SEE FAR

EXECUTIVE SUMMARIES

THE WELLS FARGO ADVISORS CENTER FOR FINANCE AND ACCOUNTING RESEARCH
In May 2012, Wells Fargo Advisors awarded a gift to Washington University in St. Louis to support Olin Business School. Olin’s newly named Wells Fargo Advisors Center for Finance and Accounting Research (WFA-CFAR) will be a catalyst for enhancing finance and accounting research and education, which benefits faculty members, students, and businesses. To that end, initiatives housed under the WFA-CFAR umbrella include:

**Specialized master’s degree programs** in finance (MSF) and accounting (MACC), which provide rigorous curricula and industry-specific knowledge to students through a 10- or 17-month format.

**The Corporate Finance and Investments Platform**, which realigns our MBA curricula to provide students with industry-specific knowledge and experiential learning opportunities, while also ensuring that these students receive a broad business education.

**Sponsored research**, which includes company-specific projects as well as research on broader topics, to ensure that Olin faculty remain at the forefront of research excellence.

**Conferences and seminars**, which bring together scholars from all over the world to share the latest ideas in finance and accounting.

To obtain copies of the original research papers summarized here or to recommend your company for a future research project, please contact Kristen Jones, Wells Fargo Advisors Center for Finance and Accounting Research Program Manager at kristen.jones@wustl.edu or 314-935-4179.
A Message from the Director

I am pleased to continue our magazine, SEE FAR. Apart from the obvious attempt to “capitalize” on the WFA-CFAR name, the name also captures the essence of our research: looking to the future rather than concentrating exclusively on current events and thinking, and focusing on big-picture issues that have far-reaching consequences.

All the articles in SEE FAR are based on finance and accounting research that has been previously published in an academic journal or as a monograph, or is currently a working paper that will be published in the future. The original papers have been rewritten as executive summaries for SEE FAR so that they are accessible to a broad audience, rather than solely to those in academia. This is no small task. Taking a paper originally written for a highly technical academic audience and converting it into something more accessible takes a great deal of skill and hard work, as we discovered while putting together this issue and our past issues. But perhaps that is why the task is so worthwhile. I firmly believe that this will not only help us build a bridge between the research of Olin Business School faculty and those in the world of practice, but also will add to the knowledge people use on a daily basis. The intellectual capital generated by our faculty members’ research is quite impressive—Olin consistently ranks among the top 10 schools in terms of our research output. For this reason, it is important that WFA-CFAR research is made available to as many of our stakeholders as possible.

CFAR has articulated a new statement of the higher purpose of the center. This statement is: To be a focal point for the support and dissemination of research in finance, accounting, and authentic higher purpose—and change the world through academic research, one idea at a time! This statement is focused on the prosocial nature of the center’s activities, including the research it promotes. The center helped organize a high-impact conference on organizational and personal higher purpose in November 2019, and will be engaged in activities that build on the insights generated during the conference.

I hope that you enjoy reading the summaries in this issue. I would like to thank my faculty colleagues who participated in helping us create this issue by providing their papers and working with us to convert them into what you will read on the following pages. I look forward to any feedback you have to help us improve this magazine. Please contact WFA-CFAR Program Manager Kristen Jones at kristen.jones@wustl.edu with your insights.

Sincerely yours,

Anjan Thakor
John E. Simon Professor of Finance, Director of Doctoral Programs, Director of the WFA Center for Finance and Accounting Research, Olin School of Business, Washington University in St. Louis
We test for a “hot hand” in Major League Baseball using panel data. We find strong evidence for its existence in all ten statistical categories we consider. The magnitudes are significant; being “hot” corresponds to between one-half and one standard deviation in the distribution of player abilities. Our results are in notable contrast to the majority of the hot-hand literature, which has generally found either no hot hand or a very weak hot hand in sports, often employing basketball shooting data. We argue that this difference is attributable to endogenous defensive responses: basketball presents sufficient opportunity for transferring defensive resources to equate shooting probabilities across players whereas baseball does not. We then document that baseball teams do respond to recent success in their opponents’ batting performance. Our results suggest teams respond in a manner consistent with drawing correct inference about the magnitude of the hot-hand except for a tendency to overreact to very recent performance (i.e., the last five attempts).

Klay Thompson, a five-time NBA all-star who plays for the Golden State Warriors, has a career field goal percentage of 46%. He just hit his last three shots from the field. How likely is he to make his next shot? The “hot hand fallacy” is that most people overestimate the likelihood that Klay will make his next shot. This tendency is frequently cited in behavioral economics as an example of a widespread cognitive mistake because it overlooks the impact of pure luck on outcomes and over-attributes outcomes to the player’s ability to be on a “streak” at that point in the game. This literature began with a seminal study by Gilovich et al., 1985’s GVT, who first documented a widespread belief in the hot hand. GVT then conducted a series of statistical tests using basketball shooting data and find little evidence for its existence. That is, recent success (e.g., making three shots in a row) does not predict future success (e.g., making the next shot). Subsequent literature on the topic follows GVT in purporting to show that sports players do not have a propensity for hot or cold streaks, contrary to the almost universally held perceptions among players, coaches, and fans alike. According to a survey of this literature, “The empirical evidence for the existence of a hot hand is considerably limited” (Bar-Eli, 2006). Consequently, the pervasive perception of a hot hand has been attributed to a basic cognitive
mistake: people infer patterns in random data. In particular, they place too much weight on recent success or failure in predicting future outcomes.

While sports have been used as a laboratory to study the hot hand fallacy, the potential ramifications of the misperception extend well beyond the arena of sports. If people over-extrapolate from recent success or failure then resources will be misallocated. Managers will (incorrectly) assign too many important tasks to employees who have had recent success. CEOs will overinvest in sectors or projects that have had a string of good luck. Investors will unjustifiably divert capital to fund managers who have performed well recently. Thus, the extent to which the purported cognitive bias exists and whether it results in suboptimal decision making are important questions for economic research.

Recent work on the hot hand fallacy has highlighted several flaws in the prior literature. First, it often fails to take account of endogenous strategic responses. For example, a hot shooter in basketball should be defended more intensely, and should take more difficult shots, both of which will lower his shooting percentage subsequent to the initial success. Therefore, we should not expect to see evidence of a hot hand in shooting data from NBA games even if the players themselves are streaky. Second, many of the tests used in this literature lack the power to reject plausible models of the hot hand and suffer from small-sample biases (Miller and Sanjurjo, 2018).

In this paper, we revisit the hot hand fallacy by examining data from major league baseball. Baseball data is particularly well-suited to address the both of the drawbacks mentioned above. First, as we will argue below in more detail, the scope for transferring defensive resources across players is far more limited in baseball than in basketball. Second, there is an abundance of baseball data. There are roughly 200,000 batter plate appearances per year and we employ twelve years of data, giving us over two million observations. Second, we look for general responses are justified or indicative of a cognitive mistake.

### Endogenous Responses: Basketball vs. Baseball

A critical difference between baseball and basketball that is central to our motivation is that the scope for transferring defensive resources across players is far more limited in baseball than in basketball. We argue that in basketball this defensive endogeneity is sufficient to equate shooting margins across players, and consequently, one should not observe hot streaks; better players, and hotter players, should be guarded more, up to the point where they do not shoot better than teammates. In contrast, in baseball, the mechanism for transferring defensive resources across players is much more limited, and consequently, different margins should exist across players, whether due to long-run ability or a hot hand.

There is an obvious and widely accepted explanation for why the NBA scoring leaders and universally acknowledged stars do not collectively shoot any better than the league average. Even a very casual basketball observer can readily observe that James Harden draws far more defensive attention, and attempts far more difficult shots, than his teammates. Indeed, a fundamental aspect of defensive strategy in basketball involves allocating more defensive attention to better shooters and less to weaker shooters. In equilibrium, this should lead to fewer on average being attempted on offensive players at any given time. If there are better marginal shots available for Player A than Player B, Player A should be shooting more and Player B should be shooting less, and defenses should be covering Player A more and Player B less, until this discrepancy vanishes.

A similar argument can be made for players who are temporarily better players as well, i.e., for hot players. A hot player, who is temporarily shooting better than he normally does, would temporarily exceed his teammates’ shooting percentage if he only took his usual shots and received his usual defensive attention. In response, he should attempt more difficult shots, and he should receive more defensive attention, until these adjustments lower his marginal shooting percentage to equate it with his that of his teammates. Consequently, for the same reason good shooters do not exhibit a significantly higher shooter percentage than average players, “hot” shooters should not exhibit a significantly higher shooting percentage than they typically realize.

In baseball the ability to transfer defensive resources across players in baseball is far more limited than in basketball, and will be insufficient to equate margins across batters. Consequently, in equilibrium marginal success rates (i.e., batting averages, home run frequencies, etc.) will not be equated across batters on a team as they are across shooters in basketball.

The primary reason for this difference is that in baseball, pitchers face batters sequentially rather than simultaneously. Unlike in basketball, there is not a fixed defensive resource that must be allocated across all the batters (offensive players). Rather, each hitter faces the pitcher and his defensive team (his fielders) one at a time, and the pitcher and all fielders can focus almost all of their defensive attention on the hitter at hand. Consequently, if Batter A is a better hitter than Batter B, the opposing team has little scope to transfer defensive resources from B to A.

This fundamental difference in the ability to transfer defensive resources across players in the two sports gives rise to a very different relationship between ability and outcomes in the two. Figure 2 shows a very strong positive relationship between batting average and both hits and plate appearances (chances for hits). This is in sharp contrast with Figure 1, where there is effectively no relationship between

---

**Figure 1:** Relation between points scored, attempts and success percentage in the National Basketball Association (NBA) using data from the 2012-13 season. Figures include data from all players who started in at least 50% of all regular season games. In neither category is the correlation coefficient statistically different from zero.

**Figure 2:** Relation between number of successes, number of attempts and success percentage in Major League Baseball using data from the 2013 season. Figures include data from all players who started in at least 50% of all regular season games. In both relationships, the coefficient is positive and statistically significant.

### References

shooting percentage and either points scored or shots taken. More generally, other measures of hitter ability (i.e., salary) correlate strongly with hitting statistics such as batting average in baseball where they do not with shooting percentage in basketball.

The limited ability to transfer defensive resources across hitters in basketball should not be confused with ability to make optimal defensive adjustments for each hitter that do not involve transfers across players. Pitchers choose their type of pitches (i.e., fastball, change-up, slider) and the location of their pitches (i.e., inside, outside, high, low) based on each hitter’s history at hitting such pitches. And fielders regularly shift in the field, based on the hitter’s proclivity for hitting to certain spots, the game situation, and the type of pitches likely to be thrown. These common adjustments, however, do not involve moving resources from one hitter to another, but rather are made optimally for each hitter.

Conceptual Framework: What is a Hot Hand?

While there does not appear to be any consensus on how to define a hot hand, the literature has generally adopted a notion in line with short-term predictability. That is, does recent success predict future success after controlling for all other predictive factors?

One factor that should be controlled for is the underlying ability of the player in question. But it is also well acknowledged that players’ abilities change over time. Should such changes be considered a hot hand? We believe that low-frequency changes in ability should not be associated with a hot hand. That Chris Davis hit five home runs in 2011 (playing only part time), 33 home runs in 2012, and a league leading 53 home runs in 2013, is not taken as evidence of streakiness, but rather, is interpreted as evidence that he was a much better hitter in 2013 than 2011.

When most players, coaches, and fans talk of a hot hand, they are usually referring to predictability at a fairly high frequency—on the order of anything from within a game to 1-2 months. For instance, both of the following examples fit most people’s intuition of a hot hand:

• In 1997, Joey Cora batted 0.475 during a 24-game span (roughly 4 weeks). He batted 0.262 during the rest of the season.

• On January 23, 2015, Klay Thompson shot a perfect 13-for-13 from the field including nine three-pointers during one quarter of a home win against the Sacramento Kings. He shot 3-for-12 from the field during the other three quarters of the game.

We will define streakiness as relatively high-frequency variation in a player’s ability (e.g., daily, weekly or monthly changes to ability) and will not consider lower frequency variation as evidence of streakiness.

In our conceptual framework, a player’s success probability in any given attempt is determined by three components:

01 Long-term player specific component (ability), e.g., raw talent, skill, speed, strength, hand-eye coordination.

02 Short-term player specific component (state), e.g., confidence, attitude, physical health, adjustments to technique.

03 External factors (controls): opponent ability, team strategy, game situation, stadium, difficulty of shot or attempt, etc.

Our primary interest is on the extent to which a player’s state predicts future performance. To do so, we will estimate the following regression equation:

\[
\text{Outcome}_{it} = \alpha + \gamma \times \text{State}_{it} + \beta \times \text{Ability}_{it} + \beta \times \text{Controls}_{it} + \epsilon
\]

We estimate the model for five different outcomes (for both batters and pitchers) corresponding to whether the player got a hit, a home run, reached base, walked, or struck out. We measure the player’s state using his recent success rate in the outcome. In selecting the history length for measuring state, one encounters the following trade-off. If the state does not change over the period, a longer period will provide a more accurate identification of the state. A hitter who has three hits in his last three at-bats might be hot, or he might have been lucky. More at-bats in the history can help distinguish between the two. However, if the state changes significantly over the length of the history, one will be using a history that is less relevant for measuring the current state and more relevant for measuring long-term ability.

For the majority of our analysis we use the last 25 attempts. This history length corresponds roughly to performance in the five most recent games for a batter (roughly the last week of performance). We then refer to a player as “hot” if recent performance is above some upper threshold and “cold” if recent performance is below some lower threshold.

We consider a variety of methods to control for the player’s ability. In the baseline model, we do so by calculating the player’s rate of success during the season not including a “window” of attempts before and after the current attempt. Our additional control variables include: ability of the opposing pitcher, stadium, time trends, and the platoon effect. Our data set provides a wide range of additional situational variables that one might expect to predict the success of a particular outcome (e.g., inning of at bat, number of outs, number of runners on base or in scoring position). However, similar to Albright (1993), these variables were not found to be statistically significant nor do they alter the predictions of the model and are generally omitted from our reported results.
Evidence of Streakiness

We find that baseball players exhibit large and strategically significant streaks across all ten statistical categories which we examine. Furthermore, these effects are of a significant magnitude: for instance, a hot hitter will exhibit an on-base percentage roughly 28 points higher than when cold after controlling for all other explanatory variables. The difference between a hot batter and a cold batter is roughly equivalent to the difference between a 70th percentile batter and a 50th percentile one. Our results have a similar degree of significance across other statistics. For example, a batter who is “hot” in home runs is 15-25% more likely to hit a home run in his next at bat. In other words, an average hitter bats more like a power hitter when he is hot in hitting home runs.

One useful way of organizing and summarizing our results across all statistics is to compare the difference in performance between a hot and a cold player to that of a one standard deviation in long-run average across players for the statistic. We report these estimates in Table 1. The difference between a hot and cold player inferred from recent history ranges from roughly one-half to one standard deviation in the variation across overall player ability. Averaged across all five statistics, we find the difference between a hot and cold player corresponds to 0.84 and 0.68 of a standard deviation in long-run average for batters and pitchers respectively.

Table 1: Normalized Magnitude of the Hot Hand Effect.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Distribution of Ability</th>
<th>Magnitude of Hot Hand Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
</tr>
<tr>
<td>Hits Bat</td>
<td>0.268</td>
<td>0.0223</td>
</tr>
<tr>
<td>On Base Bat</td>
<td>0.339</td>
<td>0.0299</td>
</tr>
<tr>
<td>Home Runs Bat</td>
<td>0.0304</td>
<td>0.0169</td>
</tr>
<tr>
<td>Strike Outs Bat</td>
<td>0.191</td>
<td>0.059</td>
</tr>
<tr>
<td>Walks Bat</td>
<td>0.0966</td>
<td>0.0302</td>
</tr>
<tr>
<td>Hits Pitch</td>
<td>0.264</td>
<td>0.028</td>
</tr>
<tr>
<td>On Base Pitch</td>
<td>0.335</td>
<td>0.0284</td>
</tr>
<tr>
<td>Home Runs Pitch</td>
<td>0.0307</td>
<td>0.0098</td>
</tr>
<tr>
<td>Strike Outs Pitch</td>
<td>0.191</td>
<td>0.0538</td>
</tr>
<tr>
<td>Walks Pitch</td>
<td>0.0965</td>
<td>0.0232</td>
</tr>
<tr>
<td>Average Bat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Pitch</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall Average</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Distribution of Ability</th>
<th>Magnitude of Hot Hand Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std. Dev.</td>
</tr>
<tr>
<td>Hits Bat</td>
<td>0.268</td>
<td>0.0223</td>
</tr>
<tr>
<td>On Base Bat</td>
<td>0.339</td>
<td>0.0299</td>
</tr>
<tr>
<td>Home Runs Bat</td>
<td>0.0304</td>
<td>0.0169</td>
</tr>
<tr>
<td>Strike Outs Bat</td>
<td>0.191</td>
<td>0.059</td>
</tr>
<tr>
<td>Walks Bat</td>
<td>0.0966</td>
<td>0.0302</td>
</tr>
<tr>
<td>Hits Pitch</td>
<td>0.264</td>
<td>0.028</td>
</tr>
<tr>
<td>On Base Pitch</td>
<td>0.335</td>
<td>0.0284</td>
</tr>
<tr>
<td>Home Runs Pitch</td>
<td>0.0307</td>
<td>0.0098</td>
</tr>
<tr>
<td>Strike Outs Pitch</td>
<td>0.191</td>
<td>0.0538</td>
</tr>
<tr>
<td>Walks Pitch</td>
<td>0.0965</td>
<td>0.0232</td>
</tr>
<tr>
<td>Average Bat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Pitch</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall Average</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Do Teams (Correctly) Respond to Streaky Hitters?

We will now ask whether defenses respond to streaky hitters. And if so, is their response consistent with them drawing a correct inference about the magnitude of the effect? Or, do they tend to over-react to recent performance in line with a hot-hand fallacy? To do so, we will utilize two facts:

01 Power hitters get walked more frequently than average hitters. This can most easily be interpreted as the defense “pitching around” the hitter. The cost to a defense of giving up a walk is significantly less than giving up a home run and so it makes sense from a defensive standpoint to err on the side pitching outside the strike zone to power hitters, which ultimately leads to more walks.

02 Average hitters bat similar to power hitters when they are hot in home runs. This was documented in the previous section.

Thus, if teams correctly respond to streaky hitters then average hitters who are hot should be walked with the same frequency as power hitters. This observation forms the basis of our empirical tests.

Our findings are as follows. First, teams clearly respond to streaky hitters. A batter who is hot in home-runs or extra-base hits is walked significantly more often then the same batter when he is not hot. Second, when measuring the batters current state using the last 25 attempts, defensive responses are remarkably consistent with defenses making the correct inference. Third, when the last 25 attempts is further decomposed into to finer intervals, we find that defenses significantly overreact to very recent performance (i.e., the last five attempts).

Hence, while we find that opposing teams appear to draw correct inferences about the hot hand of opposing batters using the week of performance, they tend to over react to streaky performances during the last game or within the current game. Moreover, this overreaction is fairly large and strongest in the last two attempts. This result is consistent with a version of the hot-hand fallacy in which significant streaky behavior exists and yet agents have a tendency to overestimate its magnitude–albeit, they only overestimate the importance of

Hot Hands in Other Settings

We focus in this paper on the distinction between basketball (for which streakiness has been extensively tested) and baseball (which we test). Yet, our research motivates a broader question on the existence of a hot hand in other sports and activities. We hypothesize that this overlap between activity will generally exhibit streakiness; that is, there will be transitory components to ability as well as long-run components. However, we would expect to find evidence of streakiness in outcomes if and only if the activity at hand does not permit an endogenous response that is likely to equate margins, as in basketball. In some of these cases the distinction between the presence and absence of a defensive response that equates margins should be obvious (there is no defense in golf or bowling), in other cases it is more subtle and depends on the details of the sport (the distinction between basketball and baseball). However, as we will now argue, there is a simple and intuitive way to identify the presence of such a defensive response, and consequently, whether one would expect to find streakiness in outcomes.

In particular, endogenous responses, when available, should equate margins both across permanent and transitory differences in abilities. Thus, to identify settings where available endogenous responses are sufficient to negate short-term streakiness is to identify settings with permanent differences in success rates. If performance in a particular statistic is correlated with other measures of ability (such as salary), and if it is persistent across long stretches of time (i.e., over years), that is indicative of a setting where endogenous responses are not sufficient to equate margins, and hence where we would also expect to find short-run streakiness. In contrast, if instead performance does not correlate with other measures of ability, and if performance does not persist across years, this is indicative of the presence of an endogenous response which equates margins, and we would not expect to find streakiness in outcomes. Testing and refining this hypothesis is a fruitful direction for future research.
We first summarize the evidence that contributes to the funding gap in biomedical research. We then describe how FDA Hedges would be structured, and the conceptual argument for how they would help encourage additional drug research. We then summarize evidence of what the risk and return characteristics of these FDA hedges would be, using detailed data on drug projects.

The Risk of the Drug Development Process
The process to develop a drug in the United States, as well as other countries, is highly regulated. In the U.S., the Food and Drug Administration (FDA) maintains regulation and oversight of any biopharmaceutical (biopharma) firm that develops a drug.

The approval process consists of three phases

As the global pandemic has prominently highlighted, medical innovation—the development of new drugs and therapeutics to battle diseases—is essential to the economy. However, despite the importance of new therapeutics, the development of them runs into significant financial challenges due to a “funding gap”—many treatments that are valuable from a societal point-of-view are not funded. This is due to the fact that the drug development process is risky, expensive, and lengthy. These aspects of the drug development process make investment performance in the sector seem meager to investors from a risk-reward standpoint.

We propose that financial innovation is a potential solution to this conundrum. To demonstrate how financial innovation may work in this way, we summarize findings from a recently published paper in The Review of Corporate Finance Studies (see Jørring, Lo, Philipson, Singh and Thakor, forthcoming). In the paper, we propose “FDA Hedges” as a financial innovation that can help investors and biopharmaceutical companies share the risk inherent in drug development, which can then spur additional research and development (R&D) to produce new drugs.

A Financial Innovation to Improve Medical Innovation
RICHARD T. THAKOR, University of Minnesota

As the global pandemic has prominently highlighted, medical innovation—the development of new drugs and therapeutics to battle diseases—is essential to the economy. However, despite the importance of new therapeutics, the development of them runs into significant financial challenges due to a “funding gap”—many treatments that are valuable from a societal point-of-view are not funded. This is due to the fact that the drug development process is risky, expensive, and lengthy. These aspects of the drug development process make investment performance in the sector seem meager to investors from a risk-reward standpoint.

We propose that financial innovation is a potential solution to this conundrum. To demonstrate how financial innovation may work in this way, we summarize findings from a recently published paper in The Review of Corporate Finance Studies (see Jørring, Lo, Philipson, Singh and Thakor, forthcoming). In the paper, we propose “FDA Hedges” as a financial innovation that can help investors and biopharmaceutical companies share the risk inherent in drug development, which can then spur additional research and development (R&D) to produce new drugs.

We first summarize the evidence that contributes to the funding gap in biomedical research. We then describe how FDA Hedges would be structured, and the conceptual argument for how they would help encourage additional drug research. We then summarize evidence of what the risk and return characteristics of these FDA hedges would be, using detailed data on drug projects.
after preclinical testing, designed to ascertain the safety and efficacy of a drug candidate. Once a drug has successfully passed the clinical trials phases, it then must undergo a final review by the FDA before it is able to be marketed to consumers. Figure 1 summarizes the FDA approval process.

As detailed by Lo and Thakor (2022), the drug development process has three key characteristics that can make it difficult for biopharma firms to undertake. First, the process is risky. The risk of the development process is shown in Table 1. As can be seen, the overall probability of failure for an average drug ranges from 74% to 95%, depending on the therapeutic area, with oncology being the riskiest.

Second, the process is expensive. Recent estimates of the average cost of developing a drug by DiMasi, Grabowski, and Hansen (2016) is $2.558 billion (in 2013 dollars). Development costs have also been increasing over time. As shown in Figure 2, Scannell et al. (2012) demonstrated how $1 billion (inflation-adjusted) of R&D spending could produce dozens of drugs in the 1950s, but could not even produce one drug by 2010.

Third, the process is lengthy. DiMasi and Grabowski (2007) shows that Phase 1 trials take an average of 1.3 years to complete, Phase 2 trials an average of 2.3 years, Phase 3 trials an average of 2.8 years, and the final FDA approval phase an average of 1.4 years. Table 2 breaks down the estimates of the length of time it takes to develop a drug.

### Table 1: Probabilities of Phase Failure by Disease Group

The table from Jørring, et al. (forthcoming) shows the average probability of failing each phase of the FDA drug development process, broken down by disease groups. These failure rates are from data from 2006-2015, and are taken from Thomas, et al. (2016).

<table>
<thead>
<tr>
<th>Disease Group</th>
<th>Phase 1</th>
<th>Phase 2</th>
<th>Phase 3</th>
<th>NDA/BLA Approval Phase</th>
<th>Overall Probability of Failure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hematology</td>
<td>27%</td>
<td>43%</td>
<td>25%</td>
<td>16%</td>
<td>74%</td>
</tr>
<tr>
<td>Infectious Disease</td>
<td>31%</td>
<td>57%</td>
<td>27%</td>
<td>11%</td>
<td>81%</td>
</tr>
<tr>
<td>Ophthalmology</td>
<td>15%</td>
<td>55%</td>
<td>42%</td>
<td>23%</td>
<td>83%</td>
</tr>
<tr>
<td>Other Disease Groups</td>
<td>33%</td>
<td>60%</td>
<td>30%</td>
<td>12%</td>
<td>84%</td>
</tr>
<tr>
<td>Metabolic</td>
<td>39%</td>
<td>55%</td>
<td>29%</td>
<td>22%</td>
<td>85%</td>
</tr>
<tr>
<td>Gastroenterology</td>
<td>24%</td>
<td>64%</td>
<td>39%</td>
<td>8%</td>
<td>85%</td>
</tr>
<tr>
<td>Allergy</td>
<td>32%</td>
<td>68%</td>
<td>29%</td>
<td>6%</td>
<td>85%</td>
</tr>
<tr>
<td>Endocrine</td>
<td>41%</td>
<td>60%</td>
<td>35%</td>
<td>14%</td>
<td>87%</td>
</tr>
<tr>
<td>Respiratory</td>
<td>35%</td>
<td>71%</td>
<td>29%</td>
<td>5%</td>
<td>87%</td>
</tr>
<tr>
<td>Urology</td>
<td>43%</td>
<td>67%</td>
<td>29%</td>
<td>14%</td>
<td>89%</td>
</tr>
<tr>
<td>Autoimmune/immunology</td>
<td>34%</td>
<td>68%</td>
<td>38%</td>
<td>14%</td>
<td>89%</td>
</tr>
<tr>
<td>Neurology</td>
<td>41%</td>
<td>70%</td>
<td>43%</td>
<td>17%</td>
<td>92%</td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>41%</td>
<td>76%</td>
<td>45%</td>
<td>16%</td>
<td>93%</td>
</tr>
<tr>
<td>Psychiatry</td>
<td>46%</td>
<td>76%</td>
<td>44%</td>
<td>12%</td>
<td>94%</td>
</tr>
<tr>
<td>Oncology</td>
<td>37%</td>
<td>75%</td>
<td>60%</td>
<td>18%</td>
<td>95%</td>
</tr>
</tbody>
</table>

### Table 2: Probabilities of Phase Failure by Disease Group

This table, from Jørring et al. (forthcoming), shows the average length of each phase in the FDA approval process for the biotech and pharma sectors. Phase length is in months (years in parentheses). NDA/BLA stand for new drug application/biologic license application. Estimates come from DiMasi & Grabowski (2007).

<table>
<thead>
<tr>
<th>Average Length of Time in Months (years)</th>
<th>Phase 1</th>
<th>Phase 2</th>
<th>Phase 3</th>
<th>NDA/BLA Approval Phase</th>
<th>Total Length of Time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>15.9 (1.3)</td>
<td>27.65 (2.3)</td>
<td>33.35 (2.8)</td>
<td>17.10 (1.4)</td>
<td>94.0 (7.8)</td>
</tr>
</tbody>
</table>
A body of academic evidence has emerged showing that these characteristics lead to important financial implications (see Lo and Thakor, 2022). First, due to the high investment costs, drug development firms are generally forced to rely on external financing in order to fund their investments. Second, this link means that the prospects of biopharma firms will be strongly linked to the state of the market, since downturns generally lead to less availability of financing. Third, frictions associated with financial markets means that some therapies, even though they are valuable from a societal perspective, are not funded.

**FDA Hedges**

To address this problem, we propose a financial innovation known as “FDA Hedges,” which are options with characteristics that are similar to those that would be attractive to both buyers and issuers. We argue that these options are particularly well suited to address the problem of funding biopharma firms, as they can provide investors with a way to hedge the risks of their investments. In order to make the case that they would be appealing to both issuers and buyers, we consider carefully the characteristics of the risk and return associated with each option.

### Evidence on the Risk and Return Characteristics of FDA Hedges

In light of the riskiness of R&D that we referred to earlier, a legitimate question is: How much risk is involved in FDA hedges? Would market participants be willing to take that risk? To address this question, we now turn to what the characteristics of FDA hedges would be if they were traded, in order to make the case that they would be appealing to both issuers as well as those who would purchase the hedges. In order to do so, we make use of detailed data which tracks the landscape of drug projects that were undertaken by biopharma firms in the U.S. over the past 20 years—a dataset of over 11,000 drugs developed by 2,800 biopharma companies. These data include not only the status of a given drug in development at any given time, but also estimates of the eventual likelihood of future FDA approval based on historical information about the drug's therapeutic area and information about its trial success thus far.

Using this information, we can calculate what the “fair” prices of FDA hedges would have been for drugs, as well as their risk characteristics. The data also enable us to simulate what the benefit for issuers would have been to sell these hedges. We obtain a number of results from our analysis.

First, we show that the risk of FDA hedges is unrelated to systematic factors such as the overall stock market or the macroeconomy. That is, they have very little systematic risk. This may increase their appeal to both buyers and issuers since investors can diversify away idiosyncratic risk and charge risk premia mainly for systematic risk. From the perspective of biopharma firms, FDA hedges allow a direct hedge against the scientific (idiosyncratic) risk of the firm's stock. The firm may appear to be a more attractive investment by reducing this risk; as a result, biopharma firms may wish to purchase FDA hedges in order to attract capital from investors. Alternatively, investors themselves may wish to purchase FDA hedges directly to offset the risk of their own investments in biopharma firms.

Second, from the perspective of issuers offering FDA options, the risk patterns we document allow issuers to hedge some of the FDA option risk, thus further improving their risk-reward trade off.

Finally, we offer “proof of concept” through case studies of instruments trading FDA risks—contingent valuation right (CVR) issued in mergers and acquisitions (M&A) deals—that pay investors pre-specified amounts when certain milestones are met as part of a M&A deal structure. Similar in many respects to FDA hedges, we show that these instruments are liquid and follow predicted pricing and volume patterns. As these milestones oftentimes include FDA approval decisions, they contain implicit FDA options.

### Implications

In summary, biopharma firms face high costs and risks, and this risk contributes to under investment in R&D in welfare-enhancing drugs. We propose and investigate a new form of financial innovation, FDA hedges, which allow biomedical R&D investors to share the pipeline risk associated with the FDA approval process with broader capital markets.

We view such financial innovations as a first step in potentially solving under funding of valuable treatments through a market-based approach. We believe that the time is right for innovations like FDA hedges to be tested in the market since they can significantly increase the flow of capital from the financial market to health-enhancing and life-saving drugs and treatment options. This would also potentially reduce reliance on government funding and provide a further boost to basic medical R&D that is applied to improve the human condition.

### References


Using Capital Requirements as a Regulatory Tool to Influence Ethics, Talent and Compensation in Banking

FENGHUA SONG, Smeal College of Business, Pennsylvania State University

In my recent paper with Anjan Thakor (Song and Thakor, 2022), we study four important issues in banking, namely ethics, talent competition/allocation, managerial compensation, and capital requirements, in a unified economic framework. The analysis provides novel insights into how capital requirements, besides their usual prudential regulation role, can influence the setting of ethical standards, the nature of labor market competition and managerial compensation in financial services, and the extent of financial innovation. A key finding of our theoretical analysis is that, while a focus on ethics is socially important and higher ethical standards can be achieved with higher capital requirements, achieving higher ethical standards is not a “free lunch” from a social welfare standpoint—an excessive focus can lead to less financial innovation and talent migration out of banking.

Although the majority of banks exhibit ethical behavior, alleged ethical violations in banking are often in the news. One often reads shocking headlines like, “Three Former ICAP Brokers Appear in British Court in Labor Manipulation Case” (New York Times, April 15, 2014), and “The Big Banks are Corrupt—and Getting Worse” (Huffington Post, May 22, 2016). Many commentators have argued that ethical violations may have been responsible for numerous failures in banking. Not surprisingly, actions have been taken by regulators since the subprime crisis: global banks paid $321 billion in fines between 2008 and 2017 for alleged legal and ethical transgressions. Of course, culpability is often hard to assign since most cases did not go to trial and banks settled by paying fines, often without admitting guilt, leading some observers to claim regulatory overreach. Besides imposing penalties, regulators have also increasingly emphasized the importance of culture and ethics in banking. For example, both Thomas Baxter, then general counsel at the Federal Reserve Bank of New York, and Federal Reserve governor, Jerome Powell, gave speeches on the same day (January 20, 2015) about the importance of ethics, with Baxter stating: “At the New York Fed, we have made ethical culture a priority for financial services.”

Meanwhile, there is growing concern among bankers about talent migration out of depository banking. The media has widely reported on the post-crisis departure of talent from bank holding
companies. Regulators have also taken note on this. The 2016 Fed Conference of State Bank Supervisors noted that a survey of community banks revealed: “A ‘brain drain’ has been blamed for problems in attracting sufficient talent for future leadership of the bank. This is a particular problem because of impending retirements of older management in many community banks.”

We develop an economic framework showing that these twin developments of ethical lapses and talent migration are two sides of the same coin. Specifically, a bank’s choice of its ethical standard, through the design of managerial compensation contracts, affects its ability to compete for managerial talent in the labor market and, consequently, its innovation capability. Regulators can influence ethical standards in banking with capital requirements, thereby affecting talent allocation across safety-net protected depositories and unprotected shadow banks.

A Unified Framework of Financial Innovation, Managerial Compensation, Banking Ethics, Talent Competition, and Capital Requirements

We focus on mis-selling financial products as a specific form of unethical banking behavior. We consider banks with varying safety-net protections, including safety-net protected depositories and unprotected shadow banks. Each bank hires a manager to develop a new financial product (i.e., financial innovation), who then, conditional on successful innovation, decides whether to sell the product to customers. Mis-selling an unsuitable product harms customers and also entails a loss for the bank (e.g., loss of business, reputation-related damage, legal and/or regulatory fines) upon detection. A higher banking ethical standard reduces the odds of mis-selling. However, since the manager is also responsible for product sale, the bank’s ethical standard is effectively set by the manager; the bank influences the manager to set a particular standard by designing her compensation contract to incentivize a choice of that standard. The bank’s capital level influences its desired ethical standard; banks with higher capital levels ceteris paribus prefer higher ethical standards to lower the odds of mis-selling, thereby protecting their capital from being eroded due to the loss incurred when mis-selling is detected. Therefore, by setting capital requirements, regulators can influence banks’ desired ethical standards, and consequently their managerial compensation contracts to implement those standards. As will be discussed later, the framework also shows that a bank’s managerial compensation contract critically affects its ability to compete for talent in the labor market, which feeds back to impact its innovation capability. Relationships among various elements of the framework are illustrated by the following figure:

Findings: Our analysis based on this framework leads to several key interesting findings which we discuss on the next few pages.

01 Triple Roles of Managerial Compensation

Managerial compensation plays key roles in our framework. The compensation contract consists of a base salary (which the manager always gets, even without innovation and product sale) and a bonus (which the manager obtains only if she develops and sells a suitable new product to the customer). The steepness (pay-for-performance sensitivity) of the contract, the ratio , plays three key roles. First, a higher leads to stronger managerial incentive to work hard to engage in innovation, since the bonus is not obtainable if the manager fails to develop the new product. Second, a higher increases the manager’s incentive to sell the new produce once it is developed, since the bonus is unobtainable if the product is not sold (even if it is developed); this increases the odds of mis-selling and results in a lower banking ethical standard. Third, a higher enables the bank to attract a more talented manager in the labor market who is ceteris paribus more capable of developing the new product, and hence selling it and getting the bonus. These three roles of managerial compensation are depicted in the following figure:

02 Trade off between Banking Ethical Standard and Financial Innovation

As discussed earlier, regulators can impose higher capital requirements to induce a bank to adopt a higher ethical standard to protect its (high) capital. However, a higher ethical standard is implemented through a managerial compensation contract with a lower , which, as just explained, not only prevents the bank from attracting more talented managers in the labor market, but also reduces the compensation contract’s incentive power in inducing any hired manager to engage in innovation. Both of these limit the bank’s innovation capability. Therefore, stimulating financial innovation may require tolerating lower ethical standards. This trade off is missed in popular arguments for improving ethics through heightened penalties and more stringent regulations. Our analysis reveals that ethics may not always be a free good; regulators need to strike a correct balance between the desire for ethics on the one end and the need for innovation on the other.

03 Talent Migration and Socially Optimal Talent Allocation

The analysis then examines talent competition among banks and their ethical standards, showing that they are two sides of the same coin. To implement a higher ethical standard, as explained earlier, a bank needs to adopt a managerial compensation contract with a lower . However, a lower also dilutes the bank’s ability to attract a more talented manager who is more likely to innovate and, hence, prefers a higher (instead, the bank will attract a less talented manager who is less likely to innovate and, therefore, prefers a contract with a relatively higher base salary which the manager can always secure). That is, our analysis reveals that a bank cannot have a higher ethical standard than its competitor and at the same time attract better talent through managerial contract design.
To make the point, we consider two types of banks: safety-net protected depositories and unprotected shadow banks. We show that regulators, who maximize aggregate welfare from the two banking sectors but otherwise have no preference over which sector innovates more, prefers to let the shadow banking sector have lower ethical standards and, therefore, attract better talent. The idea is that the socially optimal talent allocation, attaching no weight to which banking sector engages in more financial innovation and sales, minimizes the exposure of the public safety net for depositories by inducing them to adopt higher ethical standards and inevitably, therefore, hire less talented managers.

Implementation of Talent Allocation through Capital Requirements

Both depositories and shadow banks, ceteris paribus, prefer to hire talented managers, and recognize that lowering the ethical standard, which is implemented through compensation contracts with a higher bonus, improves the odds of hiring a more talented manager. Thus, if talent competition is sufficiently strong, each bank may prefer to set an ethical standard lower than its competitor in order to “beat” the competitor in the labor market and attract better talent. This will set in motion a “race to the bottom” among banks in ethical standards. To prevent this race which is socially destructive, regulators need to set the gap between capital requirements for depositories and shadow banks wide enough to preempt depositories from starting the race in the first place. In doing so, regulators deliberately put depositories at a competitive disadvantage relative to shadow banks in terms of capital requirements by imposing much higher capital requirements on depositories. This will inevitably result in less financial innovation being produced in depositories than in shadow banks.

Existing studies suggest that higher capital requirements for depositories will lead to regulatory arbitrage, causing banking activities to migrate to shadow banks. In our analysis, we have shut down this activity migration channel by assuming that the same financial innovation can be performed in both sectors. What we show is that this capital requirement gap may nonetheless be needed for the regulator to implement the socially optimal talent allocation. The following figure summarizes these findings:

Policy Implications

Our analysis has numerous policy implications. First, regulators need to be thoughtful about how to use regulatory instruments to encourage higher ethical standards in banking. Even though a higher standard leads to fewer ethical transgressions, it is not a “free lunch.” The costs are twofold: less innovation and migration of talent to an affiliated but distinct sector where ethical transgressions may be socially less costly or simply harder to detect.

Second, regulators need to explicitly consider the role of capital requirements in influencing ethical standards in banking. In our analysis, capital requirements play not only the usual role of diminishing the risk appetite of banks, but also lead to higher ethical standards.

Third, since the 2007-09 financial crisis, there has been considerable attention on regulating executive compensation in banking to change managerial behavior for the better. This has obvious challenges because it often requires regulators to “micromanage” compensation design, with all of the attendant information impediments. Our analysis shows that a potentially better alternative for regulators to influence managerial behavior may be more oblique than direct regulation of compensation. Specifically, capital requirements will induce a voluntary change in compensation practices that will improve ethical standards in banking without requiring regulators to have all the information needed for efficient compensation design.

Fourth, regulators should realize that it may be socially optimal to impose higher capital requirements on depositories than on shadow banks, even though the consequently higher ethical standards in depositories will cause talent to migrate to shadow banks. In other words, prudential regulation of depositories as well as shadow banks will have unavoidable labor market consequences in financial services.

Finally, the rich interaction among capital requirements, ethical standards, and talent competition and allocation highlighted by our analysis suggests that policy coordination is needed between consumer protection regulators and prudential regulators. If a consumer protection agency like CFPB in the U.S. pursues a policy of penalizing mis-selling of financial products, these penalties will deter innovation beyond the dampening effect of capital requirements. For any given capital requirement imposed by prudential regulators, CFPB penalties will cause depositories to innovate less and attract less talent than intended by prudential regulators, without coordination between the two types of regulators. With coordination, CFPB penalties may enable prudential regulators to achieve the targeted ethical standard in depositories (hence, the desired financial stability) with lower capital requirements.

References

In our paper, we develop new methodology to estimate arbitrage portfolios by using the information in observable firm characteristics for both abnormal returns and betas (and smart beta risk premia). Our methodology gives maximal explanatory power to risk-based interpretations of characteristics’ predictive power before attributing to mispricing. The method can accommodate investors’ learning, which may lead to time-varying predictive strength of the characteristics. When we apply our methodology to a large dataset of U.S. equity returns from 1968 to 2018, we find evidence of significant mispricing.

One of the fundamental goals of asset pricing research is to understand why some assets earn different rates of return on average compared to other assets. Is it because of differences in risk? What kinds of risk matter for determining expected returns on assets? How should these risks be measured? These are the kinds of questions investors are interested in answers to. The classical theories of asset pricing point to a clear answer to the first question: risk is the primary determinant of forward-looking expected returns. More precisely, in response to the second question, these theories say that it is the exposure to sources of systematic risk, i.e., risk that cannot be eliminated even in a well-diversified portfolio such as a large index fund, that drives average returns. In theoretical models, these risk exposures (commonly labelled “betas”) are typically known to the agents who collectively “determine” prices. As a consequence, if any variable is found to be correlated with future returns, the classical theories predict that it must because it is a proxy for some underlying exposure to a source of systematic risk. Turning to the measurement issue raised in the third question above, testing these theories for the U.S. stock market however, has proven to be quite a delicate task.

Characteristic-Based Returns: Alpha or Smart Beta?

ANDREAS NEUHIERL, Olin Business School, Washington University in St. Louis

In our paper, we develop new methodology to estimate arbitrage portfolios by using the information in observable firm characteristics for both abnormal returns and betas (and smart beta risk premia). Our methodology gives maximal explanatory power to risk-based interpretations of characteristics’ predictive power before attributing to mispricing. The method can accommodate investors’ learning, which may lead to time-varying predictive strength of the characteristics. When we apply our methodology to a large dataset of U.S. equity returns from 1968 to 2018, we find evidence of significant mispricing.

One of the fundamental goals of asset pricing research is to understand why some assets earn different rates of return on average compared to other assets. Is it because of differences in risk? What kinds of risk matter for determining expected returns on assets? How should these risks be measured? These are the kinds of questions investors are interested in answers to. The classical theories of asset pricing point to a clear answer to the first question: risk is the primary determinant of forward-looking expected returns. More precisely, in response to the second question, these theories say that it is the exposure to sources of systematic risk, i.e., risk that cannot be eliminated even in a well-diversified portfolio such as a large index fund, that drives average returns. In theoretical models, these risk exposures (commonly labelled “betas”) are typically known to the agents who collectively “determine” prices. As a consequence, if any variable is found to be correlated with future returns, the classical theories predict that it must because it is a proxy for some underlying exposure to a source of systematic risk. Turning to the measurement issue raised in the third question above, testing these theories for the U.S. stock market however, has proven to be quite a delicate task.
particular, empirical researchers in asset pricing have produced a slew of variables (mostly observable firm characteristics) that appear to predict future returns in the cross section, and these variables are difficult to explain as sources of systematic risk. Figure 1 illustrates this situation, below.

The central challenge that has emerged out of the empirical analysis is that portfolios formed by using these predictors appear to carry too little risk to rationalize their high average returns, where realized average returns are typically viewed as a proxy for expected future returns. This has led to a substantial debate in the academic literature on how to interpret a candidate return predictor. Is it predicting average returns because it measures risk (beta)? Or is it predicting average returns because some assets are priced incorrectly in the market (alpha)? Figure 2 illustrates the two competing channels. Since the two interpretations have vastly different consequences, researchers have tried to develop methods to disentangle alpha and beta in the data.

The most commonly used approach to disentangle risk from mispricing is the so-called double sorting approach. Assets are sorted into portfolios based on lagged beta estimates and firm characteristics. Long-short portfolios made of portfolios with similar beta exposure, but different levels of characteristics should measure the pure return due to the characteristics. Similarly, long-short portfolios made of portfolios with similar characteristics, but different levels of beta exposure should measure the pure (systematic) risk premia. Despite the great intuitive appeal of this approach, it can be shown very easily that it systematically misclassifies risk as mispricing and will erroneously detect alpha, even when there is no mispricing at all. This commonly happens in situations in which characteristics are cross-sectionally correlated with betas. Table 1 shows a simple simulation to illustrate this phenomenon. In this simulated example, beta is identical to the observed characteristic and there is no mispricing at all. Along the rows of Table 1, the beta exposure is constant, and the characteristic value is varying, whereas the characteristic is constant across columns and the beta is varying. We see a large and significant spread in average returns as one varies the characteristics, but almost no spread in average return as one varies beta. This finding erroneously suggests that characteristics are useful to predict returns above and beyond beta, so using the information in the characteristic to build portfolios leads to alpha, i.e., abnormal returns without incurring additional risk. Obviously, this conclusion is false, since the characteristics and the betas are chosen to be identical and thus contain exactly the same information about average returns. Since the workhorse approach is plagued by these shortcomings, the goal of our paper is to develop an approach that can reliably disentangle the information in characteristics and correctly attribute them to risk and mispricing. Our approach combines two fundamental ideas in financial modelling. “Fundamental beta” on the one hand and latent factor models on the other.

Table 1

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Low</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>10-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>0.24</td>
<td>0.25</td>
<td>0.29</td>
<td>0.34</td>
<td>0.26</td>
<td>0.15</td>
<td>0.34</td>
<td>0.25</td>
<td>0.18</td>
<td>0.14</td>
<td>0.23</td>
<td>-0.03</td>
</tr>
<tr>
<td>2</td>
<td>0.36</td>
<td>0.40</td>
<td>0.32</td>
<td>0.39</td>
<td>0.36</td>
<td>0.37</td>
<td>0.39</td>
<td>0.29</td>
<td>0.32</td>
<td>0.25</td>
<td>0.47</td>
<td>0.06</td>
</tr>
<tr>
<td>3</td>
<td>0.42</td>
<td>0.37</td>
<td>0.41</td>
<td>0.42</td>
<td>0.41</td>
<td>0.46</td>
<td>0.28</td>
<td>0.47</td>
<td>0.46</td>
<td>0.48</td>
<td>0.46</td>
<td>0.09</td>
</tr>
<tr>
<td>4</td>
<td>0.45</td>
<td>0.46</td>
<td>0.45</td>
<td>0.45</td>
<td>0.36</td>
<td>0.42</td>
<td>0.44</td>
<td>0.39</td>
<td>0.45</td>
<td>0.58</td>
<td>0.47</td>
<td>0.02</td>
</tr>
<tr>
<td>5</td>
<td>0.47</td>
<td>0.48</td>
<td>0.37</td>
<td>0.48</td>
<td>0.45</td>
<td>0.52</td>
<td>0.51</td>
<td>0.47</td>
<td>0.48</td>
<td>0.44</td>
<td>0.47</td>
<td>-0.01</td>
</tr>
<tr>
<td>6</td>
<td>0.53</td>
<td>0.56</td>
<td>0.51</td>
<td>0.61</td>
<td>0.43</td>
<td>0.47</td>
<td>0.47</td>
<td>0.56</td>
<td>0.59</td>
<td>0.47</td>
<td>0.56</td>
<td>-0.01</td>
</tr>
<tr>
<td>7</td>
<td>0.58</td>
<td>0.52</td>
<td>0.54</td>
<td>0.61</td>
<td>0.58</td>
<td>0.60</td>
<td>0.60</td>
<td>0.51</td>
<td>0.56</td>
<td>0.71</td>
<td>0.59</td>
<td>0.07</td>
</tr>
<tr>
<td>8</td>
<td>0.59</td>
<td>0.56</td>
<td>0.54</td>
<td>0.56</td>
<td>0.63</td>
<td>0.60</td>
<td>0.46</td>
<td>0.66</td>
<td>0.70</td>
<td>0.62</td>
<td>0.56</td>
<td>0.01</td>
</tr>
<tr>
<td>9</td>
<td>0.67</td>
<td>0.67</td>
<td>0.68</td>
<td>0.59</td>
<td>0.71</td>
<td>0.66</td>
<td>0.64</td>
<td>0.71</td>
<td>0.69</td>
<td>0.67</td>
<td>0.70</td>
<td>0.03</td>
</tr>
<tr>
<td>High</td>
<td>0.78</td>
<td>0.78</td>
<td>0.74</td>
<td>0.68</td>
<td>0.83</td>
<td>0.75</td>
<td>0.74</td>
<td>0.85</td>
<td>0.78</td>
<td>0.80</td>
<td>0.86</td>
<td>0.08</td>
</tr>
<tr>
<td>10</td>
<td>0.78</td>
<td>0.78</td>
<td>0.74</td>
<td>0.68</td>
<td>0.83</td>
<td>0.75</td>
<td>0.74</td>
<td>0.85</td>
<td>0.78</td>
<td>0.80</td>
<td>0.86</td>
<td>0.08</td>
</tr>
</tbody>
</table>

***p <0.01; **p <0.05, *p <0.1
other. Fundamental betas aim to tie risk to observable firm characteristics. Whereas latent factor estimation refers to an unsupervised learning technique that is often used in financial econometrics to determine factor portfolios that explain the data well. Our approach is specifically developed for the asset pricing application. If we apply it to the example above, we correctly find no mispricing. More generally, our approach can reliably distinguish between risk (beta) and possible mispricing (alpha).

Another shortcoming of the double sorting approach, which is implicit in the example on the previous page, is that it can only be applied when the number of candidate predictors is small. The introductory figures, however, show clearly that there are now very many candidate predictors that should be considered simultaneously. John Cochrane, a past president of the American Finance Association, dubbed this the multidimensional challenge and called for the development of new methods—our estimator responds to this multidimensional challenge.

The approach gives the characteristics maximal explanatory power for risk premia before we attribute any explanatory power to alpha. By applying this procedure, we can build arbitrage portfolios, i.e., portfolios that have no exposure to any source of systematic risk but have high predicted returns. Studying the return properties of the arbitrage portfolio allows us to answer the fundamental question illustrated in Figure 2: Do firm characteristics predict returns because they proxy for (unknown) underlying factor loadings or because they contain information about mispricing. If firm characteristics are merely proxies for risk and there is no mispricing, then the excess returns of the arbitrage portfolio should be close to zero. However, if we find large returns to the arbitrage portfolio, we can conclude that the characteristics contain information above and beyond risk and can predict returns because there is mispricing.

To test this hypothesis, we apply our approach to a large sample of stocks listed on the U.S. exchanges for the time period between January 1968 and December 2018. Our data set contains monthly stock returns from the Center for Research in Securities Prices (CRSP) and the most used predictors from the academic literature. These are precisely the predictors for which there has been no consensus among researchers if they predict because of risk or mispricing.

A key innovation of our procedure is that we can estimate betas and alphas over a short time window. This allows us to feature time-varying betas and the relationship between characteristics and returns can change over time. For example, some predictor might be related to mispricing in the early part of the sample and is arbitaged away as more market participants are learning about it. We implement the estimation of the arbitrage portfolio using a rolling window approach in which we estimate the arbitrage portfolio weights $\hat{w}$ over the past 12 months and the compute the returns over the next month, Figure 3:

We then repeat this procedure every month and then study the returns over the full sample period. Table 2 below summarizes the results. The number of “Eigenvectors” is the number of factors we estimate.

Table 2 shows economically large excess returns to building an arbitrage portfolio using our methodology. If characteristics only contained information about risk exposures, we should observe excess returns close to zero for the arbitrage portfolio since the arbitrage portfolio is constructed in a way that forces all the risk exposures to be zero. From the results in Table 2 we can conclude that characteristics not only contain information about risk exposures, but also carry significant information about mispricing. A common concern is that we may have missed a possible source of risk when we construct the arbitrage portfolio. We therefore confront the arbitrage portfolio with a battery of popular factor models, such as the Fama-French models and various extensions. However, none of these methods can explain the returns of the arbitrage portfolio. We consistently find monthly alphas greater than 1%, and possibly equally important, known factor models can explain at the most one third of the variation in the returns to the arbitrage portfolio. We have therefore robustly established
that firm characteristics not only contain information about, but that there is a sizeable mispricing component that can be uncovered from them.

To shed more light on the properties of the arbitrage portfolio, we also examine its behavior in the time series. Importantly, we find that its returns are not systematically different during recessions as illustrated in the Figure below. The Figure shows the cumulative returns of the arbitrage portfolio (black) and the U.S. stock market (red). The areas shaded in gray are recession periods. From the Figure (above) we can clearly see that the returns of the stock market are lower during recessions, but the returns of the arbitrage portfolios are not. In a further analysis, we find that the returns of the arbitrage portfolio have declined slightly over time. This suggests (Figure below, on the previous page) that markets have become more efficient and using information on observable characteristics does not lead to equally large abnormal returns as in earlier periods. It is however noteworthy that the trend we document is relatively small and that returns even towards the end of the sample are still economically meaningful.

An important feature of our approach is the explanatory power of characteristics for both alpha and beta can ebb and flow over time. We find that some characteristics have relatively consistent relationship to beta, in particular past returns, firm size, return volatility and operating leverage. The characteristics that are important for alpha show a more transitory relationship except for total assets, sales to total assets and market capitalization. Interestingly, book-to-market equity (value) has a very transitory relationship with alpha when we control for the full set of characteristics.

Concluding Thoughts
Our analysis has numerous policy implications. First, regulators need to be thoughtful about how to use regulatory instruments to encourage higher ethical standards in banking. Even though a higher standard leads to fewer ethical transgressions, it is not a “free lunch.” The costs are twofold: less innovation and migration of talent to an affiliated but distinct sector where ethical transgressions may be socially less costly or simply harder to detect.

A large literature in asset pricing has reached vastly differing conclusions on how to interpret characteristics that predict returns in the cross section. One of the main reasons that these researchers have reached conflicting conclusions is because they employed an approach that has several shortcomings. In particular, it cannot reliably distinguish between risk and mispricing. A further shortcoming is that it can only be applied to a small number of predictors, a clear deficiency in the age of big data.

In our research we propose new econometric methodology to reliably separate risk from mispricing even in a setting with many candidate predictors simultaneously. Contrary to other methods it can be applied over a short time span using a large cross section. This allows us to accommodate factor momentum or alpha decay, which may occur through learning.

When applied to U.S. equities over the period from 1968 to 2018, we find that characteristics carry significant information about mispricing. While some characteristics are consistently related to risk exposures, the relationship to mispricing is more transitory, which underscores the importance of studying such relationships over short time windows.

Our work has a number of practical implications for portfolio managers. If a portfolio manager aims to use the information in firm characteristics to build alpha portfolios, they should use a larger number of characteristics jointly, rather than focusing on single characteristics at a time. Moreover, it is critically important to account for time-variation. Sources of alpha change quickly over time and it is therefore important to frequently use predictive models than can detect which information may currently carry alpha and which sources of alpha have dissipated. Of the 61 characteristics we study, firm size, sales to assets and operating leverage are related to alpha, but due to the strong time variation, investors should not only focus on these.
The study of abnormal stock returns—the amount by which the actual returns on a stock exceed those predicted by asset pricing models based on the risk in the stock—is of great interest to investors and fund managers. The existence of such abnormal return is dubbed an “anomaly” in finance research. An extensive academic research literature has successfully detected various anomalies from financial statements, firm characteristics, and stock price trends by leveraging modern computing power and big data. Of course, to the extent that investors are aware of abnormal returns, we would expect them to trade on them. Trading volume measures how active investors are in trading based on those anomalies. This trading can affect the magnitude and persistence of these anomalies. An important economic question is: How does trading volume affect abnormal returns in terms of the nature and magnitude of the effect?

We find that volume amplifies mispricing (see Han, et. al., 2022, for a more formal analysis). The stock market is a place where investors trade to find the true value of assets, but there is always mispricing. Our finding is that the mispricing is concentrated in high-volume stocks. The intuition is that, on the one hand, high volume implies extensive trades from many investors who may disagree on the correct valuation, and consequently, there is likely mispricing during the trading process. On the other hand, if investors have little interest in trading a particular asset, it is likely that most investors believe that the current valuation is correct. The theoretical model of Atmaz and Basak (2018) may be used here to explain our finding if we interpret trading volume measures investor disagreement and mispricing measures investor expectation bias for asset pricing to be away from their fundamental values. In equilibrium, after good news, optimistic investors vindicate their beliefs and become relatively wealthier via their investments, which in turn increases their bias to be more optimistic. This implies that investor disagreement has an amplification effect on the bias, consistent with our empirical results.
Our finding is that the mispricing is concentrated in high-volume stocks. The intuition is that, on the one hand, high volume implies extensive trades from many investors who may disagree on the correct valuation, and consequently, there is likely mispricing during the trading process.

**Research design**

First, we need a good measure of mispricing. Following Stambaugh, Yu, and Yuan (2015), we use a representative measure based on 11 accounting fundamentals and stock price trends. They cover broad aspects of firms’ performance, management, financial, and accounting decisions. For example, our measures cover firms’ profitability with return on asset and gross profit, corporate decisions with net stock issues, accounting decisions with accrual, operating decisions with net operating assets, financial decisions with financial distress, and stock price trends with the stock price momentum in the past year.

With the information in those 11 variables, we compute an overall mispricing score for each stock. The mispricing score ranges from one to 100. A score of one indicates the most underpriced, and 100 indicates the most overpriced each month. We measure trading volume as the turnover of shares, which is defined as the number of shares traded divided by the number of shares outstanding.

Each month, we form quintile portfolios based on our mispricing score using all stocks. We also create quintile portfolios based on the trading volume. We take the intersection of the two independent sorts to construct 25 mispricing-volume portfolios. We then follow the excess returns of these 25 portfolios in the following month and track their value-weighted excess returns. We buy the underpriced stocks and sell the overpriced stocks for each volume group, and the long-minus-short return is referred to as UMO.

To remove systematic risks involved in our trading strategies, we regress the excess returns of our UMO portfolios on common asset pricing factors. We apply two leading factor models: The Fama and French (2015) five-factor model, and the Hou, Xue, and Zhang (2015) Q-factor model. The former includes the market factor, a size factor, a value factor, an investment factor, and a profitability factor. The latter excludes the value factor and has a different profitability factor. More specifically, the Fama and French five factor model propose to adjust risk exposure using the following model:

\[ R_i = \alpha + \beta_{MKT} + \beta_{SMB} + \beta_{HML} + \beta_{CMA} + \beta_{RMW} + \epsilon_i, \]

where \( R_i \) is the return of a portfolio in excess of the risk-free rate, \( \beta_{MKT} \) is the market excess return, \( SMB \) is the size factor that captures the return difference of small and big firms, \( HML \) is the value factor that captures the return difference between stocks with high book-to-market ratio and low book-to-market ratio, \( CMA \) is the investment factor that captures the return difference between stocks with high book-to-market ratio and low book-to-market ratio, \( RMW \) is the profitability factor that captures the return difference between stocks with high profitability and low profitability, the \( \beta_s \) are the exposures on these risk factors and the \( \alpha \) is the return left after removing systematic risk exposures, and it represents the abnormal returns. Theoretically, the Hou, Xue, and Zhang model differs from the Fama and French model with a different profitability factor based on quarterly return on assets.

**Results**

Figure 1 plots the returns of UMO in each volume group. It shows clearly that buying underpriced stocks and selling overpriced stocks works only among stocks with high trading volume. Among the stock with low trading volume, the return of UMO is 0.27%; among stocks with high trading volume, the return of UMO is 1.17%. In addition, the return of UMO increases monotonically as trading volume increases. In the same panel, we observe the FF alpha is 1.1%, economically significant, for UMO among high volume stocks, and is only 0.13% among low volume stocks. Consistent with this result, the Q-Alphas are 0.87% and 0.12%, respectively.

Table 1 (next page) presents a more detailed analysis. Panel A reports the performance of the entire double-sorted 25 portfolios. For a given level of mispricing, the average returns increases monotonically with trading volume, and this pattern is unambiguous for the UMO. We also find that anomaly returns come from both the short and long legs. Panel B presents the t-stat values on the statistical significance of those results. These values are typically about three or four standard deviations away from zero, indicating that the results of Panel A are highly statistically significant.

**Discussions**

The trading volume is a multi-faceted variable related to liquidity, volatility, and disagreement. The higher the trading volume, the stronger the liquidity. When liquidity is strong, it is cheap for arbitragers to trade. The bid-ask spread does not deter the arbitragers from trading. We shall not observe significant mispricing. Hence, liquidity alone does not explain our results.

Figure 1: Anomaly Returns by Volume Groups

This figure plots the long-minus-short return of anomalies by volume groups. Group 5 has the highest trading volume. The UMO refers to the underpriced-minus-overpriced spread portfolio.
Volatility is also positively related to trading volume. High volatility indeed deters arbitrageurs from trading. As traders face pressures from the annual review, they will be cautious when they trade stocks that have large volatility. Since stock prices may move against what they believe, and the traders may not be able to meet their performance goals for the year. So, volatility may drive our results, and we must control for the confounding effect from the volatility. To that end, we regress trading volume on stock return volatility and using the residual from that regression to replace trading volume in its place. This residual captures about half of the abnormal returns. So volatility alone is not the whole story.

Here is our economic explanation. Using analyst price targets to compute expected returns, we find trading volume is related to the dispersion of the expected return among analysts. The views can produce large underpricing or overpricing. Suppose we have two stocks. One stock has a mispricing score of 70 so overpricing with small dispersion of expected returns, and the other stock has a mispricing score of 70 but with large dispersion of expected returns. In the latter, some investors hold extreme views, as overpricing starts to correct, extreme pessimistic investors win, and extreme optimistic investors lose. Since the buyers of these overpriced stocks are hurt and may not recover to support stock price, we observe large price declines in the second stock. In our regression results that use this disagreement measure in the place of the trading volume, we can see a similar return amplification effect and render the volume-amplification effect insignificant.

Conclusion
Based on Han et al. (2022), we show empirically that there is the heterogeneity of the volume-return relation across stocks: it is positive among underpriced stocks and negative among overpriced stocks. The fact that mispricing is concentrated in high-volume stocks provides valuable information to investors and fund managers for making optimal portfolio decisions. In addition, our research is also of interest for investment houses that wish to engineer new products. To the extent that volume matters significantly, products capturing volume and mispricing of the financial industry, similar to stock price momentum and quality, are likely to be valuable for hedging and as well as investments.

References
Author Biographies

Brett Green (page 6) is a financial economist and an Associate Professor of Finance at the Olin Business School. His research spans a variety of fields including corporate finance, microeconomics, contract theory, and development economics. He received a PhD in economics from Stanford Graduate School of Business. While at Stanford, he also completed a master’s degree in financial mathematics. He earned a bachelor’s of science in engineering and economics from Duke University.

Richard T. Thakor (page 14) is an Assistant Professor of Finance at University of Minnesota, Carlson School of Management. He is a Research Affiliate at the MIT Laboratory for Financial Engineering (LFE), and is a Research Affiliate at the WFA Center for Finance and Accounting Research (WFA-CFAR) at Washington University in St. Louis. He received his PhD in financial economics from MIT.

Fenghua Song (page 20) is an Associate Professor of Finance and Joseph F. Bradley Fellow of Finance at Smeal College of Business, Pennsylvania State University. He conducts research in the areas of banking and corporate finance. He received his PhD in finance from Washington University in St. Louis.

Andreas Neuhierl (page 26) is an Assistant professor of Finance at the Olin Business School. His research interests are in asset pricing, machine learning and fintech. He received his PhD in finance from Northwestern University.

Guofu Zhou (page 34) joined Washington University in 1990 and has been teaching and conducting research at the Olin Business School ever since. His research interests are empirical asset pricing, investor sentiment, big data, asset allocation, predictability, behavioral finance, household finance, mutual funds, hedge funds, and Chinese financial markets. He received his PhD in finance from Duke University.

SAVE THE DATE

35th Annual Nick Dopuch Accounting Conference
Fall 2022
Examine topics such as firm volatility, accounting conservatism, and pricing/mispricing.

SAVE THE DATE
19th ANNUAL CORPORATE FINANCE CONFERENCE
Nov. 18-19, 2022
- Debate and discuss ideas
- Explore real-world issues with corporate executives
CFAR Practicum

The creative application of advanced analytical tools.

Cutting-edge business strategy. State-of-the-art analytical tools. Intellectual property rights. Access to future talent. Washington University's student consulting teams—the next generation of finance and accounting experts—are ready to apply advanced analytics and actionable insights to your business challenges. Conducted through Olin Business School’s Wells Fargo Advisors Center for Finance and Accounting Research, Practicum projects are customized, hands-on, student led and faculty guided.

Confront challenge, create change.

Our students are informed by numbers, driven by principle. We challenge them to look beyond the bottom line and carefully consider what will make the biggest impact and do the most good.

Your Practicum project will combine the analytical perspective of talented students in our Master of Science in Finance program with the expertise of our world-renowned finance faculty. Students will closely study your situation and employ a variety of analytical tools to offer solutions to your business challenge.

How it works

Step 1: Scope problems
A client liaison faculty member will manage scoping your project, maintaining a high level of client service and ensuring your student team communicates effectively and meets all project deadlines.

Step 2: Collect data
You will share appropriate data with your faculty-supervised team of students who sign a nondisclosure agreement to protect your confidentiality. They will then apply theory-based models to your business problems.

Step 3: Analyze information/data
Your student team will study the data, analyze the situation and draw conclusions to formulate recommendations to combat your business challenge.

Step 4: Report results
Students will share their insights and recommendations with you through video conference calls. You will own the intellectual property resulting from your Practicum project.

Recent Clients

Advisory Research Investment Management
CoverCress
Detalus
Edward Jones
Neocova
NuPeak Therapeutics
Opera Theatre of Saint Louis
RGA International Corporation
St. Louis Symphony Orchestra
St. Louis Trust Company

FINANCE AND ACCOUNTING RESEARCH
THE WELLS FARGO ADVISORS CENTER FOR FINANCE AND ACCOUNTING RESEARCH

“WFA-CFAR finance consulting projects bring together some of America’s most distinguished finance research faculty and gifted graduate students to collaborate with business partners to solve complex problems facing organizations.”

ANJAN THAKOR, DIRECTOR OF DOCTORAL PROGRAMS & CFAR, JOHN E. SIMON PROFESSOR OF FINANCE, WASHINGTON UNIVERSITY

5 to 10 students per team
2 project time frames
3 STEM-designated programs
5 to 10
14-week projects
fall and spring
Master of Science in Finance—Quantitative
Master of Science in Finance—Wealth and Asset Management
Master of Science in Business Analytics—FinTech Analytics

WFA-CFAR PRACTICUM PROGRAM OVERVIEW

intellectual property belongs to your organization

$12,000 standard consulting fee

Find out more.

olin.wustl.edu/cfar
Timothy G. Solberg, CFA
Professor of Practice in Finance
314-935-7270
solbergtg@wustl.edu

olin.wustl.edu/cfar

42
Practicum Projects with CFAR

"The Finance and Accounting Practicums vary from corporate finance research, fintech coding, highly quantitative analysis of trading plans, metrics of wealth management product offerings to writing business commercialization plans for biotech and agtech start-ups in the Cortex center with WashU researchers. CFAR also conducts non-profit organizational practicums to benefit our community. The goal is to present professional-level consulting results to our corporate sponsors in the course of a fourteen week semester. Students conduct analysis and research under the supervision of an expert professor in the consulting topic. CFAR’s clients are receiving enthusiastic student efforts in sophisticated and practical business solutions. We are expanding the hands-on experiential learning of the CFAR Finance Practicum to add Fintech, with students from Olin’s rapidly expanding Masters of Science in Analytics program, which has tripled in size from 30 to over a hundred students in three years. CFAR is also expanding geographically to include firms nationally and in China. We are very excited as we develop more corporate and non-profit client relationships."

Professor Timothy G. Solberg, CFA
Professor of Practice in Finance and Academic Director of the Corporate Finance and Investments Platform

We have partnered with the Center for Finance and Accounting Research over the last several years to offer students the opportunity to tackle current strategic and operational projects with industry-specific quantitative and technical components for the firm. Each year our student team has exceeded our expectations thinking outside the box to come up with creative solutions to help move our business forward."

Julie Winkler,
Investment Advisory–Strategic Initiatives
Edward Jones

Olin Business School Internationally recognized for scholarship and research, Olin faculty members help you turn business problems into practical applications. Their far-reaching research addresses priority issues and emerging business challenges, producing timely and relevant material that functions far beyond the classroom—for sustainable improvement and growth for companies. Through the efforts of Olin’s faculty-led research centers such as WFA-CFAR, an organization’s top priorities and business challenges can drive new areas of study. To discuss offering your organization’s data for a new project with Olin’s world-renowned finance and accounting faculty, contact WFA-CFAR Program Manager Kristen Jones at 314-935-4179 or kristen.jones@wustl.edu.

Full-time Finance Faculty

Doniz Aydin
Assistant Professor of Finance
PhD, Stanford University
Research interests: finance, empirical macroeconomics and applied microeconomics

Taylor Begley
Assistant Professor of Finance
PhD, University of Michigan
Research interests: empirical investigation of financial contracting models

Jian Cai
Senior Lecturer in Finance
PhD, Washington University in St. Louis
Research interests: corporate finance, corporate governance, executive compensation, career concerns, financial intermediation, financial institutions and empirical asset pricing

Charles J. Camy
Senior Lecturer in Finance
PhD, Stanford University
Research interests: capital structure, financial innovation, employee stock options

Jeremy Degenhart
Professor of Practice in Finance
Research interests: venture capital, private equity

Jason R. Donaldson
Associate Professor of Finance
PhD, London School of Economics
Research interests: contract theory, corporate finance theory

Philip H. Dybvig
Booth’s Blawshares Professor of Banking and Finance
PhD, Yale University
Research interests: banking, corporate finance, asset pricing, banking, financial markets, fixed-income securities, industrial organization, portfolio management

Nicole Garibau
Professor of Finance
PhD, Stanford University
Research interests: asset pricing, finance/investments, financial economics, option pricing

Armando R. Gomes
Associate Professor of Finance
PhD, Harvard University
Research interests: corporate finance, mergers and acquisitions, corporate governance, economic theory

Radhakrishnan Gopalan
Professor of Finance and Academic Director of the IIT-Bombay-Washington University Executive MBA Program
PhD, University of Michigan
Research interests: corporate finance, corporate governance, emerging-market financial systems, mergers and acquisitions, corporate restructuring and entrepreneurial finance, household finance

Todd Gurney
Professor of Finance, Area Chair and Academic Director of GMP
PhD, Massachusetts Institute of Technology
Research interests: corporate governance, empirical methods, mutual funds

Brett Green
Associate Professor of Finance
PhD, Stanford University
Research interests: financial economics, information economics, corporate finance, contract theory, development economics, sports economics

Xing Huang
Assistant Professor of Finance
Research interests: behavioral finance, investor behavior, market efficiency, information acquisition, mutual funds, household finance, asset pricing

Ohad Kadan
H. Frederick Hagemann, Jr. Professor of Finance and Vice Dean for Education and Globalization
PhD, The Hebrew University of Jerusalem
Research interests: corporate finance, asset pricing, market microstructure, economics of information and game theory

(continued on next page)