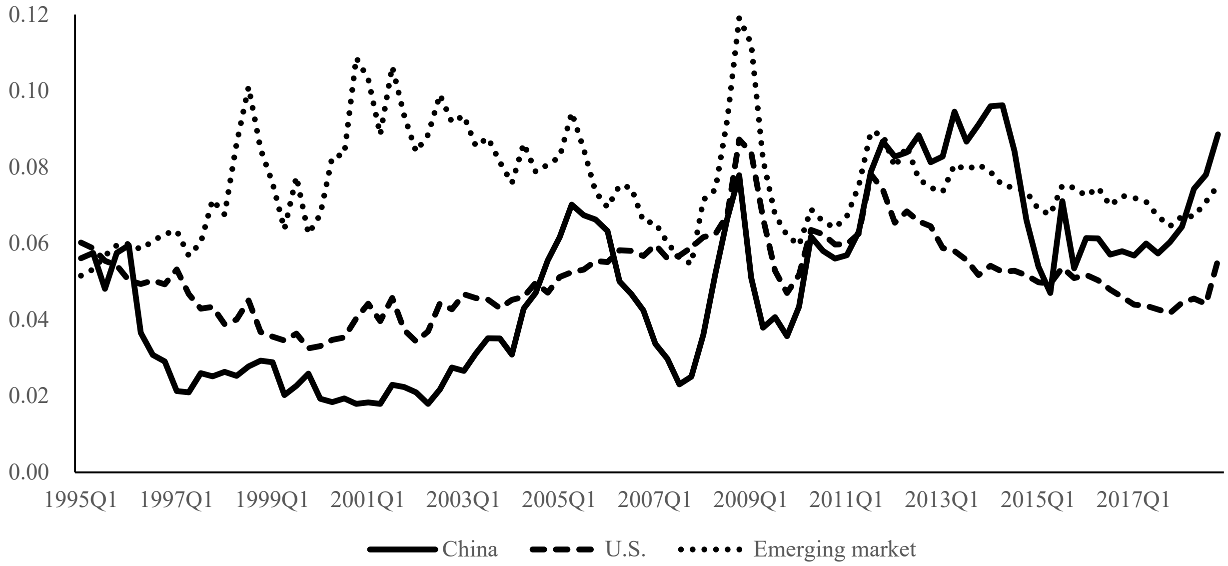


Figure 2. Change in Earnings Yield Differentials by Sector

This figure shows the evolution of earnings yield (EY) differentials for different sectors. The X-axis shows the average EY differentials (Chinese sector level EY minus U.S. sector level EY) during the first five years of our sample (1995-1999). The Y-axis shows the average EY differentials during the last five years of our sample (2014-2018).

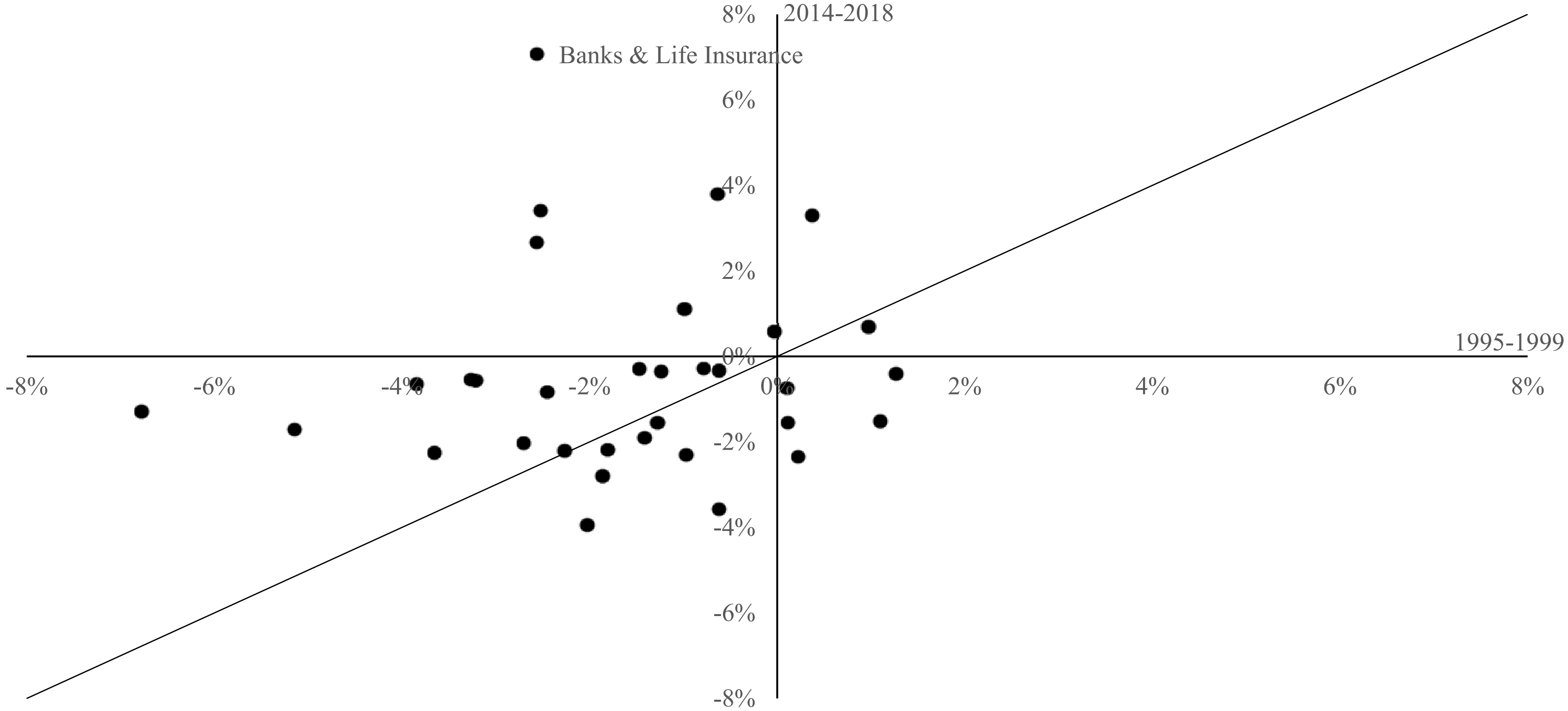


Figure 3. Time-Series Plots of China Aggregate Variables

This figure shows the time-series plots of aggregate international accessibility, ownership, turnover, and growth prospects measures for Chinese stock market. Firm level variable construction details are shown in the Appendix C. To obtain market level values, we value weight firm level variables.

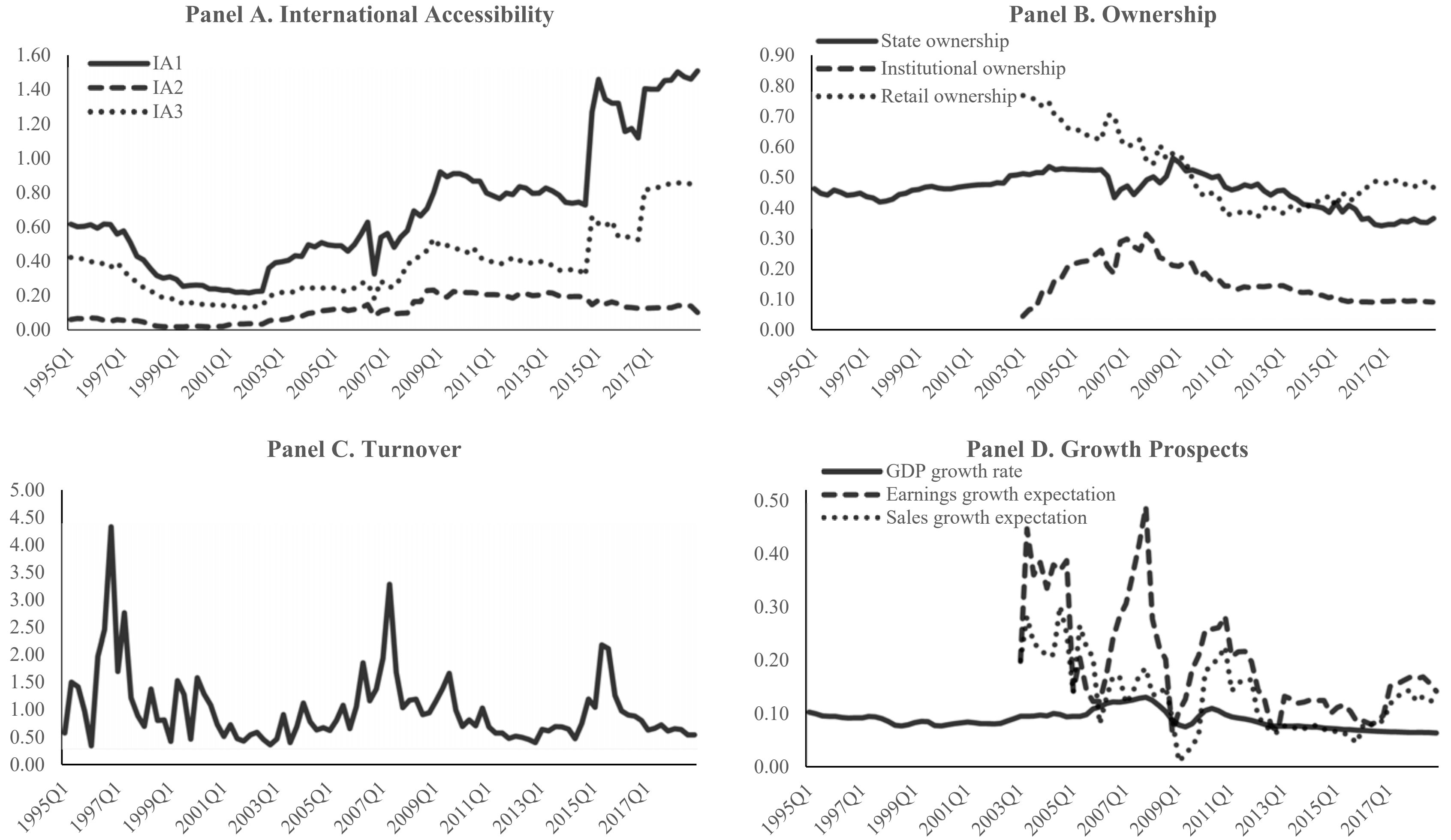
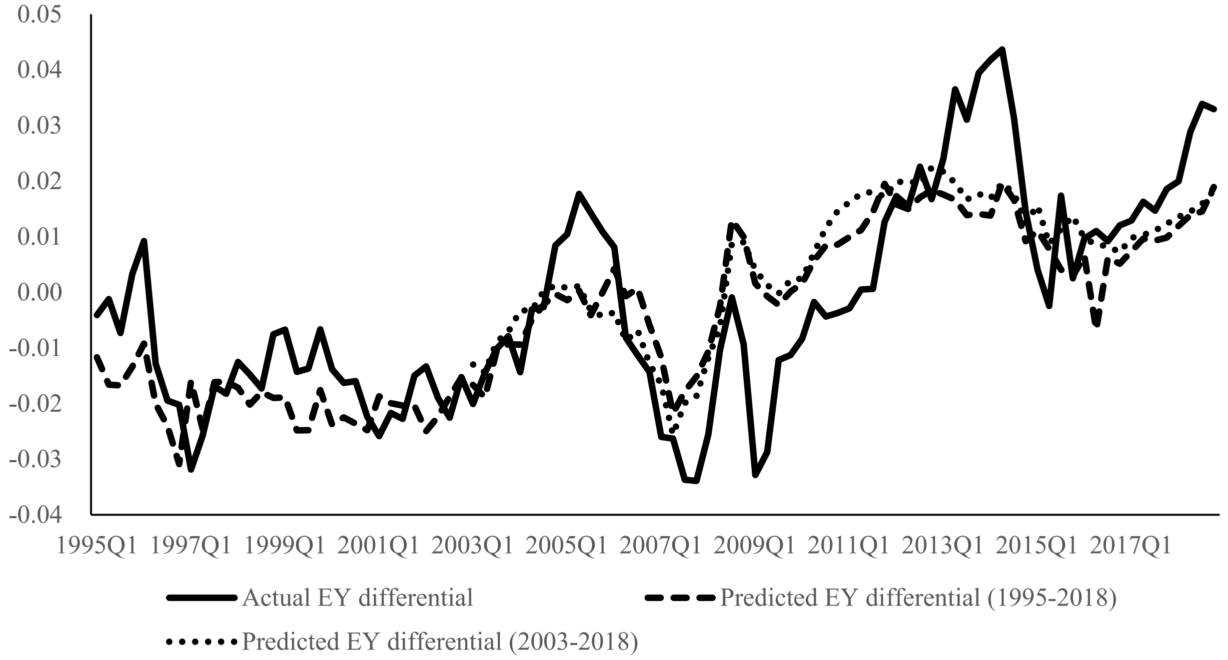


Figure 4. Model Fitness

This figure compares the actual earnings yield differentials and Chinese earnings yields with their predicted values. Panel A presents the actual earnings yield differential and its predicted values from the Valuation Difference model. Panel B presents the actual Chinese earnings yield and its predicted values from the China EY model.

Panel A. Valuation Difference Model



Panel B. China EY Model

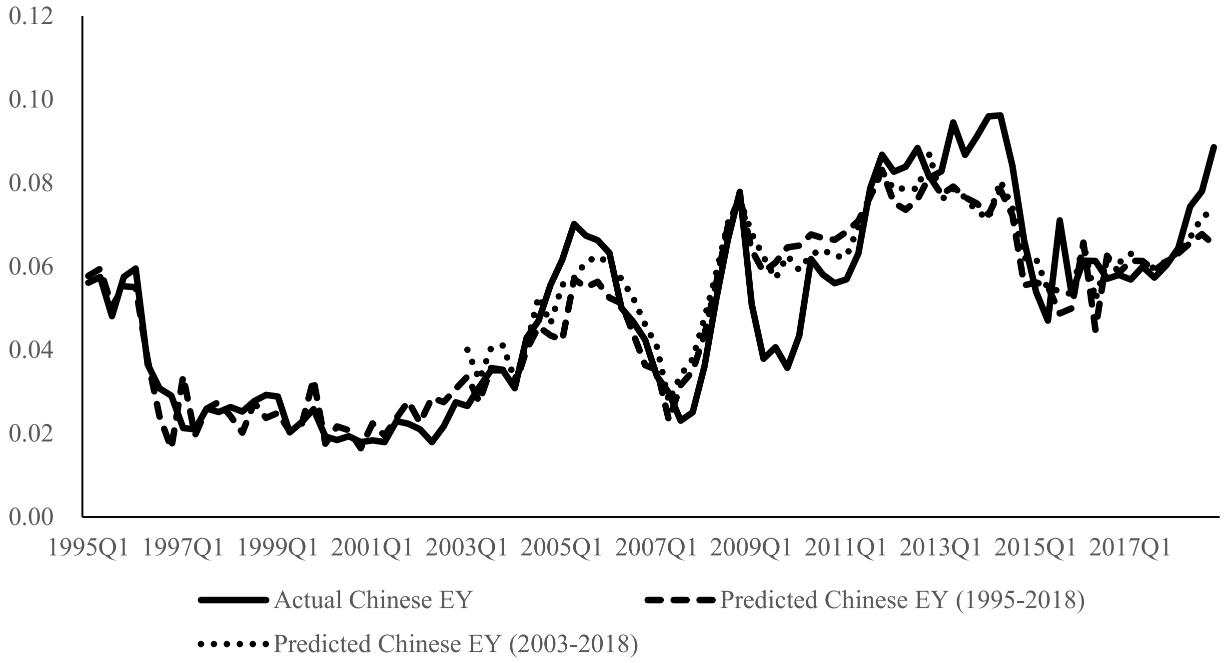
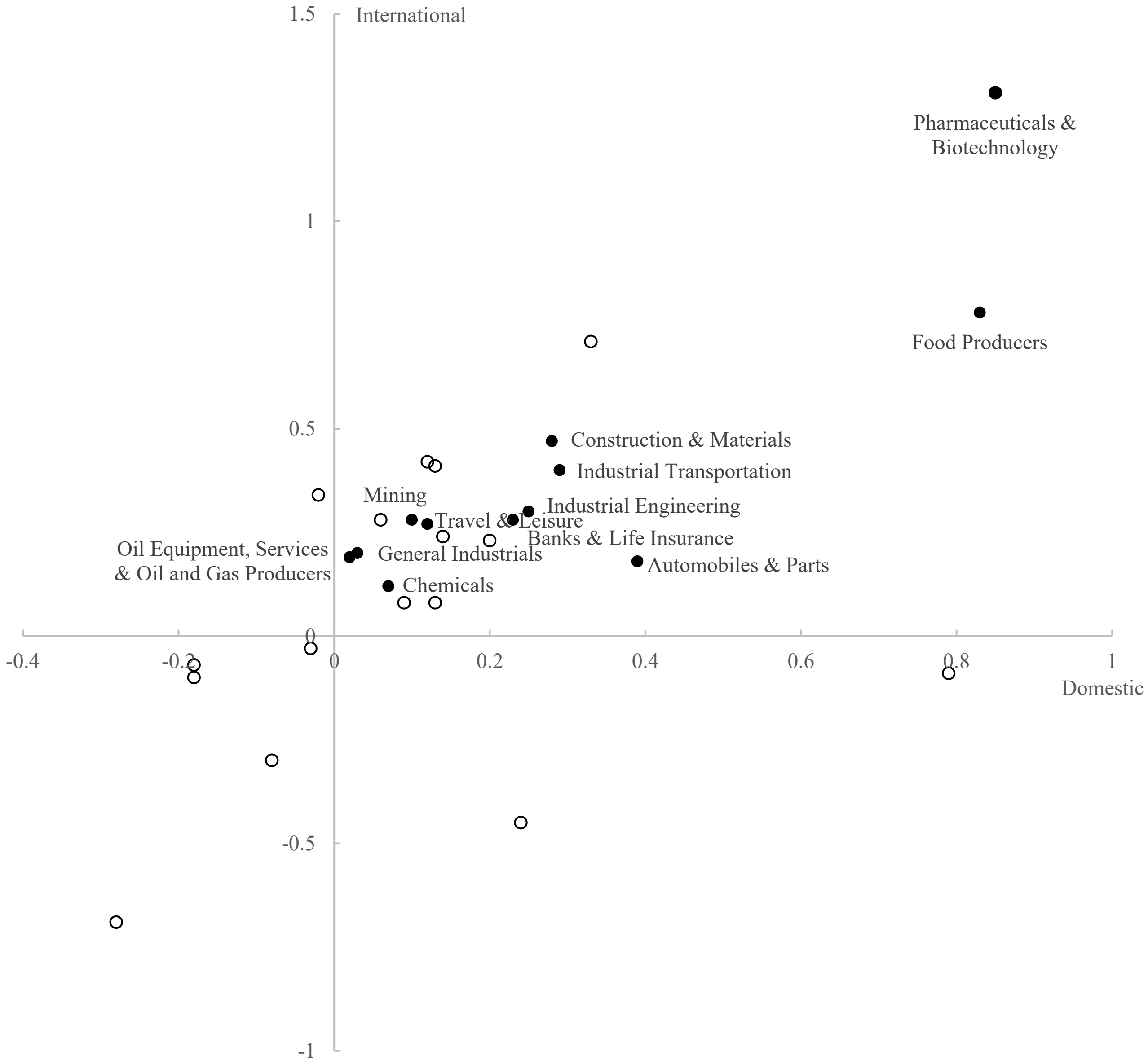


Figure 5. International Accessibility

This figure shows the industry spillover of international accessibility and the impact of IPO waves on international accessibility. Panel A shows the slope coefficients from portfolio-specific time-series regressions of earnings yield on a constant and industry international accessibility (IA2). The vertical (horizontal) axis shows the slope coefficient for a portfolio of international (domestic) firms within an industry. The solid dots represent industries with both coefficients positive and statistically significant at the 10% level and their names are shown next to the dots. Panel B shows the time-series plots of market international accessibility (IA2) with (solid line) and without (dashed line) the firms that conducted their IPOs between 2006Q3 and 2007Q4. Firm level variable construction details are shown in the Appendix C. To obtain market level values, we value weight firm level variables.

Panel A. Industry Spillover



Panel B. International Accessibility with and without IPO Firms

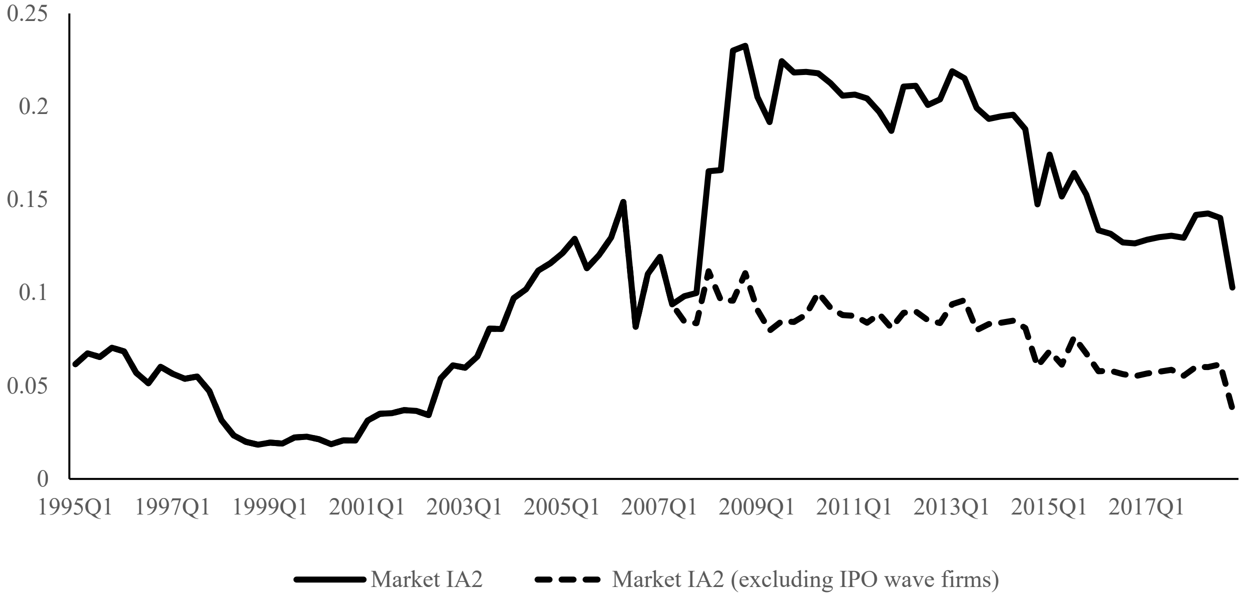
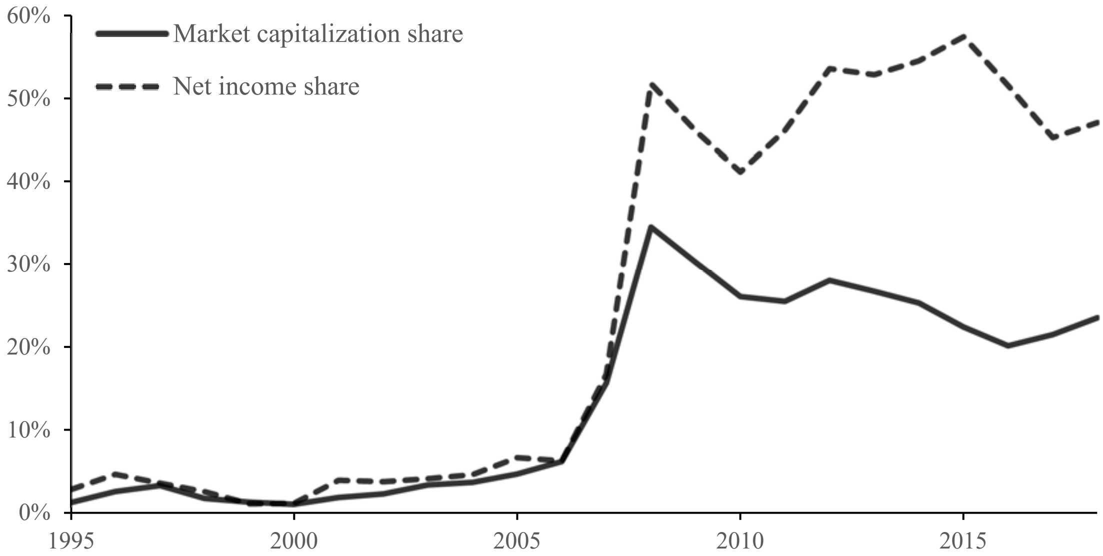


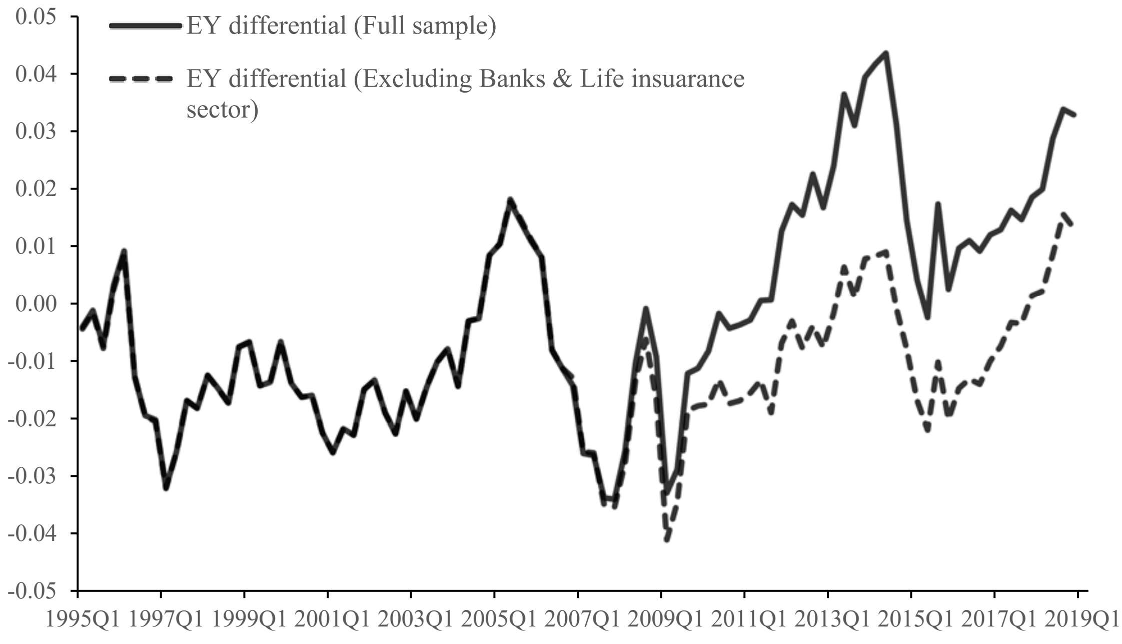
Figure 6. The Role of Banks & Life Insurance Sector

This table presents the market share of the "Banks & Life Insurance" sector and the market earnings yield differential with and without this sector. Panel A shows the time-series of the market share for the "Banks & Life Insurance" sector, both in terms of market capitalization and net income. The solid line shows the market share in terms of market capitalization, calculated as the sum of market capitalization of firms in the "Banks & Life Insurance" sector divided by the total market capitalization of the entire market. The dash line shows the market share in terms of net income, calculated as the sum of net income of firms in the sector divided by the total net income of the entire market. In Panel B, the solid (dash) line shows earnings yield differentials for the whole market with (without) firms in the "Banks & Life Insurance" sector.

Panel A. Market Share of Banks & Life Insurance Sector



Panel B. Earnings Yield Differential with/without Banks & Life Insurance Sector



Appendix

Appendix A. Construction of Characteristics Portfolios

We split the whole sample of firms with available quarterly earnings into portfolios based on their characteristics: state ownership (SO), retail ownership (Retail), institutional ownership (IO), international accessibility (IA, whether a firm has B share, H share or ADR), illiquidity, market capitalization (Size), industries (Tech, whether a firm belongs to TMT industry), and listed boards (Board)

1. State Ownership, International Accessibility and Listing Board Portfolios

We formed 4 state ownership portfolios based on firm-level state ownership (0, 0-10%, 10%-50% and >50%), 2 portfolios based on firm-level IA1 (0, and >0) and 3 listing board portfolio based on the board in which they are listed. Since these three variables are only available for Chinese firms, we only form the portfolios for China. After the SO portfolios, IA portfolios and listing board portfolios are formed, within each portfolio, portfolio-level variables for China are generated. The calculation procedure for China is as follows: among all the variables, portfolio-level leverage, IA1, IA2, state ownership, institutional ownership, retail ownership, standardized number of shareholders, zeros, turnover, number of analysts, and forecast dispersion are calculated as the weighted average of the corresponding firm-level variables, using the lagged firm market value as weights. For other variables (earnings yield, earnings growth volatility, minimum number of stocks, IA3, earnings growth expectation, and sales growth expectation), we calculate their portfolio-level measures directly in the same way that we applied for sectors using individual firm data.

Next, we match each China portfolio with a U.S. portfolio benchmark, which has the same sector composition. Specifically, within each portfolio in China, for each sector j, we sum up the market capitalization of firm i and get the sector-level weight $VW_{j,t}^{CN} = \sum_i VW_{i,j,t}^{CN}$. For the next step, we use the sector-level weight to form the U.S. benchmark as $EY_t^{US} = \sum_j VW_{j,t}^{CN} EY_{j,t}^{US}$. We carry out this procedure for all the variables for the benchmark U.S. portfolios.

2. Retail, IO, Illiquidity, Turnover, Size, and Tech Portfolios

We formed 2 retail portfolio based on quarterly firm-level retail ownership: ≤ Lower country 30% percentile of retail ownership and >= Higher country 30% percentile of retail ownership, 2 IO portfolio based on quarterly institutional ownership: ≤ Lower country 30% percentile of institutional ownership and >= Higher country 30% percentile of institutional ownership, 2 illiquidity portfolios based on quarterly firm-level zeros: ≤ Lower country 30% percentile of zeros and >= Higher country 30% percentile of zeros, 2 turnover portfolio based on quarterly firm-level turnover: ≤ Lower country 30% percentile of turnover rate and >= Higher country 30% percentile of turnover rate, 2 size portfolio based on quarterly firm-level market value: ≤ Lower country 30% percentile of market value and >= Higher country 30% percentile of market value, 2 tech portfolio based on whether firms are in high-tech industry. High-tech industry is defined as "TMT" industry, including "Fixed and Mobile Telecom", "Media", "Software & Computer Services", "Technology Hardware & Equipment" 4 sectors. Since both China and the U.S. have these classifications, we can form these portfolios for both China and the U.S. After these portfolios are formed, within each portfolio, we calculate portfolio-level variables for China and the U.S.

Portfolio-level leverage, IA1, IA2, state ownership, institutional ownership, retail ownership, standardized number of shareholders, zeros, turnover, number of analysts, and forecast dispersion are calculated as the weighted average of the corresponding firm-level variables, using the lagged firm market value as weights. For other variables (earnings yield, earnings growth volatility, minimum number of stocks, IA3, earnings growth expectation, and sales growth expectation), we calculate their portfolio-level measures directly in the same way that we applied for sectors using individual firm data.

Next, we match each China portfolio with a U.S. portfolio benchmark, which has the same "within-portfolio" sector composition. Specifically, within each portfolio in China, for each sector j, we sum up the market capitalization of firm i and get the sector-level weight $VW_{j,t}^{CN} = \sum_i VW_{i,j,t}^{CN}$. For the next step, we use the sector-level weight to form the U.S. benchmark as $EY_t^{US} = \sum_j VW_{j,t}^{CN} EY_{j,t}^{US}$, where $EY_{j,t}^{US}$ is the "within-portfolio" sector-level earnings yield. We carry out this procedure for all the variables for the benchmark U.S. portfolios.

Appendix B. Bai, Lumsdaine and Stock (1998) Structural Break Test

Considering the following specification[1]:

$$y_t = (G_t' \otimes I_n)\theta + d_t(k)(G_t' \otimes I_n)S'S\delta + \varepsilon_t \tag{1}$$

where y_t is n×1, G_t' is a row vector containing a constant, lags of y_t, and row t of the matrix of exogenous regressors X, I_n is a n×n identity matrix. $d_t(k) = 0$ for t < k and $d_t(k) = 1$ for t ≥ k, and $\sum$ is the covariance matrix of error term ε_t. θ and δ are parameter vectors with dimension r. S is a selection matrix containing zero and ones. It is used to identify (via the placement of the ones) which of the r parameters are allowed to change in the regression. For our case, we consider two specifications: $\mathrm{S} = \mathrm{s} \otimes I_n, s = (1,0, \dots 0)$ when only the intercept is allowed to break and $\mathrm{S} = I_r$ (all parameters break).

We can write the system more compactly as

$$y_t = Z_t'(k)\beta + \varepsilon_t \tag{2}$$

where $Z_t'(k) = \left(\left((G_t' \otimes I_n), d_t(k)(G_t' \otimes I_n)S'\right)\right.$ and $\beta = (\theta', (S\delta)')'$. If we let $\mathrm{R} = (0, \mathrm{I})$ be the selection matrix associated with β, then $\mathrm{R}\beta = \mathrm{S}\delta$ and the F-statistic testing $\mathrm{S}\delta = 0$ is

$$\hat{F}_T(k) = T\{R\hat{\beta}(k)\}' \left\{R\left(T^{-1}\sum_{t=1}^{T} Z_t \hat{\Sigma}_k^{-1} Z_t'\right)^{-1} R'\right\}^{-1} \{R\hat{\beta}(k)\} \tag{3}$$

where $\hat{\beta}(k)$ and $\hat{\sum}_k$ denote the estimators of β and $\sum$, respectively, evaluated at $\hat{k}$. We here focus on max $\hat{F}_T(k)$, which is our sup-Wald statistics.

Bai, Lumsdaine and Stock (1998) show that the confidence interval is as follows:

$$\hat{k} \pm \alpha_{\frac{\pi}{2}}[(S\hat{\delta}_T)'S(\hat{Q} \otimes \hat{\Sigma}_k^{-1})S'(S\hat{\delta}_T)^{-1}]$$

where $\hat{Q} = (1/T)\sum_{t=1}^{T} G_t G_t'$

[1] These notations are largely based on Bai, Lumsdaine and Stock (1998) and Bekaert, Harvey and Lumsdaine (2002). See their papers for more details.

In our case, to find the break in the valuation differential, we estimate the following specification[2]:

$$DIFEY_t = \delta + \theta Break_t + \sum_{j=1}^{n} \rho_j DIFEY_{t-j} + \varepsilon_t, \qquad (4)$$

where $DIFEY$ is the earnings yield differential between China and the U.S., $Break$ is a dummy variable equal to one (zero) after (before) the break date detected by the methodology, and ε is the error term. The optimal length *n* for the AR process is selected by the BIC criterion, and we always find a first-order process to be optimal.

[2] Following Bekaert, Harvey and Lumsdaine (2002), we also try another specification which in addition allows the lag terms to break. The break dates are robust.

Appendix C: Valuation Model

We first consider a simple dynamic valuation problem where discount rates vary through time and there is a persistent, time-varying cash flow component. For now, we omit all indictors of country or industry, simply focusing on the pricing problem.

For simplicity, we assume a constant pay-out ratio and, without loss of generality, set the payout ratio equal to 1. Let D_t represent dividends and let δ_t be the continuously compounded discount rate. The log-dividend growth rate (which equals earnings growth under our assumptions) is denoted by Δd_t.

Under these assumptions, as is also shown in BHLS (2011), rational pricing implies:

$$P_t = E_t\left[\sum_{i=1}^{\infty} \exp\,(-\sum_{j=1}^{i} \delta_{t+j-1}) D_{t+i}\right] \tag{1}$$

$$PE_t \equiv \frac{P_t}{D_t} \equiv E_t[\sum_{i=1}^{\infty} exp\,\{\sum_{j=1}^{i}(-\delta_{t+j-1} + \Delta d_{t+j})\}] \tag{2}$$

We can rewrite (2) as

$$\frac{P_t}{D_t} \equiv \sum_{i=1}^{\infty} V_{t,i} \tag{3}$$

where $V_{t,i} \equiv E_t[\exp\,\{\sum_{j=1}^{i}(-\delta_{t+j-1} + \Delta d_{t+j})\}]$

To solve (3), we assume the following data generating processes for Δd_t and δ_t:

$$\begin{gathered} \Delta d_t = cf_{t-1} + \epsilon_t^d \\ cf_t = \mu_{cf}\left(1 - \phi_{cf}\right) + \phi_{cf} cf_{t-1} + \sigma_{cf} \varepsilon_t^d \\ \delta_t = \mu_\delta(1 - \phi_\delta) + \phi_\delta \delta_{t-1} + \varepsilon_t^\delta \end{gathered} \tag{4}$$

where cf indicates cash flow; μ_{cf}, ϕ_{cf}, σ_{cf}, μ_δ and ϕ_δ are constants.

That is, the discount rate and the persistent component of cash flows follow simple autoregressive processes. There are only two shocks ε_t^d and ε_t^δ, which are assumed $\mathrm{N}(0, \sigma_k^2)$ for $\mathrm{k} = \mathrm{d}, \delta$. Allowing an additional shock for expected cash flows, cf, is a straightforward extension.

Under these assumptions, the solution to (3) is of the form:

$$\frac{P_t}{D_t} = \sum_{i=1}^{\infty} \exp\left(a_i + b_i cf_t + c_i \delta_t\right) \tag{5}$$

where $a_0 = b_0 = c_0 = 0$ and a_i, b_i, c_i follow difference equations that are easily derived, First note that:

$$V_{t,1} = E_t[\exp(-\delta_t + \Delta d_{t+1})] = \exp\left(\frac{1}{2}\sigma_d^2 + cf_t - \delta_t\right)$$

So that, $a_1 = \frac{1}{2}\sigma_d^2$, $b_1 = 1$ and $c_1 = -1$.

The induction step of the proof starts from (5), to deduce:

$$V_{t,i+1} = E_t[\exp(-\delta_t + \Delta d_{t+1})\, V_{t+1.i}] = \exp(a_{i+1} + b_{i+1} cf_t + c_{i+1}\delta_t) \tag{6}$$

Solving the first part of (6) delivers the required solutions for a_i, b_i, c_i:

$$\begin{aligned} a_{i+1} &= a_i + b_i \mu_{cf}\left(1 - \phi_{cf}\right) + c_i \mu_f (1 - \phi_\delta) \\ &\quad + \frac{1}{2}\sigma_d^2\left(1 + b_i^2 \sigma_{cf}^2\right) + \frac{c_i^2}{2}\sigma_\delta^2 \\ b_{i+1} &= \phi_{cf} b_i + 1 = \frac{1 - \phi_{cf}^{i+1}}{1 - \phi_{cf}} \\ c_{i+1} &= \phi_\delta c_i - 1 = -\frac{1 - \phi_\delta^{i+1}}{1 - \phi_\delta} \end{aligned} \tag{7}$$

Thus, the b_{i+1} and c_{i+1} coefficients have simple closed-form solutions; a_{i+1} does too, but the expression is complicated.

Our empirical specification uses earnings yields, the reciprocal of the price earnings ratio. Given the closed form solution in (3), a tractable expression for the earnings yield can be derived using a first-order Taylor series approximation around $(\overline{cf}, \bar{\delta})$:

$$EY_t = \frac{1}{PE_t} = f(cf_t, \delta_t) = \mu_{ey} + \gamma_{cf} cf_t + \kappa_\delta \delta_t \tag{8}$$

where
$$\mu_{ey} = \overline{EY} + \overline{EY}^2 * \left(\overline{PE}_{cf} + \overline{PE}_\delta\right)$$

$$\gamma_{cf} = -\overline{EY}^2 * \overline{PE}_{cf}$$

$$\kappa_\delta = -\overline{EY}^2 * \overline{PE}_\delta$$

$$\overline{EY} = \frac{1}{\sum_{i=1}^{\infty} \exp\left(a_i + b_i \overline{cf} + c_i \bar{\delta}\right)}$$

$$\overline{PE}_{cf} = \sum_{i=1}^{\infty} b_i \exp\left(a_i + b_i \overline{cf} + c_i \bar{\delta}\right)$$

$$\overline{PE}_\delta = \sum_{i=1}^{\infty} c_i \exp\left(a_i + b_i \overline{cf} + c_i \bar{\delta}\right)$$

Note that because $b_i > 0, c_i < 0$, it is always true that $\gamma_{cf} < 0$ and $\kappa_\delta > 0$, consistent with intuition.

While Equation (8) seems already close to our empirical model, it really is not tractable empirically because all coefficients depend on both country c, and portfolio j. We consider two different settings to go from Equation (8) to our empirical model.

(1) *A country-specific "long-run" model*

We assume that all parameters are country but not portfolio specific. Each portfolio has its own shocks but the parameters are "long-run" country parameters. Performing the linearization around country averages then turns Equation (8) into our Chinese earnings yield model:

$$EY_{j,t}^{CN} = \mu^{CN} + \gamma_{cf}^{CN} cf_{j,t} + \kappa_{\delta}^{CN} \delta_{j,t} \quad (9)$$

and our various explanatory variables capture portfolio-dependent short-term valuation in $cf_{j,t}$ and $\delta_{j,t}$, using linear functions of observable information. For example, the cash flow function admits a country-wide (GDP growth) and portfolio specific earnings and sales growth expectation information. For discount rates, we model the deviations from the country average to be a function of the ownership structure but also accommodate market-wide time series variation as a function of reforms and cross-sectional differences as a function of liquidity. Of course, even if the parameter restrictions do not hold literally and γ_{cf}^{CN} and k_{δ}^{CN} are still dependent on portfolio characteristics through the dependence of a_i, b_i and c_i on portfolio-specific parameters, our panel model with multiple explanatory variable may provide a good approximation of either $\gamma_{j,cf}^{CN} cf_{j,t}$ and $\kappa_{j,\delta}^{CN} \delta_{j,t}$.

(2) *A model under integration*

A setup under the null of economic and financial integration as in BHLS (2011) is also embedded in this framework. BHLS assume:

$$\begin{gathered} cf_{j,t}^{c} = cf_{w,t}; \\ \delta_{j,t}^{c} = (1 - \beta_j) r_f + \beta_j \delta_{w,t} \end{gathered} \quad (10)$$

However, the dividend shock is still portfolio (industry) specific. They are interested in removing the dependence on the country so that they can look at differences in valuation ratios.

Under their assumptions, the b_i and c_i coefficients in Equation (5) only depend on the industry, but not the country. However, the a_i coefficient should still feature country-specific information, as long as earnings volatility is not equalized across countries. This is the reason why earnings growth volatility differentials are used as a control in their empirical specifications, a practice we adopt as well. This model provides a justification of formulating the model in differences, but note that under the linearization, the a_i term still appears in the γ_{cf} and κ_δ coefficients. Implicitly, its variation must be captured by some of our explanatory variables or its effect on those coefficients must be second order (which is likely given that they involve innovation variances)

In addition, the model must then recognize variation in β's across portfolios. While some of our explanatory variables may well capture such variation, Section V.B considers such variation explicitly.

In fact, we go one step further and argue that the integrated model in (9) is unlikely to work for China and also consider (partial) segmentation models where:

$$\delta_{j,t}^{CN} \equiv \left(1-\beta_j^{CN}\right) r_f + \beta_j^{CN} \delta_t^{CN} + \beta_j^{US} \bar{\delta}_t^{US} \tag{11}$$

where $\bar{\delta}_t^{US}$ is the U.S. risk premium. We consider fully segmented models, where $\beta_j^{US} \equiv 0$, and models where the β's vary through time. If true, allowing for these additional sources of cross-sectional and time-series variation should improve model fit.

Appendix D. Variable Descriptions

Variable	Description
Earnings yield differential (DIFEY) (Sector/portfolio level, 95-18)	$DIFEY_{j,t} = EY_{j,t}^{CN} - EY_{j,t}^{US}$ This variable measures the sector level earnings yield differentials between China and the U.S. In each country, sector valuation EY is the sum of earnings across all firms in the sector over sector market capitalization. Earnings at quarter t is calculated as the trailing annualized net income by summing up net income from quarter t-4 to quarter t-1. Negative earnings are set to be 0 before aggregating into the sector level. Because Chinese firms only reported semi-annual reports before 2002 and they reported accumulated net income in their semi-annual reports, for missing quarterly earnings data before 2002, we assume that earnings in the first and second quarter of the year are one half of the earnings reported in the semi-annual reports and the earnings for the third and fourth quarter are one half of the total earnings generated in the second half of the year, which is the difference between the earnings reported in the firm's annual reports and that in the semi-annual reports. Frequency: Quarterly.
Control Variables	
Leverage (Sector/portfolio level, 95-18)	For non-financial firms, sector level leverage is calculated as the value-weighted (using last-quarter market cap as weight) ratios of long-term debt plus short-term debt over total assets. For China, the direct items measuring short-term debt and long-term debt are not available, so we add up four items: short-term borrowing, long term borrowing, debt due in future one-year and bond payable to measure total debt. For financial firms, sector level leverage is calculated as the value-weighted (using last-quarter market cap as weight) ratios of total liability over total assets. We winsorize leverage at the 1 and 99 percentiles. Frequency: Quarterly. Sources: WIND and COMPUSTAT.
Earnings growth volatility (Sector/portfolio level, 95-18)	To compute earnings growth volatility, we first calculate annualized firm level net income. Annualized firm level net income at quarter t is calculated by summing up firm level net income from quarter t-4 to quarter t-1. In each quarter, we then compute sector level annualized net income (NI) by adding up firm level annualized net income within each sector. The sector earnings growth at quarter t is calculated as $\log\left(\frac{NI_t * CPI_{t-4}}{NI_{t-4} * CPI_t}\right)$. We calculate the volatility of sector NI growth each quarter by calculating the standard deviation of the log growth rate over the past twenty quarters. For the 10th -19th observation of each sector, we use all available observations to calculate the standard deviation. For the first 10 observations, we use the standard deviation computed for the 10th observation. Frequency: Quarterly. Source: WIND and COMPUSTAT.

Variable	Description
Minimum number of stocks (Sector/portfolio level, 95-18)	Natural logarithm of the minimum of the number of stocks in each sector of China and the U.S. Frequency: Quarterly.
Financial Openness	
REGOPEN (Market level, 95-18)	REGOPEN is only available for China. Based on the major events listed in Appendix E, this cumulative regulation dummy variable is constructed as follows: take the value of 0 from 1995Q1 to 2000Q4, the value of 1 from 2001Q1 to 2002Q3 (Bshares), the value of 1.5 from 2002Q4 to 2003Q2 (the announcement of QFII), the value of 2 from 2003Q3 to 2006Q1 (the first transaction by QFII), the value of 2.5 in 2006Q2 (the announcement of QDII), the value of 3 from 2006Q3 to 2011Q3 (market execution of QDII), the value of 4 from 2011Q4 to 2014Q1 (the announcement and market execution of RQFII), the value of 4.67 from 2014Q2 to 2014Q3 (the announcement and regulation execution of Shanghai-Hong Kong Connect), the value of 5 from 2014Q4 to 2016Q2 (the official start of Shanghai-Hong Kong Connect), the value of 5.67 in 2016Q3 (the announcement and regulation execution of Shenzhen-Hong Kong Connect), the value of 6 from 2016Q4 to 2017Q1 (the official start of Shenzhen-Hong Kong Connect), the value of 6.67 from 2017Q2 to 2018Q1 (the announcement of incorporating A share into MSCI index), and the value of 7 from 2018Q2-2018Q4 (233 stocks listed in A-share market was officially incorporated MSCI emerging markets index and MSCI All Country World Index). Frequency: Quarterly. Source: constructed by authors.
International accessibility (Sector level, 95-18)	International accessibility variables are only available for China. There are two IA variables at the firm level. Firm IA1 is calculated by adding up four dummy variables, Bshare, Hshare, ADR and CHconnect. Variable Bshare (Hshare, ADR) is equal to 1 if the stock has B shares (H shares, ADR) issued. Variable CHconnect is equal to 1 if the stock is included in the China-HK connect program. Firm IA1 takes the minimum value of 0, meaning the stocks have no international accessibility, and takes the maximum of 4, which indicates that the stocks have B shares, H shares, ADRs and are incorporate into the China-HK connect program. Firm IA2 is the ratio of market capitalization of B shares, H shares and ADRs to firm total market capitalization. There are three sector-level IA variables. Sector IA1 (IA2) is the weighted average of the firm-level IA1 (IA2) within the sector, using the firm market capitalization of last quarter as weight. Sector IA3 is the market share of firms with positive firm-level IA1 within the sector. Frequency: Quarterly. Source: calculated by author using data from WIND.

Variable	Description
Real Interest rate (Market level, 95-18)	The difference between the real interest rate between China and the U.S. For the nominal interest rate in China, we use the 1-year institution and individual deposit rate, obtained from People's Bank of China. For US, we use the 1-year Treasury constant maturity Rate from FRED Economic Data. The real interest rate is calculated by subtracting inflation from the nominal interest rate. The inflation rate is calculated as the percentage change of quarterly CPI over the same quarter in the previous year. Inflation rate(t) = CPI(t)/CPI(t-4)-1. We obtained the quarterly CPI data from China National Bureau of Statistics and US Bureau of Labor Statistics. Frequency: Quarterly. Source: People's Bank of China, China National Bureau of Statistics, FRED and US Bureau of Labor Statistics.
Overall political rating (Market level, 95-18)	The sum of all 12 ICRG subcomponents, with a total score of 100 and the maximum score for each subcomponent displayed in parenthesis: Government Stability (12), Socioeconomic Conditions (12), Investment Profile (12), Internal Conflict (12), External Conflict (12), Corruption (6), Military in Politics (6), Religious Tensions (6), Law and Order (6), Ethnic Tensions (6), Democratic Accountability (6), and Bureaucracy Quality (4). ICRG currently only provides data till 2018Q3. We fill in the 2018Q4 numbers making an assumption that they are equal to that in 2018Q3. Frequency: Annual. Source: ICRG.
Ownership Structure	
State ownership (Sector/portfolio level, 95-18)	State ownership is only available for China. It is measured as the fraction of total shares that are owned by the state. It is calculated in three steps. First, from CSMAR, we collect information on the ten largest shareholders (including their numbers of holding shares, the nature of the shares) and use this information to calculate a measure of state ownership for a given company. Second, since the financial statement discloses how many shares are state-owned shares among the non-tradable shares, we use this information to calculate another version of state ownership by only taking the non-tradable shares into account. Then, we take the larger value of the first and second measure to proxy for the state ownership of a given firm. This variable is only available for China. Frequency: Quarterly. Source: CSMAR.
Institutional ownership (Sector/portfolio level, 03-18)	Institutional ownership is measured as fraction of tradable shares that are owned by institutions. For China, we use institutional holding data from WIND to calculate institutional ownership. Institutions are defined as professional money managers including mutual funds, QFII, insurance companies, banks, hedge funds, investment trust companies, pension funds and security company and wealth management products of security company. Notice that only mutual funds and wealth products of security companies have

Variable	Description
Institutional ownership (Sector/portfolio level, 03-18)	an obligation to report their holdings in China. For other types of institutions, WIND can only collect institutional holdings from ten largest tradable shareholders information disclosed in firms' quarterly financial statements, so we can only use the ten largest tradable shareholders holding information, which might underestimate the institutional ownership. For the U.S., following Ferreira, Miguel and Matos (2008), we use FactSet data to calculate institutional holdings. Specifically, institutional holding is calculated as the market value of the sum of 13f holdings and non-13f fund holdings, divided by total market value. Frequency: Quarterly. Source: WIND, FactSet LionShares
Retail ownership (Sector/portfolio level, 03-18)	Retail ownership is measured as the fraction of tradable shares that are owned by the retail investors. For China, retail investor ownership is defined as follows: 1 - institutional ownership - state ownership - insider ownership. State ownership here is the fraction of tradable shares that are owned by the state, which differs from our Chinese state ownership measure mentioned above. Institutional ownership is fraction of tradable share owned by institutional investors, which is defined above. Insiders are defined as directors, supervisors or managers in a company, or large individual shareholders who show up in the firms' ten largest shareholders' profile. We obtain insiders information from CSMAR. For the U.S., retail investor ownership is defined as follows: 1 - institutional ownership - insider ownership. Insider information for U.S. is extracted from Thomson Reuters. Frequency: Quarterly. Source: WIND, FactSet LionShares, CSMAR, Thomson Reuters
Standardized number of shareholders (Sector/portfolio level, 03-18)	Standardized number of shareholders is only available for China. It is calculated as number of shareholders divided by total tradable shares and multiplied by 1000. Frequency: Quarterly. Source: WIND
Financial Development	
REGDEV (Market level, 95-18)	REGDEV is only available for China. Based on the major events listed in Appendix E, this cumulative regulation dummy variable is constructed as follows: take the value of 0 from 1995Q1 to 2005Q1, the value of 1 from 2005Q2 to 2008Q3 (the Split-share Reform), the value of 1.5 from 2008Q4 to 2009Q4 (the announcement of the Margin Trading and Short-selling Program), the value of 2 from 2010Q4 to 2015Q3 (the official start of the Margin Trading and Short-selling Program), the value of 2.5 from 2015Q4 to 2018Q4 (The Standing Committee of the People's Congress authorize the central government to apply a registration-based initial public offering (IPO) system Frequency: Quarterly. Source: constructed by authors.

Variable	Description
Number of public firms (Market level, 95-18)	The log of the number of publicly traded firms at the end of each quarter in a given country. Frequency: Quarterly. Source: World Bank
Adjusted market development (Market level, 95-18)	Let MC_t^{CN} be the stock market capitalization of China relative to GDP and MC_t^{US} be the stock market capitalization of the US relative to GDP. Let $MCRatio_t = MC_t^{CN}/MC_t^{US}$ .Then, standardize $MCRatio_t$ by subtracting its mean and divided by standard deviation over 1995Q1 to 2018Q4. For the next step, take one year past cumulative market return in China and divide by one-year cumulative market return in the US. This ratio is denoted as $RetRatio_t$. This variable measures recent trends in returns. Then, standardize $RetRatio_t$ by subtracting its mean and divided by standard deviation over 1995Q1 to 2018Q4. The Adjusted market development is the difference between standardized $MCRatio_t$ and standardized $RetRatio_t$. Frequency: Quarterly. Sources: WIND and CRSP.
Zeros (Sector/portfolio level, 95-18)	Following Bekaert, Harvey, and Lundblad (2007), we calculate zeros as the proportion of zero daily returns observed over the relevant quarter for each security. For each sector/portfolio in each quarter, we calculate the market capitalization-weighted (using the market cap from last quarter) proportion of zero daily returns across all firms. Frequency: Quarterly. Source: WIND and CRSP
Turnover (Sector/portfolio level, 95-18)	We first calculate firm-level turnover rate as the ratio of market value traded to total tradable shares market capitalization in each quarter. For sector-level and market-level, we take the value-weighted average of all the firms in the sector and in the market. Frequency: Quarterly. Source: WIND, CRSP
Growth Prospects	
GDP growth rate (Market level, 95-18)	GDP growth rate in quarter t is calculated as [GDP(t)+GDP(t-1)+GDP(t-2)+GDP(t-3)]/[GDP(t-4)+GDP(t-4)+GDP(t-6)+GDP(t-7)]-1, using quarterly real GDP. Frequency: Quarterly. Source: China National Bureau of Statistics and Bureau of Economic Analysis.
Earnings growth expectation (Sector/portfolio level, 03-18)	We use analysts' earnings forecasts to calculate earnings growth expectation. Earnings growth expectation is calculated as the weighted average of annualized earnings growth rate expectation in the next 3 years using the most recent comparable CPI growth rates as deflators. Specifically, in each quarter, we first sum up the median firm-level earnings forecasts while setting those negative values to zeros within each sector. Then, for most nearby fiscal year t, the real earnings growth rate expectation EG_t is calculated as (median analyst earnings forecast for fiscal year t * CPI_{t-2}) / (actual earnings in fiscal year t-1* CPI_{t-1}) - 1. For fiscal year t+1, the earnings growth rate expectation EG_{t+1} is defined as [(median

Variable	Description
	analyst earnings forecast for fiscal year t+1 $^*CPI_{t-3}$/actual earnings in fiscal year t-1 $^*CPI_{t-1}$) - 1]/2. For fiscal year t+2, the earnings growth rate expectation EG_{t+1} is defined as [(median analyst sales forecast for fiscal year t+2 $^*CPI_{t-4}$ / actual earnings in fiscal year t-1 $^*CPI_{t-1}$) - 1]/3. We use the number of quarters that actually have to be predicted as weight. In the first quarter of every year, the weighted earnings growth expectation is defined as $4/12^*EG_t$ + $4/12^*EG_{t+1}$ + $4/12^*EG_{t+2}$. For the second quarter, it is $3/11^*EG_t$ + $4/11^*EG_{t+1}$ + $4/11^*EG_t$, $2/10^*EG_t$ + $4/10^*EG_{t+1}$ + $4/10^*EG_t$ for the third quarter and $1/9^*EG_t$ + $4/9^*EG_{t+1}$ + $4/9^*EG_t$ for the fourth quarter. For China, we obtain the analyst forecast data from CSMAR and supplement it using analyst data from Suntime. Suntime have more forecast records for a given firm than CSMAR, but it has a shorter sample started from 2006. For sample 2003-2005, we use CSMAR data. For sample 2006-2018, we use Suntime data. U.S. analyst data is collected from I/B/E/S. Frequency: Quarterly. Source: CSMAR, Suntime and I/B/E/S.
Sales growth expectation (Sector/portfolio level, 03-18)	We use analysts' sales forecasts to calculate the sales growth expectation. Sales growth expectation is calculated as the weighted average of the annualized sales growth rate expectation in the next 3 years using the most recent comparable CPI growth rates as deflators. Specifically, in each quarter, we first simply sum up the median firm-level sales forecasts within each sector. Then, for the most nearby fiscal year t, the real sales growth rate expectation SG_t is calculated as (median analyst sales forecast for fiscal year t $^*CPI_{t-2}$) / (actual sales in fiscal year t-1 $^*CPI_{t-1}$) - 1. For fiscal year t+1, the sales growth rate expectation SG_{t+1} is defined as [(median analyst sales forecast for fiscal year t+1 $^*CPI_{t-3}$/actual sales in fiscal year t-1 $^*CPI_{t-1}$) - 1]/2. For fiscal year t+2, the sales growth rate expectation SG_{t+1} is defined as [(median analyst sales forecast for fiscal year t+2 $^*CPI_{t-4}$ / actual sales in fiscal year t-1 $^*CPI_{t-1}$) - 1]/3. We use the number of quarters that actually have to be predicted as weight. In the first quarter of every year, the weighted earnings growth expectation is defined as $4/12^*SG_t$ + $4/12^*SG_{t+1}$ + $4/12^*SG_{t+2}$. For the second quarter, it is $3/11^*SG_t$ + $4/11^*SG_{t+1}$ + $4/11^*SG_t$, $2/10^*SG_t$ + $4/10^*SG_{t+1}$ + $4/10^*SG_t$ for the third quarter and $1/9^*SG_t$ + $4/9^*SG_{t+1}$ + $4/9^*SG_t$ for the fourth quarter. For China, we obtain the analyst forecast data from CSMAR and supplement it using analyst data from Suntime. Suntime have more forecast records for a given firm than CSMAR, but it has a shorter sample starting from 2006. For the 2003-2005 sample, we use CSMAR data. For the 2006-2018 sample, we use Suntime data. U.S. analyst data is collected from I/B/E/S. Frequency: Quarterly. Source: CSMAR, Suntime and I/B/E/S.

Variable	Description
Number of analysts (Sector/portfolio level, 03-18)	Number of analysts that reported forecasts for a given firm in each quarter. We take the value-weighted average number of analysts across all firms in each sector to obtain the sector-level measure. For China, we obtain the analyst forecast data from CSMAR and supplement it using analyst data from Suntime. Suntime have more forecast records for a given firm than CSMAR, but it has a shorter sample started from 2006. For sample 2003-2005, we use CSMAR data. For sample 2006-2018, we use Suntime data. U.S. analyst data is collected from I/B/E/S. Frequency: Quarterly. Source: CSMAR, Suntime and I/B/E/S.
Forecast dispersion (Sector/portfolio level, 03-18)	We calculate this measure as the standard deviation of reported EPS forecast for Fiscal year 1 (forecast period indictor, FPI=1) in each quarter, standardized by the absolute value of the average forecast across analysts for a given firm in each quarter. We take the value-weighted forecast dispersion across all firms in each sector to obtain the sector-level measure. We winsorize this variable at the 1 and 99 percentiles. For China, we obtain the analyst forecast data from CSMAR and supplement it using analyst data from Suntime. Suntime have more forecast records for a given firm than CSMAR, but it has a shorter sample started from 2006. For sample 2003-2005, we use CSMAR data. For sample 2006-2018, we use Suntime data. U.S. analyst data is collected from I/B/E/S. Frequency: Quarterly. Source: CSMAR, Suntime and I/B/E/S.

Appendix E. Major Events Related to Stock Market Development and Openness in China

Category	Date	Description	Keywords
Policy	2001.02.21	Citizens in mainland China were permitted to invest in B shares.	B shares
Policy	2002.11.05	People's Bank of China (PBOC) and China Securities Regulatory Commission (CSRC) jointly published "the Administration of Domestic Securities Investments of Qualified Foreign Institutional Investors (QFII) (Trial)", indicating the official start of QFII.	QFII
Market	2003.07.09	The first investment of QFII was generated by UBS.	QFII
Policy	2005.04.29	CSRC announced the official start of the trial run of Split-Share Structure Reform.	Split-Share
Policy	2006.04.13	People's Bank of China announced for the first time that qualified funds and other fund-raising institutions can trade in stocks, bonds and funds and other securities outside of China, indicating the official start of QDII.	QDII
Market	2006.08	The first trial QDII fund, Hua'an International Fund, was established by Hua'an Fund.	QDII
Policy	2008.10.05	CSRC announced that the program of dual margin trading and short selling in stock market would start at some point in the future.	Margin Trading and Short Selling
Market	2010.03.31	It was the first day of margin trading and short selling. Hundreds of transactions went through and the total trading value of margin trading and short selling was about 6.59 million RMB.	Margin Trading and Short Selling
Policy	2011.12.16	CSRC announced "Administration of Domestic Securities Investment by Fund Management Companies and Securities Companies as RMB Qualified Foreign Institutional Investors (RQFII) (Trial)"	RQFII
Market	2014.04.10	Shanghai-Hong Kong Stock Connect was announced to be started in the future.	Shanghai-Hong Kong Connect
Policy	2014.06.13	"Regulations of Shanghai-Hong Kong Stock Connect" was published and executed.	Shanghai -Hong Kong Connect
Policy	2014.11.17	Shanghai-Hong Kong Stock Connect officially started.	Shanghai -Hong Kong Connect

Category	Date	Description	Keywords
Market	2015.12.27	The Standing Committee of the People's Congress authorize the central government to apply a registration-based initial public offering (IPO) system.	Registration-based IPO
Market	2016.08.16	Shenzhen-Hong Kong Stock Connect was announced to be started in the future	Shenzhen -Hong Kong Connect
Policy	2016.08.26	"Regulations of China mainland-Hong Kong Stock Connect" was published and executed.	Shenzhen -Hong Kong Connect
Policy	2016.12.05	Shenzhen-Hong Kong Stock Connect officially started.	Shenzhen -Hong Kong Connect
Market	2017.06.21	A-share was announced to be incorporated into MSCI index in June, 2018	MSCI Index
Market	2018.06.01	233 stocks listed in A-share market was officially incorporated MSCI emerging markets index and MSCI All Country World Index	MSCI Index

Appendix F. Model Selection: General-to-Specific Search Algorithm (PcGets Procedure)

Steps	Significance Level
1 Eliminating collinear variables	0.800
Going variable by variable, if multiple variables are correlated more highly than the cut-off value, we select the one variable that features the highest absolute t statistic in the univariate regression and drop other variables.	
2 Estimate a general model with all variables (M1)	
a. Test significance of individual coefficient estimates: t-test. If all coefficients are individually significant, M1 is the final model.	0.025
b. Test M1 against the null of "all coefficients are zero" and the null of "all coefficients but intercept are zero": F-test. If the null is not rejected, M1 is the final model.	0.500
3 Pre-search tests	
a. Top-down tests. Rank the p-values of all coefficients in M1 from largest to smallest. Test joint significance of expanding list of coefficient estimates from largest p-value (least significant) to smallest p-value (most significant): F-test. If F-test is not rejected, remove variables on the current list. (M2)	0.500
b. Repeat top-down tests. Estimate M2 and rank the p-values of all coefficients from largest to smallest. Test joint significance of expanding list of coefficient estimates from largest p-value (least significant) to smallest p-value (most significant): F-test. If F-test is not rejected, remove variables on the current list. (M3)	0.250
c. Bottom-up tests. Rank the p-values of all coefficients in M3 from smallest to largest. Test joint significance of decreasing list of coefficient estimates from smallest p-value (most significant) to largest p-value (least significant): F-test. If F-test is not rejected, remove variables on the current list. (M4)	0.025
4 Multiple-path tests	
a. Estimate M4. If all estimates are individually significant, M4 is the final model.	0.025
b. Otherwise, initiate search paths. Remove blocks of variables with increasing p-values of t-statistics and reestimate the model: remove one insignificant variable each time until all insignificant variables are removed and commenced a path.	
c. Repeat step 3b as long as insignificant variables survive.	0.025
d. The algorithm arrives to a terminal model if all coefficients are individually significant: t-test (M5)	0.025

Appendix G. Extensions to the China EY Model

We consider different CAPM models for discount rates. Our main empirical model uses a set of variables to capture portfolio specific variation in discount rates but these variables may not perfectly capture cross-sectional and time-series variation in exposures to systematic risk factors. Here, we view them as capturing variation in aggregate Chinese discount rates, which differ across portfolios because different investors (e.g. retail versus foreign investors) might use different discount rates. Meanwhile, different industries may also exhibit different exposures to these aggregate discount rates.

We start with a local CAPM models; that is, only the Chinese stock market beta, β_{jt}^{CH}, matters for the portfolio level discount rate, $\delta_{j,t}$, as follows:

$$\delta_{j,t} = r_f + \beta_{j,t}^{CN}\left(\delta_{MKT,t}^{CN} - r_f\right) \tag{12}$$

Here we use the one-year deposit rate from People's Bank of China as the risk-free rate, r_f, and the market capitalization weighted returns of all A-shares to proxy for the Chinese market returns. In the second model, we assume that both global and local market factors affect portfolio returns in China, verifying whether the inclusion of global factors increases the model's explanatory power for explaining the China EY. Assume the U.S. market is a proxy for the global market, then both local and U.S. market betas are relevant for the portfolio j's discount rate, as follows:

$$\delta_{j,t} = r_f + \beta_{j,t}^{CN}\left(\delta_{MKT,t}^{CN} - r_f\right) + \beta_{j,t}^{US}\left(\delta_{MKT,t}^{US} - r_f\right). \tag{13}$$

To estimate the market betas above, $\beta_{j,t}^{CN}$ and $\beta_{j,t}^{US}$, we consider two cases. In the first case, we estimate the betas using weekly returns on both individual portfolios and market portfolios from the whole sample. As a result, each portfolio has one constant beta estimate, the "unconditional beta", which accommodates cross-sectional but not time-series variation. In the second case, we estimate the betas using weekly returns on both individual portfolios and market portfolios from the previous 52 weeks. The rolling estimation generate time-series betas for each portfolio, the "conditional betas", which allows both cross-sectional and time-series variation. Interestingly, while showing some time variation, the conditional betas of Chinese portfolios on the Chinese stock market return are invariably close to 1.0. This indicates that beta variation may be a second-order effect in our analysis.

To understand how the different specifications of betas affect the dynamics of the Chinese EY, we first multiply our original portfolio level discount rate variables, as introduced in Section IV, by the betas, except for the real interest rate which is multiplied by one minus beta (as implied by Equation (12)). The last term in Equation (13) also involves the U.S. discount rate times the portfolio specific beta with respect to the U.S. market. Instead of directly using the U.S. market return, we assume the U.S. discount rate is constant over time and simply add the U.S. beta to the regression, so that the regression coefficient essentially provides an estimate of the U.S. discount rate[3]. We then re-estimate the China EY Model and conduct the PcGets procedure.

[3] We also use two additional methods to measure U.S. discount rates. First, we extract an estimate for the U.S. discount rate from the U.S. earnings yield. In particular, we add the aggregate analysts expected earnings growth to the U.S. earnings yield, and use their sum as a measure of U.S. discount rate. This is inspired by the approach in Ferreira and Santa Clara (2011). Second, we use the discount rate provided by Martin (2017). We use his measure with a horizon of 12-months plus the U.S. risk-free rate as a measure of the U.S. discount rate. For these two measures, we simply add a series multiplying the portfolio specific U.S. betas with the U.S. discount rate to the regression. The results are robust.

Online Appendix

Table OA1. Summary Statistics of Price-to-Earnings (PE) Ratio

This table reports the time-series average of price-to-earnings (PE) ratio for the market and industry portfolios in China and the U.S., from 1995Q1 to 2018Q4. The Chinese sample covers all firms that listed in the A share market. The U.S. sample includes all common stocks listed in New York Stock Exchange, NASDAQ and American Stock Exchange. All variables are constructed on a quarterly basis. We calculated earnings at quarter t as the annualized net income by summing up net income from quarter t-4 to quarter t-1. MV is calculated as the sum of all stocks' price multiplied by common shares outstanding, converted to U.S. dollars using the quarter-end exchange rate. PE ratio is market value for common equity divided by total net income. Negative values of firm earnings are set to zero before being aggregated into sector level.

	China	U.S.
Market	25.9	20.3
Industry Portfolios		
Aerospace & Defense	64.7	19.4
Alternative Energy	68.5	54.9
Automobiles & Parts	26.9	20.7
Banks & Life Insurance	22.2	14.6
Beverages	33.6	25.0
Chemicals	36.8	18.6
Construction & Materials	32.0	20.5
Electricity	22.6	16.3
Electronic & Electrical Equipment	45.4	26.8
Financial Services	44.9	16.2
Fixed and Mobile Telecom	59.9	23.3
Food & Drug Retailers	46.7	21.8
Food Producers	39.8	18.7
Forestry & Paper	35.4	56.4
Gas, Water and Multiutilities	39.3	18.2
General Industrials	28.5	22.5
General Retailers	36.8	23.4
Health Care Equipment & Services	82.3	23.4
Household Goods & Home Construction	22.6	18.7
Industrial Engineering	37.0	18.4
Industrial Metals & Mining	52.8	19.6
Industrial Transportation	30.7	17.4
Leisure Goods	45.6	30.8
Media	59.4	29.3
Mining	28.2	44.0
Oil Equipment, Services & Oil and Gas Producers	27.0	20.7
Personal Goods	32.6	21.2
Pharmaceuticals & Biotechnology	37.7	27.0
Real Estate Investment & Services	33.4	36.8
Software & Computer Services	65.8	29.0
Support Services	41.0	25.7
Technology Hardware & Equipment	43.4	30.6
Travel & Leisure	41.9	20.4

Table OA2. Summary Statistics of Earnings Yield (EY) for Characteristic Portfolios

This table reports the time-series average of earnings yield (EY) for each characteristic portfolio for China and the U.S., from 1995Q1 to 2018Q4. The Chinese sample covers all firms that listed in the A share market. The U.S. sample includes all common stocks listed in New York Stock Exchange, NASDAQ and American Stock Exchange. All variables are constructed on a quarterly basis. We calculated earnings at quarter t as the annualized net income by summing up net income from quarter t-4 to quarter t-1. MV is calculated as the sum of all stocks' price multiplied by common shares outstanding, converted to U.S. dollars using the quarter-end exchange rate. EY is total earnings divided by market value for common equity. The detailed description of the portfolio formation is in Appendix A.

	China	U.S.
State-own Portfolio 1 (SO=0)	3.50	5.15
State-own Portfolio 2 (0<SO<=10%)	3.61	5.14
State-own Portfolio 3 (10%<SO<=50%)	4.36	5.42
State-own Portfolio 4 (SO>50%)	5.96	5.92
Retail Portfolio 1(Retail ownership <= country level lower 30%)	7.04	5.16
Retail Portfolio 2(Retail ownership >= country level upper 30%)	3.53	4.60
IO Portfolio 1(IO <= country level lower 30%)	4.46	4.69
IO Portfolio 2(IO > country level upper 30%)	6.22	5.14
International Accessibility Portfolio 1 (IA1=0)	3.65	5.29
International Accessibility Portfolio 2 (IA1>0)	5.91	6.02
Illiquidity Portfolio 1(Zeros <= country level lower 30%)	4.06	5.06
Illiquidity Portfolio 2(Zeros > country level upper 30%)	5.74	5.17
Turnover Portfolio 1(Turnover <= country level lower 30%)	5.69	4.99
Turnover Portfolio 2(Turnover >= country level upper 30%)	3.02	5.15
Size Portfolio 1(Market value <= country level lower 30%)	1.87	4.70
Size Portfolio 2(Market value >= country level upper 30%)	5.60	5.19
Tech portfolio 1(Not TMT industry)	5.06	5.23
Tech portfolio 2(TMT industry)	2.88	4.47
Listing board portfolio 1(Main board)	5.30	5.69
Listing board portfolio 2(SME board)	3.06	5.44
Listing board portfolio 3(ChiNext board)	2.10	5.11

Table OA3. Summary Statistics of Domestic Ownership Variables

This table reports the time-series average of domestic ownership variables. Definitions of all these variables are described in detail in Appendix D. Institutional ownership, retail ownership and standardized numbers of shareholders are available from 2003Q1 to 2018Q4 while Chinese state ownership starts from 1995Q1 to 2018Q4.

	China				U.S.	
	State ownership	Institutional ownership	Retail ownership	Standardized # of shareholders	Institutional ownership	Retail ownership
Industry Portfolios						
Aerospace & Defense	0.519	0.152	0.546	0.147	0.892	0.107
Alternative Energy	0.132	0.100	0.736	0.244	0.646	0.347
Automobiles & Parts	0.441	0.142	0.574	0.128	0.805	0.193
Banks & Life Insurance	0.411	0.205	0.354	0.083	0.750	0.247
Beverages	0.518	0.213	0.430	0.131	0.703	0.294
Chemicals	0.405	0.123	0.634	0.162	0.829	0.170
Construction & Materials	0.407	0.126	0.586	0.159	0.874	0.119
Electricity	0.597	0.115	0.463	0.104	0.772	0.226
Electronic & Electrical Equipment	0.302	0.127	0.657	0.170	0.881	0.108
Financial Services	0.364	0.122	0.586	0.127	0.863	0.128
Fixed and Mobile Telecom	0.576	0.120	0.485	0.110	0.697	0.302
Food & Drug Retailers	0.293	0.186	0.582	0.129	0.888	0.109
Food Producers	0.303	0.143	0.654	0.162	0.771	0.225
Forestry & Paper	0.305	0.097	0.747	0.164	0.957	0.042
Gas, Water and Multiutilities	0.475	0.059	0.632	0.170	0.737	0.259
General Industrials	0.182	0.105	0.722	0.153	0.736	0.263
General Retailers	0.269	0.160	0.641	0.143	0.753	0.222
Health Care Equipment & Services	0.319	0.181	0.654	0.149	0.920	0.077
Household Goods & Home Construction	0.144	0.196	0.643	0.160	0.771	0.225
Industrial Engineering	0.393	0.137	0.611	0.168	0.862	0.129
Industrial Metals & Mining	0.540	0.127	0.556	0.148	0.745	0.252
Industrial Transportation	0.502	0.132	0.505	0.114	0.852	0.144
Leisure Goods	0.331	0.075	0.714	0.202	0.898	0.096
Media	0.363	0.164	0.575	0.150	0.772	0.225
Mining	0.500	0.155	0.507	0.156	0.912	0.084
Oil Equipment, Services & Oil and Gas Producers	0.654	0.139	0.340	0.087	0.777	0.221
Personal Goods	0.264	0.099	0.708	0.170	0.884	0.110
Pharmaceuticals & Biotechnology	0.296	0.181	0.636	0.149	0.823	0.176
Real Estate Investment & Services	0.344	0.136	0.637	0.167	0.848	0.146
Software & Computer Services	0.179	0.139	0.645	0.179	0.795	0.144
Support Services	0.381	0.087	0.662	0.205	0.898	0.091
Technology Hardware & Equipment	0.328	0.148	0.625	0.184	0.824	0.173
Travel & Leisure	0.458	0.139	0.543	0.166	0.831	0.158

	China				U.S.	
	State ownership	Institutional ownership	Retail ownership	Standardized # of shareholders	Institutional ownership	Retail ownership
Characteristic Portfolios						
State-own Portfolio 1 (SO=0)	0.000	0.148	0.745	0.156	0.816	0.176
State-own Portfolio 2 (0<SO<=10%)	0.040	0.135	0.744	0.151	0.812	0.181
State-own Portfolio 3 (10%<SO<=50%)	0.334	0.144	0.565	0.143	0.815	0.179
State-own Portfolio 4 (SO>50%)	0.663	0.162	0.386	0.118	0.802	0.194
Retail Portfolio 1(Retail ownership <= country level lower 30%)	0.560	0.205	0.294	0.095	0.973	0.022
Retail Portfolio 2(Retail ownership > country level upper 30%)	0.252	0.037	0.931	0.208	0.256	0.737
IO Portfolio 1(IO <= country level lower 30%)	0.441	0.004	0.547	0.170	0.246	0.736
IO Portfolio 2(IO > country level upper 30%)	0.416	0.269	0.511	0.114	0.972	0.026
International Accessibility Portfolio 1 (IA1=0)	0.374	0.144	0.633	0.151	0.819	0.175
International Accessibility Portfolio 2 (IA1>0)	0.527	0.179	0.425	0.115	0.799	0.196
Turnover Portfolio 1(Turnover <= country level lower 30%)	0.485	0.163	0.443	0.115	0.491	0.493
Turnover Portfolio 2(Turnover > country level upper 30%)	0.356	0.106	0.719	0.209	0.903	0.092
Illiquidity Portfolio 1(Zeros <= country level lower 30%)	0.414	0.183	0.557	0.149	0.827	0.167
Illiquidity Portfolio 2(Zeros > country level upper 30%)	0.491	0.131	0.488	0.120	0.681	0.310
Size Portfolio 1(Market value <= country level lower 30%)	0.269	0.030	0.757	0.213	0.416	0.562
Size Portfolio 2(Market value > country level upper 30%)	0.500	0.181	0.473	0.115	0.815	0.181
Tech portfolio 1(Not TMT industry)	0.462	0.158	0.514	0.131	0.810	0.185
Tech portfolio 2(TMT industry)	0.345	0.151	0.615	0.161	0.795	0.189
Listing board portfolio 1(Main board)	0.482	0.155	0.493	0.128	0.807	0.188
Listing board portfolio 2(SME board)	0.112	0.179	0.572	0.137	0.838	0.155
Listing board portfolio 3(GEM board)	0.042	0.156	0.672	0.118	0.865	0.124

Table OA4. Summary Statistics of Market Development Variables

This table reports the time-series average of market development and liquidity variables for China and the U.S. from 1995Q1 to 2018Q4. Definitions of all the variables are described in detail in Appendix D.

	China	U.S.
REGDEV	1.04	
Number of public firms	1,759	5,438
Adjusted market development	0.00	
Zeros	0.05	0.02
Turnover	1.00	0.47

Table OA5. Extensions to China EY Model

This table shows PcGets results for different China EY model specifications. Column (1) shows the benchmark result. To measure portfolio-specific discount rates, variables in the discount rate groups are multiplied by Chinese betas, except for real interest rate which is multiplied by one minus Chinese betas. Specifically, in columns (2) and (3), we assume that the Chinese stock market is totally segmented. Chinese betas are estimated by running Chinese portfolio level excess returns on Chinese market level excess returns. In columns (4) and (5), we assume that the Chinese market is partially segmented, and we obtain Chinese betas and U.S. betas by running regressions of Chinese portfolio level excess returns on Chinese market level excess returns and U.S. market level excess returns. To incorporate U.S. expected return into the PcGet procedure, in columns (4) and (5), we assume the U.S. expected return is constant and add the estimated U.S. beta into the PcGets procedure. We estimate both an unconditional version and a conditional version. In an unconditional version, we estimate the beta from a full sample regression. In the conditional version, we obtain betas using the following procedure: in each quarter, we run the excess portfolio level excess return on the Chinese market excess return (also U.S. market excess return for columns 4 and 5) using data of the past 52 weeks.

Panel A. 1995-2018

Models	(1)	(2)	(3)	(4)	(5)
	Benchmark	Local Factor Only		Local and Global Factors	
	Table 6	Uncond. local MKT beta	Cond. local MKT beta	Uncond. Local and U.S. MKT beta	Cond. Local and U.S. MKT beta
Financial openness	42.0%	27.1%	39.1%	27.1%	39.1%
Domestic ownership structure	6.3%	4.5%		4.5%	6.3%
Market development	32.8%	55.0%	37.1%	55.0%	37.1%
Growth prospects	5.4%	4.2%	4.1%	4.2%	4.1%
Control variables					
Portfolio fixed effects	13.5%	9.1%	19.9%	9.1%	19.9%
Number of observations	4,873	4,873	4,857	4,873	4,857
Adjusted R-square	0.609	0.608	0.571	0.608	0.571

Panel B. 2003-2018

Models	(1)	(2)	(3)	(4)	(5)
	Benchmark	Local Factor Only		Local and Global Factors	
	Table 6	Uncond. local MKT beta	Cond. local MKT beta	Uncond. Local and U.S. MKT beta	Cond. Local and U.S. MKT beta
Financial openness	38.1%	38.2%	36.5%	39.3%	35.5%
Domestic ownership structure	31.7%	27.4%	13.8%	24.1%	8.9%
Market development	13.7%	13.6%	13.0%	24.9%	15.9%
Growth prospects	14.1%	50.4%	50.6%	7.0%	53.9%
U.S. expected return				26.4%	-2.3%
Portfolio fixed effects	2.4%	-29.6%	-13.9%	-21.6%	-11.9%
Number of observations	3,339	3,339	3,323	3,339	3,323
Adjusted R-square	0.650	0.642	0.615	0.642	0.623

Table OA6. Lagged Independent Variables

This table reports the results of portfolio-quarter panel regressions of valuation differentials on discount rate and growth prospects variables from 1995Q1 to 2018Q4. The dependent variable is the portfolio level earning yield differential between China and the U.S., DIFEY. The independent variables are lagged one-quarter differences between China and the U.S. except for "Dummy: 2009Q3" which is a dummy variable equal to 1 (0) during period after (before) 2009Q3. Control variables include leverage differentials, earnings growth differential and minimum number of stocks. Panel A (B) shows the results using discount rate (growth prospects) variables. Definitions of all these variables are described in detail in Appendix C. The regressions include portfolio fixed effects and the standard errors are double clustered by portfolio and time. T-statistics are in parentheses. ***, ** and * indicate significances at the 1%, 5% and 10% levels using two-tailed tests.

Panel A. Explaining the Valuation Gap: Changes in Discount Rates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dummy: after 2009Q3		0.007		0.024***		0.004		0.021***
		(1.266)		(5.126)		(0.808)		(3.675)
REGOPEN	0.004***	0.003**						
	(3.396)	(2.149)						
IA1	0.007	0.008						
	(0.805)	(0.924)						
IA2	0.135***	0.123***						
	(4.866)	(4.253)						
IA3	-0.014	-0.013						
	(-1.264)	(-1.176)						
Real interest rate	0.014	0.010						
	(0.250)	(0.178)						
Overall political rating	-0.032	-0.015						
	(-1.619)	(-0.781)						
Chinese state ownership			0.025	0.057***	0.080***	0.083***		
			(1.663)	(3.722)	(3.493)	(3.606)		
Institutional ownership					-0.072***	-0.066***		
					(-4.745)	(-4.420)		
Retail ownership					-0.024*	-0.020		
					(-1.924)	(-1.537)		
SNS					-0.039*	-0.032		
					(-1.758)	(-1.377)		
REGDEV							0.013***	0.004
							(3.316)	(1.007)
Number of public firms							-0.008	-0.010
							(-1.269)	(-1.655)
Adj. market development							-0.002***	-0.002***
							(-2.727)	(-2.697)
Zeros							0.106***	0.140***
							(3.405)	(4.705)
Turnover							-0.006***	-0.006***
							(-4.478)	(-4.650)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Portfolio fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample years	1995-2018	1995-2018	1995-2018	1995-2018	2003-2018	2003-2018	1995-2018	1995-2018
Number of observations	4,819	4,819	4,819	4,819	3,354	3,354	4,819	4,819
Adjusted R-square	0.375	0.379	0.241	0.333	0.443	0.444	0.339	0.364

Panel B. Explaining the Valuation Gap: Changes in Growth Prospects

	(1)	(2)	(3)	(4)	(5)
Dummy: after 2009Q3		0.018***		0.010**	0.011**
		(4.147)		(2.477)	(2.223)
GDP growth rate	-0.161***	-0.133**	-0.308***	-0.274***	-0.280***
	(-2.703)	(-2.629)	(-4.505)	(-4.954)	(-4.819)
Earnings growth expectation			-0.002**	-0.002*	0.003
			(-2.152)	(-1.958)	(1.089)
Sales growth expectation			-0.017	-0.012	-0.008
			(-1.556)	(-1.195)	(-0.698)
Earnings growth expectation×number of analysts					-0.009***
					(-3.215)
Earnings growth expectation×forecast dispersion					0.029***
					(2.990)
Number of analysts					0.002
					(0.799)
Forecast dispersion					-0.028***
					(-3.698)
Controls	Yes	Yes	Yes	Yes	Yes
Portfolio fixed effects	Yes	Yes	Yes	Yes	Yes
Sample years	1995-2018	1995-2018	2003-2018	2003-2018	2003-2018
Number of observations	4,819	4,819	3,263	3,263	3,150
Adjusted R-square	0.250	0.311	0.413	0.428	0.477