



GMO Quarterly Letter

February 20, 2018

by Ben Inker

of GMO

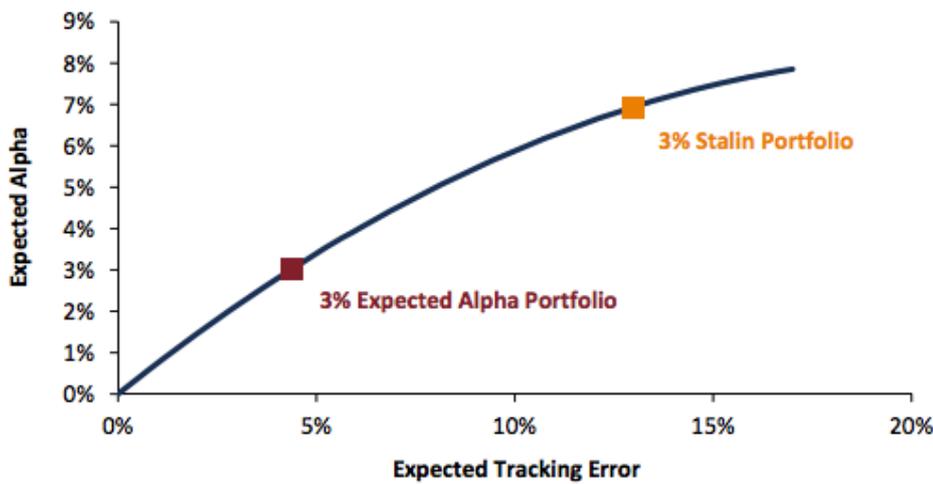
In Jeremy Grantham's 2017 third quarter letter, he posed a question: What should you do if you are tasked with managing Stalin's pension portfolio? The specific scenario he painted was that Stalin has ordered you to achieve inflation plus 4.5% on his pension assets over the next decade. If you succeed, you get a nice dacha on the Black Sea and a pension of your own. If you fail, you get shot. It is an entertaining problem set-up, and for the more quantitative among us, leads to a refreshingly simple utility function – maximize the probability of meeting or beating the target return. But if Stalin actually had any intention of living off of his pension, it is a lousy ultimatum to give his chief investment officer (CIO). It completely ignores the impact of magnitudes, and magnitudes matter. Barely missing the goal and achieving 4.4% above inflation is importantly different than returning 15% below inflation, and for Stalin, they are the same – failure. But Stalin may be on to something... just not quite the way he originally intended.

It is hard to imagine why any individual or institution should want an overall portfolio managed in as high-risk a fashion as Stalin's pension implies, but it might make sense to hire a number of managers who do act that way. While I admit that I haven't come across any managers who actually think about things quite the way Jeremy does in posing the Stalin problem, there is a class of managers who manage their portfolios as if they were solving that problem. They are the concentrated, high tracking error, "high conviction" managers that have gained a significant following, particularly in the endowment and foundation universe. So should you hire that type of manager? It depends. Hiring such managers only makes sense if you think you are good at identifying talented active managers, of course. But just as important and somewhat less intuitively, it also requires that you are significantly better than the average CIO at avoiding the performance chasing that traditionally accompanies the firing and hiring of managers. If (and only if) you can meet those two requirements, a portfolio built from a number of Stalinesque pieces will give you higher returns without much downside in the form of greater volatility or tracking error. If you can't meet both requirements, such aggressive managers are more likely to hurt you than help. In GMO's asset allocation group, we historically haven't run "Stalin-style" portfolios, but for clients who believe they meet the requirements to handle them well, we would be happy to discuss whether we could put together a portfolio that would make sense for them.

What does Stalin make you do?

I'm going to change Jeremy's framing of the Stalin problem slightly to make the analysis a little simpler. Rather than frame it as a total return, I'm going to look at it in relative return terms versus an ownable benchmark.¹ Let's frame the Stalin problem as maximizing the likelihood of outperforming a benchmark by at least 3%. While 3% alpha doesn't seem like an exceedingly high target level for a "Stalinesque" portfolio, the reality is that under almost all circumstances, the portfolio that maximizes that probability will have an expected alpha significantly higher than 3%. This may sound a little weird, but it becomes clearer when you think about it. If you run a portfolio targeting an expected alpha of 3%, you should have a roughly 50% chance of achieving 3% or higher alpha. If you run a portfolio targeting an expected alpha of 7%, your chance of achieving or exceeding 3% alpha is almost certainly higher than 50% – you could underperform your alpha target by up to 4% and still achieve 3% alpha. The probability of achieving at least a target return is a function of both your expected alpha and your volatility, where higher alpha is good and higher volatility is bad. Exhibit 1 shows an efficient frontier with a "normal" and Stalin version of achieving 3% alpha.²

Exhibit 1: Traditional and Stalin Portfolios

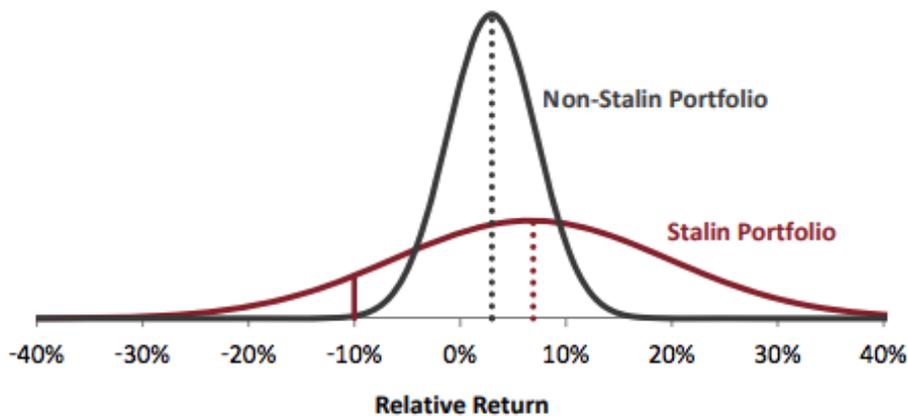


Source: GMO

The portfolio to the lower left is the “normal” way of thinking about achieving 3% alpha. It is the lowest tracking error portfolio that has an expected alpha of 3%. It has a tracking error of 4.4% and an information ratio (IR) of 0.68. The portfolio to the upper right is the Stalin version – the portfolio that maximizes the probability of achieving at least 3% alpha. It has an expected alpha of 6.9%, and a tracking error of 13% for an IR of 0.53.

Let’s first spend a moment thinking about why you wouldn’t want to manage your entire portfolio like the Stalin portfolio. The simple answer is that a 13% tracking error is an awful lot! Exhibit 2 shows the return distribution of the Stalin and non-Stalin versions of the 3% portfolio.

Exhibit 2: Distribution of 3% Stalin and Non-Stalin Portfolios



Source: GMO

Despite the high expected alpha, the Stalin portfolio has a 10.6% chance of underperforming by 10% or more, while the traditional portfolio has a 0.1% chance of doing that badly. Show me the CIO of an institution who wants to be in a position to have to explain underperforming its benchmark by 10% more than once a decade, and I’ll show you a CIO who is counting on being in the job less than 10 years.

Stalin as a piece of the portfolio

So let’s assume you do not want to Stalinize your whole portfolio. That doesn’t mean Stalin-esque pieces can’t be used to build a less scary overall portfolio. Let’s imagine you are an institution looking to outperform your benchmark by at least 1% over time. You will do this by hiring 20 outside managers who are all trying to beat their respective benchmarks.³ You’re pretty good at hiring talented managers, with a 50% hit rate. The combination of trading costs and management fees for the portfolios is 1%. How should you go about achieving 1% excess return? One way would be to tell your managers to shoot for 3% alpha. The 50% of your managers with skill will average 3% alpha. The 50% with no skill will average -1% due to costs, and therefore your portfolio will achieve 1% alpha on average. Your overall tracking error will be 1%, for a very nice IR of 1.0.⁴ But let’s say that instead of the “sensible” thing of aiming for your target return, you instead tell your managers to pretend you are Stalin, and to maximize their probability of achieving 3% alpha. Those portfolios will be meaningfully less efficient than the 3% expected alpha portfolios – in my example their IR will fall from around 0.68 to around 0.53. Moreover, let’s assume that the costs of the

strategy go up as well, so that the losers will lose by 1.5% instead of 1%.⁵ Despite those handicaps, the overall portfolio expected alpha more than doubles from 1% to 2.7%, and the IR is only slightly lower at 0.9. Table 1 shows the results in an easier to read form.

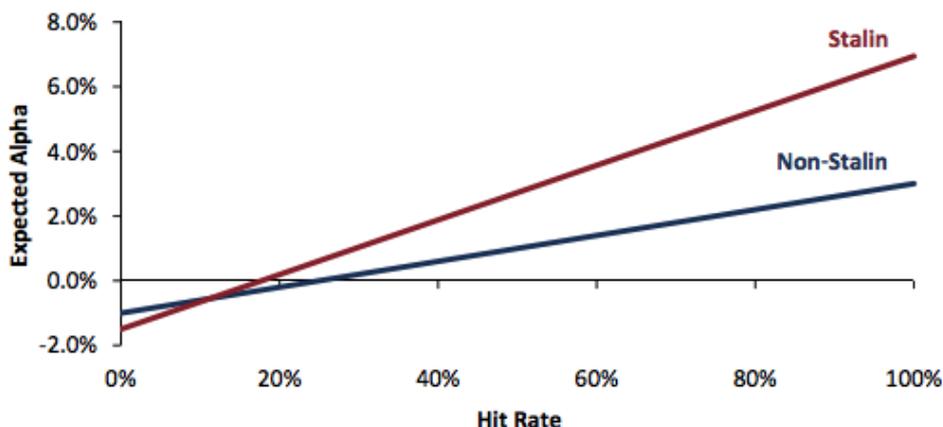
Table 1

	Non-Stalin	Stalin
Expected Alpha	1.0%	2.7%
Expected Tracking Error	1.0%	2.9%
Information Ratio	1.0	0.9

More than twice the alpha and almost the same IR? Sign

me up! But a 50% hit rate for your active managers is pretty high. Perhaps this strategy only makes sense for exceptionally good CIOs? Maybe the rest of us are better off sticking with the lower tracking error managers? Interestingly enough, that turns out not to be the case. Exhibit 3 shows what the alpha looks like for the two strategies as the assumed hit rate in manager hiring is altered.

Exhibit 3: Expected Alpha as a Function of Hit Rate



Source: GMO

For any hit rate where it makes sense to hire active managers in the first place, Stalin seems to dominate the non-Stalin strategy. Yes, you’ve got higher tracking error,⁶ but the alpha is much higher.

The verdict seems easy. Go Stalin or go home. If you believe you have a hit rate hiring talented active managers above 20%, hire Stalin managers. If you believe your hit rate is lower than 20%, index your portfolio.

Not so fast Stalin!

Having your managers go Stalin seems like an easy choice – as long as you think you have at least a 20% hit rate on finding talented managers, you’ll have higher alpha and generally a comparable IR versus hiring “normal” managers. But there are behavioral requirements that are a lot tougher than they appear at first blush. The first point is that if we assume that Stalin managers have a lower IR than non-Stalin active managers (this seems a very reasonable assumption), then basing hiring and firing decisions significantly on the basis of performance is more apt to give bad results. It is not enough to find talented managers. It is also critically important to time your hiring and firing of managers well.

Exhibit 4 shows a randomly generated return path of a talented Stalin manager with an alpha of about 5% and volatility of about 12%.⁷

Exhibit 4: Return Pattern of Talented Stalin Manager



Source: GMO

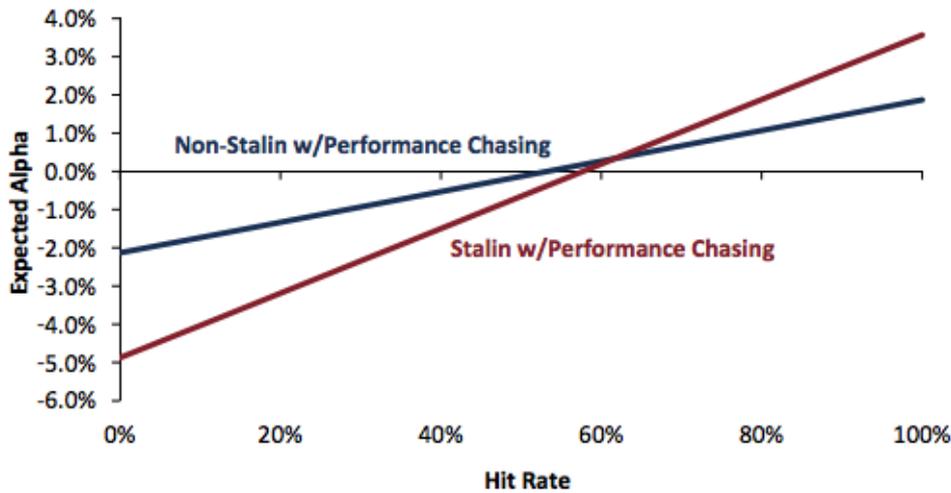
In the long run, it's a wonderful ride, with the manager leaving a client who hired them on day 1 with over 8 times as much money as the benchmark over 50 years. But it's easy to imagine a client actually losing money with this manager despite the strong alpha. If the manager didn't get on investors' radar screens unless they had 5% alpha since inception and at least a 3-year track record (why hire a Stalin manager who hasn't shown strong performance, after all?), it's not until year 15 of the manager's career that investors will really take notice. The due diligence then might take another year, with the hiring occurring in year 16. Over the next 10 years, the manager loses 27% relative to the benchmark. You'd have to somehow hold on for over 20 years to so much as break even. And this from a manager who had skill!

The combination of manager aggressiveness and a difficulty disentangling luck from skill can be a deadly one; in Appendix A, I will go over a couple of studies that demonstrate how the average institution can fall victim to performance chasing in its hiring and firing decisions. The short version is that managers get hired after periods of unsustainably good performance, and some of that is likely to reverse in future. Similarly, managers get fired after periods of unsustainably bad performance, some of which will also reverse.

So how does this affect our earlier analysis that suggested Stalinization was a good idea? That analysis effectively assumed a single decision at the start of the period to hire 20 managers and that all managers were then held for the duration. In reality, managers will be hired and fired along the way.⁸ Let's see how that alters things. Assume for the moment that the institution will fire 3 of their 20 managers each year and that managers will have been fired after underperforming by one standard deviation over 3 years.⁹

Why does this matter? Let's assume, as per the evidence from Appendix A, that those managers are fired after a period of bad luck, and 50% of the bad luck return would have reversed in time.¹⁰ And given that managers need to get replaced, let's assume that the managers who are hired are hired after a period of good luck, 50% of which will reverse in future. This means that by replacing the manager, we now can expect 0.5 standard deviations of net bad luck from our new manager instead of what would have been 0.5 standard deviations of good luck had we retained the original manager.¹¹ This changes the analysis significantly, as we can see in Exhibit 5.

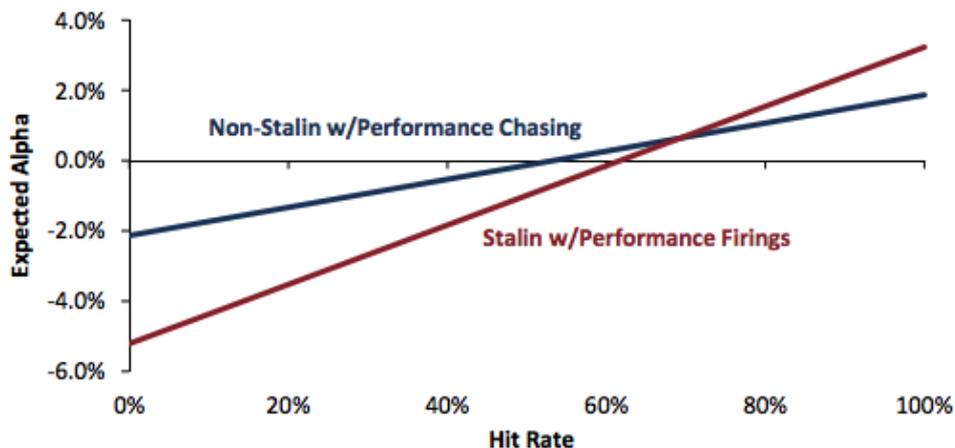
Exhibit 5: Expected Alphas Incorporating Performance Chasing



Source: GMO

It now takes a hit rate of around 50% to justify hiring non-Stalin active managers, and about a 60% hit rate for Stalin managers. But I'd submit that there is another problem for an institution that goes Stalin. The high tracking errors of the Stalin managers mean that "unacceptably bad" performance is simply more likely. So let's assume in the Stalin case, manager turnover is 10% higher. Exhibit 6 shows what happens when we incorporate these two effects.¹²

Exhibit 6: Expected Alpha as a Function of Hit Rate

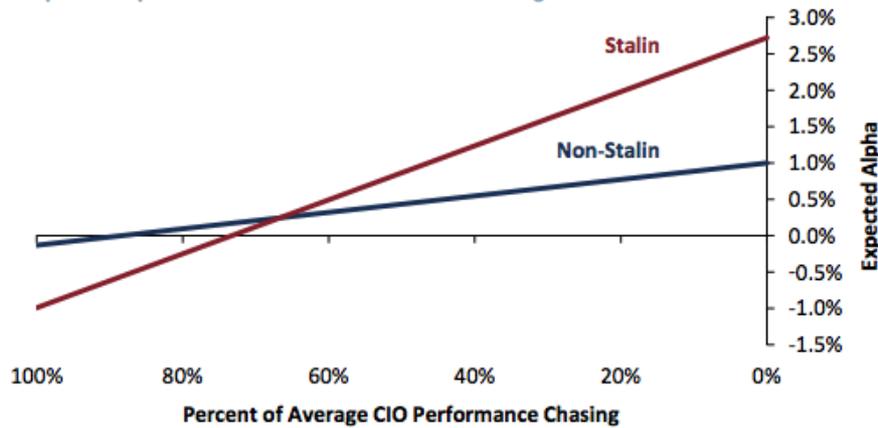


Source: GMO

Now, your hit rate has to reach 70% on picking talented managers just to break even in alpha terms with a portfolio built from non-Stalin managers, whereas before accounting for performance chasing, any hit rate above 20% left Stalin looking superior. And even when the hit rates get implausibly high at 80% or above, the gap in performance is not particularly large. At a highly implausible 90% hit rate, the expected Stalin alpha is 2.4% versus 1.5% for normal managers and the IR is 0.8 versus 1.5. Maybe the higher alpha Stalin version is better, but the CIO who went with the lower alpha and higher IR is going to sleep awfully well at night.

While I suppose it is possible to believe you are going to have a hit rate of 90% in your manager decisions, I'd suggest it's a bad idea to assume you have one that high. If you want to hire aggressive, high tracking error managers, the key element is ensuring you do not chase performance like the average institution. Let's assume for the moment our original 50% hit rate in hiring talented managers and look at how our expected alpha changes as we alter the size of the performance chasing problem in hiring and firing decisions.

Exhibit 7: Expected Alpha as a Function of Performance Chasing



Source: GMO

If you can get your chasing behavior down below about 60% of “normal,” Stalin starts to make sense.¹³

If you can hit the Holy Grail of 0% bias in your hiring and firing decisions, you can get the full initial benefit I showed way back in Exhibit 3 and earn more than twice the alpha with a very high IR. The slope of the Stalin line in Exhibit 12 is impressively steep. Every 10% you can improve your performance chasing earns you 0.4% of additional return. As you can see, the stakes when hiring nonStalin managers are lower. You can get away with a bigger bias in your decisions, and while lowering your bias is still helpful, it buys you only about 0.13% of additional return per 10% reduction in bias.

What does all this really mean? It is intuitively obvious that you need to believe you have skill in finding talented active managers in order to justify hiring them in your portfolio. It is less intuitive, but no less true, that the presence or absence of performance chasing in your manager decisions is just as important in determining whether you will achieve alpha in the end. How can you avoid performance chasing? It may be tempting to claim that if you properly account for the style biases of your managers you can get rid of a lot of the problem. In Appendix B, I show that the case of GMO’s Global Developed Equity Allocation composite makes it clear that things are seldom so simple.

The standard footnote to investment ads says “Past Performance Is Not Indicative of Future Results.” In reality, past performance almost certainly has something important to say about future results. A manager who has exhibited very strong historical performance is more likely than the average manager to be skilled, and a manager who has underperformed is less likely to be skilled. But despite that fact, the footnote is still excellent advice. A manager who has just done exceptionally well was probably benefitting from a trend that will reverse in time, and one that has done badly is likely to benefit as trends reverse as well. Hiring just after exceptionally good performance and firing just after exceptionally bad performance may well cost your portfolio dearly.

Conclusion

Managing your entire portfolio as if you were given Jeremy’s Stalin problem seems like a bad idea. It causes you to be extremely aggressive and it is hard to imagine the individual or institution that should really want that for their whole investment portfolio. Hiring a number of different aggressive managers who “go Stalin,” on the other hand, seems like an appealing idea at first blush. The benefits of diversification mean that your expected returns rise and your volatility is not too materially impacted. Digging a little deeper, however, you have to believe not only that you have decent skill in hiring talented managers, but that you are also good at avoiding the performance chasing most institutions have shown in their hiring and firing decisions. The performance chasing problem is arguably a bigger impediment to the advisability of “going Stalin” than the basic need for manager selection skill. Performance chasing means there is a pretty meaningful chance you can invest with a talented manager and still lose, and the scale of that loss will increase with the aggressiveness of the manager. Stalin-esque portfolios need to be treated with particular care.

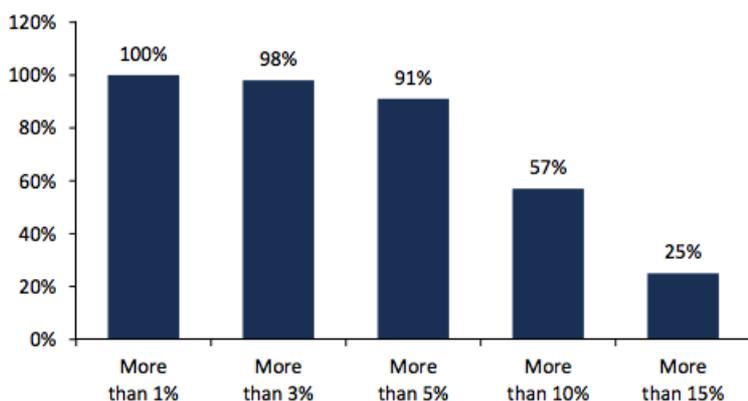
But that isn’t to say that they are inherently a bad idea. Over the last 15 or 20 years, a number of institutions, particularly some of the large endowments and foundations, have shown that they can do well hiring very aggressive managers. I don’t have the detailed data to decode how precisely they have done it, but I’m going to go out on a limb and guess that it has been due to a combination of three factors. First, they have had strong hit rates on their manager selection, in some cases well north of 50%. Second, they have done a good job of mitigating the performance chasing problem, both by not simply hiring managers after periods of extremely strong performance and by firing on the basis of factors other than extremely bad performance. And third, they have almost certainly benefitted from some good luck, which is never to be underestimated in investing.

Within GMO's Asset Allocation team, we have historically avoided managing Stalinesque portfolios, largely out of the fear that even if the portfolios themselves outperform over the long term, the clients in them will not, due to their timing in entering and leaving the portfolios. I still believe that is a greater risk for very aggressive portfolios than it is for portfolios with more moderate tracking error. But the interesting thing about these aggressive portfolios as we study them is that it isn't really the case that it is only special investment managers who can manage them. All managers have their best ideas, one way or another. Running a portfolio that consists entirely of those ideas may well be less efficient in an IR sense than running a more diversified portfolio, but if the manager has skill in the first place, the expected returns should be higher. I'd argue the bigger hurdle to wanting such Stalin strategies in your portfolio is that it is only a subset of clients who can hire and fire them without generating too much bad luck along the way. For clients who believe they can avoid the pitfalls, going Stalin makes sense. For our part, we would be happy to talk to such clients about what a GMO "Stalin portfolio" would look like and whether it might fit in their overall investment strategy.

Appendix A: Behavioral Bias in Manager Hiring and Firing Decisions

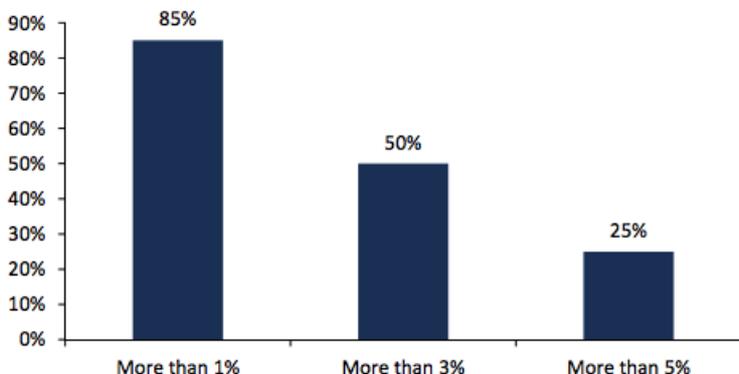
There are a number of studies that are relevant for understanding behavioral bias in hiring and firing decisions, but I'm going to focus on two that seem particularly relevant. The first, by Aaron Reynolds,¹⁴ looks at drawdowns relative to benchmarks and peer medians of 370 mutual funds that had 10-year track records of outperforming by at least 1%. Every manager in the group outperformed over the full 10-year period studied, but that doesn't mean they were outperforming the whole time. A full 81% of the managers had at least one 3-year period during the 10 years in which they underperformed their peer median. Exhibits 8 and 9 go into a little more detail on some of the interim underperformance of this strong-performing group.

Exhibit 8: Percent of Winning Managers Underperforming Style Benchmark over 1 Year



Source: Aaron Reynolds, "The Truth About Top-Performing Mutual Fund Managers," *AAL Journal*, July 2011.

Exhibit 9: Percent of Winning Managers Underperforming Style Benchmark over 3 Years of Annualized Performance



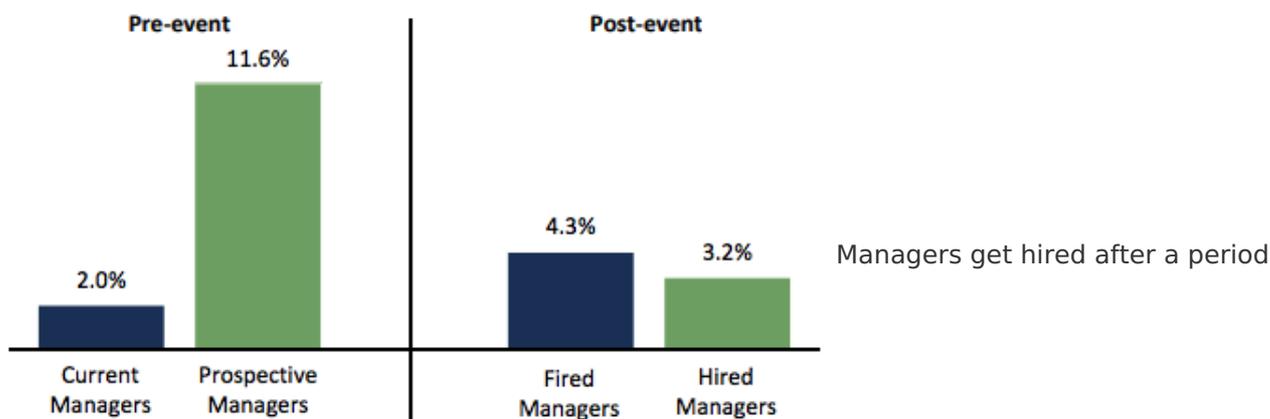
Source: Aaron Reynolds, "The Truth About Top-Performing Mutual Fund Managers," *AAL Journal*, July 2011.

Despite the fact that every manager in the study was ex-post "talented" over the decade,¹⁵ 25% underperformed by 15% or more in a year in the period and 57% by 10% or more. Similar percentages underperformed by 3% and 5% annualized over 3-year periods. And while 5% underperformance for 3 years doesn't sound impossible to hang on through, remember that these aren't Stalin managers. I haven't fully gone through the track record of each of the funds in his study, but a quick survey suggests they had average tracking errors of less than one-third of what a Stalinesque manager would have had. Multiply those underperformance thresholds by three and it suddenly seems a little harder to assume you'd never be tempted to fire one of them along the way. Three years of -15% annualized performance is tough to ignore,

and really tough to explain away to your investment committee.

The second study, by Goyal and Sunil, looks directly at hiring and firing decisions by institutions.¹⁶ This paper showed an event study of manager hiring and firing decisions by 3,400 institutions between 1994 and 2003. Their findings aren't hugely surprising, but they are eye-opening nevertheless.¹⁷ Exhibit 10 shows some of the results.

Exhibit 10: 3-Year Cumulative Excess Returns



Data reflect cumulative excess returns for investment managers from a sample of 412 round-trip firing/hiring decisions between 1996 and 2003.

Source: Goyal, Amit and Wahal, Sunil, "The Selection and Termination of Investment Management Firms by Plan Sponsors," *Journal of Finance*, August 2008.

of extremely good performance. Managers were hired after posting an average 3-year alpha of 3.3%, which dropped to 0.6% in the 3 years after being hired. Sunil and Goyal found that the higher the returns from managers prior to hiring, the more returns fell in the subsequent period.

Managers that were fired, not surprisingly, had done substantially worse than those that were newly hired, underperforming by 1% per year in the 2 years before they were fired. This small figure is slightly misleading, however, as managers were terminated for a wide variety of reasons. Those terminated primarily for performance underperformed by 3.5% per year. Unfortunately they don't show the subsequent performance of the cohort of managers terminated for performance, but for the entire group of terminated managers, the subsequent 3-year alpha was 1.1% per year, twice as much as the newly hired managers.

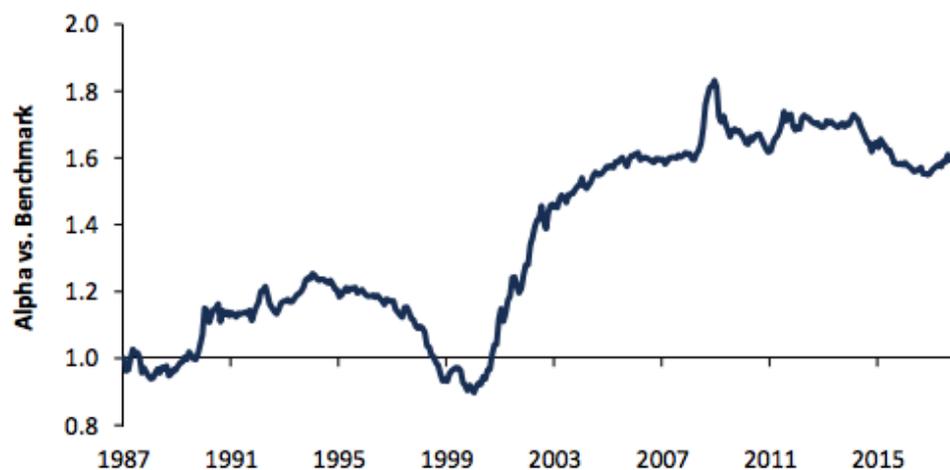
The alpha figures in this study all seem fairly small on both the positive and negative side, but it is important to recognize that this cohort of managers was decidedly not Stalinesque, with an average tracking error in the range of 1%. This seems to be impacted by a large number of fixed income managers in the sample, but as near as I can tell, even the domestic equity managers had tracking errors of less than 2% on average.

Given the low tracking error of these managers on average, the impact of the performance chasing effect isn't huge, at about 0.4-0.5% per year underperformance of the newly hired managers relative to the managers they replaced, but that's still more than half a standard deviation of "bad luck." When dealing with tracking errors an order of magnitude larger as they would be in a Stalinesque portfolio, those numbers would really start to hurt.

Appendix B: GMO Global Equity Case Study

Disentangling the underlying factors behind manager returns is a difficult task. The often highly concentrated portfolios of Stalinesque managers can make it a particular challenge, but just because a manager has a lower tracking error and more diversification doesn't make it easy. GMO's Global Developed Equity Allocation composite is a good case in point. It's got a nice long history, a clear style bias, and yet understanding why the portfolio has performed the way it has has been challenging to many of our clients. Exhibit 11 shows the value added of the GMO Global Developed Equity Allocation composite since its inception in 1987.¹⁸

Exhibit 11: GMO Global Developed Equity Allocation Composite



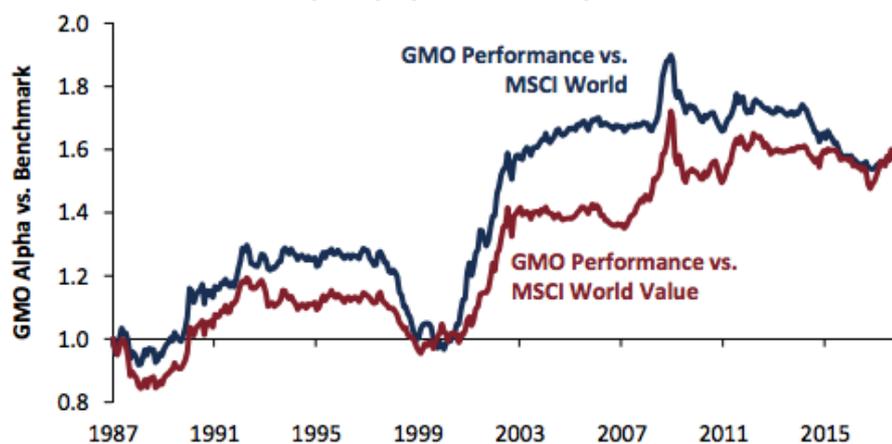
As of 12/31/17

Source: GMO, MSCI

Over the full history, the strategy has had an alpha of 2.3% on a gross of fee basis¹⁹ and a tracking error of just under 5%.²⁰ Clients hired us in the early to mid-1990s on the back of pretty good trailing performance, and more than half fired us in 1999 and 2000 after a number of difficult years against the benchmark in the late 1990s. Even more clients hired us between 2004 and 2007 given our extremely strong performance off of the tech bust in 2000, and a number have in turn fired us in the last couple of years after weaker performance since 2009.

A client who invested a dollar at inception would now have \$18.40, almost exactly double the \$9.57 of the benchmark. Unfortunately, despite the fact that the strategy has outperformed, many of our clients did not outperform because of the timing of their hiring and firing decisions. Jeremy Grantham once wrote “90% of what passes for brilliance or incompetence in investing is the ebb and flow of investment style.”²¹ As this portfolio has had a strong value bias over its history, perhaps our performance looks less cyclical relative to a value benchmark? Unfortunately, there are more styles in heaven and earth than are dreamt of in your Barra analysis.²²

Exhibit 12: GMO Global Developed Equity Allocation Composite Value Added vs. Benchmark



As of 12/31/17

Source: GMO, MSCI

Exhibit 12 shows the performance (net of fees) of our composite against both MSCI World and MSCI World Value.²³ The overall performance record is the same - we have had as much alpha over the value benchmark as we have had over the broad market. Looking at our performance against the value benchmark does help explain some of our more extreme performance moves. We still underperformed the value index in the late 1990s, but the magnitude was meaningfully less striking. And we still did extremely well in the early 2000s, but against value it looks a little bit less miraculous. Comparing us to value does put some context around our performance, but is far from a magic bullet. The years 2016 and 2017 are an excellent example of how style analysis can't be a substitute for actually digging into what we are doing in our portfolio and why. Value had a good year in 2016, with MSCI World Value outperforming MSCI World by 4%. Our composite nevertheless underperformed MSCI World by 1.7% gross of fees. In contrast, 2017 was a tie for the worst year for MSCI World Value since 1999, with value underperforming the broader index by a disheartening 6%. And yet for the year, our strategy beat MSCI World by 2.8%. On the face of it, that seems quite odd. Digging in, it is possible to see what was going on. The underlying regional portfolios that make

up the overall strategy do indeed have a value bias, and their performance reflects this. But we also bias our portfolio toward the broad groups and regions that look cheap to us. In 2016, despite the fact that value worked well as a stock selection parameter, it worked poorly for selecting regions, with the most expensive market, the US, performing best. In 2017, while value worked poorly as a stock selection parameter - particularly in the US and emerging markets - it worked well selecting regions, as non-US markets beat the US and emerging in particular had a very strong year. It certainly helped that our stock portfolios also substantially outperformed simple value in 2017, but our alpha came from bets that MSCI Value doesn't capture.

The point is not that it is impossible to understand why GMO outperforms and underperforms when it does. Clients who know us well and know what is going on in our portfolios shouldn't have been particularly surprised by how we performed in either year. We have a value bias, but we have other biases that are harder to systematically strip out. It's hard to believe we are unique in this. While it is tempting to stick managers in a style box, most managers are more than a simple style, and understanding why they perform the way they do takes work. And remember, the strategy we are talking about here is a "normal" tracking error, decidedly non-Stalin strategy. In a very aggressive, probably highly concentrated strategy, the difference between a great year and a horrible one could easily be a single stock. How to decide whether that stock's impact was luck or skill (or lack of skill) is sure to be a challenge. If such a portfolio is going to be a helpful part of your overall portfolio, finding a way to answer the skill or luck question without resorting to simply looking at performance is crucial.

Ben Inker. Mr. Inker is head of GMO's Asset Allocation team and a member of the GMO Board of Directors. He joined GMO in 1992 following the completion of his B.A. in Economics from Yale University. In his years at GMO, Mr. Inker has served as an analyst for the Quantitative Equity and Asset Allocation teams, as a portfolio manager of several equity and asset allocation portfolios, as co-head of International Quantitative Equities, and as CIO of Quantitative Developed Equities. He is a CFA charterholder.

Performance data quoted represents past performance and is not predictive of future performance. Returns are presented after the deduction of a model advisory fee and a model incentive fee if applicable. Net returns include transaction costs, commissions and withholding taxes on foreign income and capital gains and include the reinvestment of dividends and other income, as applicable. A GIPS compliant presentation of composite performance has preceded this presentation in the past 12 months or accompanies this presentation, and is also available at www.gmo.com. Actual fees are disclosed in Part 2 of GMO's Form ADV and are also available in each strategy's compliant presentation. Fees paid by accounts within the composite may be higher or lower than the model fees used. The information above is supplemental to the GIPS compliant presentation that was made available on GMO's website in August of 2017.

Disclaimer: The views expressed are the views of Ben Inker through the period ending February 2018, and are subject to change at any time based on market and other conditions. This is not an offer or solicitation for the purchase or sale of any security and should not be construed as such. References to specific securities and issuers are for illustrative purposes only and are not intended to be, and should not be interpreted as, recommendations to purchase or sell such securities.

Copyright © 2018 by GMO LLC. All rights reserved.

¹ The "absolute return" Stalin problem gets somewhat complicated for a couple of reasons. First, if the required return is greater than the expected return of any ownable asset (and you can't lever the portfolio) it immediately becomes a "risk-seeking" strategy, where its preference between two portfolios of equal expected return is for the one with greater volatility. This generally is seen as going against the concept of fiduciary responsibility. Even in the case where the return is achievable, the strategy gravitates to very high absolute risk portfolios, assuming risk premia are positive. This leads to high levels of leverage in portfolios that allow for it and 100% equity portfolios in no-leverage portfolios where equities have a higher expected return than other assets.

² The simulations in this quarterly required a number of assumptions to compute. I'm confident that the findings are not particularly sensitive to the specific assumptions I chose, but I will list them along the way. The frontier in Exhibit 1 (and all subsequent simulations that stem from it) assumes a 0.75 net IR at a 1% tracking error, with the IR falling 0.018 per 1% increase in tracking error. Throughout the piece I assume that returns are normally distributed.

³ For this simulation, I'm going to claim that each manager's alpha is completely independent of the others'. This is unlikely to be true in reality, but on the other hand, larger institutions generally have more than 20 managers. Consider 20 to be the figure for the "effective" number of independent managers.

⁴ This assumes that the managers with skill have an IR of 0.68. There is nothing magical about this assumption, and actually the IR changes slowly as you change the assumed IR.

⁵ This could be due to higher transaction costs for the less liquid, more aggressive portfolios. But frankly, I think a bigger reason will be higher fees. If you really expect your managers to swing for the fences, they are generally going to demand upside to make up for their greater odds of getting fired for a run of bad luck. Even though this would likely come in the form of “carry” rather than a flat management fee, it still increases your costs from unskilled managers because some of them will generate alpha by good luck, and you’ll be paying them for it.

⁶ The tracking error is unaffected by the percentage of managers who are skilled. I am assuming that whether skilled or not, they hit their tracking error targets.

⁷ This path was built by a simple Monte Carlo using the Stalin portfolio tracking error and alpha from Exhibit 1. If you try to do this, you’ll find that you almost never wind up all that close to your input tracking error and alpha over a short (50- year) simulation. This path was the first one I found that wound up close to the original inputs. Obviously there are tons of potential paths, but I didn’t have to cherry-pick this one to demonstrate my point.

⁸ This analysis is going to seem to come to the conclusion that the secret to outperforming is to never fire a manager. This is unlikely to be the right solution in the real world. In reality you should update your belief in the talent of a manager over time and when that belief drops below a threshold, it makes sense to fire him or her. You can put together a more complicated scenario in which you have some initial hit rate on managers, you have some other assumed skill at updating your assessments, and managers have some additional likelihood of either becoming talented or untalented if they started off the opposite. It rapidly becomes a pretty complicated simulation and set of assumptions, fun for the more quantitatively oriented, but really hard to explain. I kept it simple, but I believe the basic findings hold for the more complex case as well.

⁹ This makes the average holding period of a manager a little less than seven years, which certainly smells about right to me, although you can feel free to stick in your own assumed holding period if you’d like. It changes the numbers slightly, but not the basic point.

¹⁰ The fact that the managers had bad luck doesn’t mean they were necessarily talented managers in the first place. Nontalented managers can have good and bad luck just as talented managers can.

¹¹ There is actually a little more going on here than meets the eye. In reality, the subsequent returns of your fired managers do not impact your portfolio. But by firing unlucky managers, the aggregate luck of the portfolio will be reduced. Let’s imagine that luck entirely washes out of your portfolio in the absence of hirings and firings. In a given period, you would expect about three of your managers to have performance worse than one standard deviation below target and three to have performance better than one standard deviation above target. Half of that “luck” will be reversed in the subsequent period. If you fire the managers who had the bad luck and replace them with managers who had good luck over the period, you have six managers whose performance should be worse than their “normal” over the subsequent period and just removed three managers who would have had better performance. Without any firings and hirings, you would have expected three better than normal and three worse than normal, which would cancel out.

¹² An extra 15% turnover rate will mean an additional “bad luck” quotient of that magnitude.

¹³ I’m effectively assuming that your hit rate and your behavioral bias are two completely separate factors. As my colleague Melanie Rudoy pointed out to me, that is almost certainly false. However, in one simple way, the two are connected in a way that means an improvement in your hit rate is likely to mean a worsening of your bias. Past returns are, in reality, a valuable piece of data for helping to determine manager skill. Because skilled managers have higher average alphas than unskilled ones, past strong performance really does increase the likelihood that the manager had skill, and bad past performance increases the odds the manager didn’t. Let’s imagine two CIOs. One completely ignores past performance in making manager decisions. The other relies entirely on past performance. The first CIO will have no behavioral bias in hiring and firing decisions, but is ignoring some very valuable data and will have a lower hit rate because of it. The second CIO, by looking at the performance data, will improve his or her hit rate but have an extremely strong behavioral bias when making decisions. While the second CIO really should have a decent hit rate finding talented managers, it is almost certain that he or she would be better off indexing given the strong bias inherent when making hiring and firing decisions.

¹⁴ Aaron Reynolds, “The Truth About Top-Performing Mutual Fund Managers,” AAI Journal, July 2011. His sample is 370 mutual funds with an alpha of at least 1% versus their benchmark and a volatility lower than their benchmark.

¹⁵ I’m defining talented managers here as those who outperform over time. Because this entire group outperformed over the 10-year period, they were all “talented” over that period regardless of whether it was due to some platonic internal “talent” or simply good fortune. In fact, this sample should be even more

positively biased than one based on true “platonic” skill, as it is quite possible for a talented manager to underperform over a 10-year period as we saw in Exhibit 4.

¹⁶ Goyal, Amit and Wahal, Sunil, “The Selection and Termination of Investment Management Firms by Plan Sponsors,” *Journal of Finance*, August 2008.

¹⁷ Actually, the truly striking thing in this paper to me is that the average alpha of managers over the period was positive. Go active management!

¹⁸ I used this composite because it is the longest running equity allocation strategy at GMO. The return patterns, not surprisingly, are similar for our other equity allocation strategies as well.

¹⁹ The portfolio’s alpha net of fees was 1.6%.

²⁰ This is for the asset-weighted composite, before fees relative to the asset-weighted benchmarks of the portfolios in the composite

²¹ Jeremy Grantham, “Everything I Know About The Stock Market in 30 Minutes,” November 7, 2003.

²² Couldn’t resist the misquote. I know that Barra does look at more factors than the value/growth axis, but it is still the case that a more complex packaged analysis would have extreme trouble understanding that our early hatred for Japan and our recent love of emerging aren’t two different time-varying biases in the portfolio but examples of the same underlying strategy of pushing the portfolio toward cheap regions and away from expensive ones. Similarly, our movement into and out of small caps, quality, and other factors can easily look like unnecessary and temporary tracking error from uncompensated risk factors to a portfolio analysis package. In reality they are inherent in our basic strategy, and an important part of our long-run alpha.

²³ Our performance versus MSCI World is slightly different from Exhibit 5, which shows the aggregate performance against the dollar-weighted average benchmark of the accounts included. Different accounts have had somewhat different benchmarks, which don’t quite average to MSCI World. As we don’t have a comparable value benchmark, it seemed fair to switch entirely to MSCI World, which does have a standard and value version.