

Rodrigo Gordillo (00:06):

Welcome to Gestalt University hosted by the team of ReSolve Asset Management where evidence inspires confidence. This podcast will dig deep to uncover investment truths and life hacks you won't find in the mainstream media. Covering topics that appeal to left brain robots, right-brained poets and everyone in between all with the goal of helping you reach excellence. Welcome to the journey.

Speaker 2 (00:28):

Mike Philbrick, Adam Butler, Rodrigo Gordillo and Jason Russell are principles of ReSolve Asset Management. Due to industry regulations they will not discuss any of Resolve's funds on this podcast. All opinions expressed by the principles are solely their own opinion and do not express the opinion of Resolve Asset Management. This podcast is for information purposes only and should not be relied upon for a basis for investment decisions. For more information visit investresolve.com.

Rodrigo Gordillo (00:54):

Hello everyone and welcome to another episode of Gestalt University. My name is Rodrigo Gordillo. Today we're going to bring you I think a very timely episode. We're going to have as a guest, our friend [Corey Hoffstein](#) and Adam Butler, our CIO of ReSolve to talk a little bit more about the concept of ensemble construction for portfolios and strategy construction.

Rodrigo Gordillo (01:16):

And the reason that this is timely is because even though this podcast was recorded a couple of weeks before, we are going to be publishing this about a week after the February market crash. So this is the last week of February where we saw the fastest market crash that we've seen in the history of the US market and it's made it so that it frames this concept of portfolio construction through ensembles in a better light.

Rodrigo Gordillo (01:44):

The reason I say that is because a lot of the simple trend equity strategies that many people follow as an on/on switch like the 200-day moving average or the 10-month moving average and the 12-month moving average. These have all recently and very abruptly triggered a Sell All signal. And what we're going to contrast in this episode is the difference between the robustness of ensemble methods that tend to act kind of like a dimmer switch upon events that start change in the trends in markets versus the light switch approach that a single system may have.

Rodrigo Gordillo (02:24):

And I think a lot of the current market action this past week and the emotions flying this past week are going to put things into perspective when you listen to the full hour, hour and a half long podcast. Before we get into it I just want to share some of the emotional turmoil that I'm hearing from investors and advisors that may invest with ReSolve but also run their own very simple models.

Rodrigo Gordillo (02:52):

It just reminded me back in the day in 2008 where I had a mentor that had built his whole business on this idea of the 10-week and 40-week moving average crossover and had done a very good job of accumulating assets under the assumption that when these things trigger, they would be able to get out. The advisor would be able to pull the trigger in order to get out and I remember vividly.

Rodrigo Gordillo (03:22):

I wasn't working for him back then but I did reach out and we're in Canada so the market didn't really roll over in Canada until August of 2008, but it finally started to roll over. All of the major Canadian stocks were rolling over, the 10 and 40 week and I reached out to this gentleman and I asked him what he's doing, he must be getting clients out of the market, and what he articulated is how difficult it was to pull the trigger because he hadn't seen a trigger in six, seven years.

Rodrigo Gordillo (03:54):

And he was going to wait it out to see maybe if this is a choppy area, the last thing he wanted to do was get clients in and out for the next few weeks only to remain long again. And to his detriment unfortunately he never got out and clients lost whatever the markets lost if not more depending on what he was holding. And that to me made it absolutely clear.

Rodrigo Gordillo (04:16):

Two things, one that you needed to have a very disciplined process but number two how difficult it is to make wholesale changes in strategies that may work over time but when you haven't seen these signals in such a long period of time, having a 100% flip in asset allocation can be a really tough, gut wrenching thing. And as I canvased my clients and my fellow advisors that continue to use these type of mandates they're finding it very, very difficult.

Rodrigo Gordillo (04:47):

Some have not even pulled the trigger yet. So I think the concept here that we're going to discuss on ensembles is the ability to not have to make these wholesale decisions. The ability to ease in like a light switch and ease out like a light switch. Make it a lot more palatable for the construction of trend equity or whatever type of strategy that you may want to implement.

Rodrigo Gordillo (05:05):

And so it's interesting, I think you'll get a lot out of it. We're not just talking about the emotional side on this podcast, we do get into a lot of the theoretical concepts behind it and hopefully it'll get allocators, investors, and practitioners to think a little bit deeper about the problem in these turbulent times. With that said, I hope you enjoy the podcast and we look forward to hearing your thoughts.

Ensembles

Rodrigo Gordillo (05:33):

We're all here, this is a long time coming, we've been collaborating on this for over a year now. The idea of ensembles and this is the first time we're actually going to riff together on it and see if we can add some more color to what it is that ensembles truly are. So I have Corey and Adam here, the two main authors on all the ensemble content to give us a little bit of background as to why we came together and how the Newfound ReSolve Robust Equity Momentum Index came to be, and why it's important to talk about ensembles, and why it's important that this Index exists as a good model for other quant's going forward. So one of you hit it off and tell us a little bit about how the research came about and how you started collaborating.

Adam Butler (06:16):

Corey, come on your going to have to just go on this thing.

Rodrigo Gordillo (06:19):

You've told the story like 20 times already so you've got it honed in baby.

Corey (06:23):

So this is a pretty unique scenario I think in the industry. We definitely see lots of strategies out there where you've got multiple sub-managers who are working on their individual sleeve. I think what is unique here is this is a truly collaborative strategy and a collaborative innovation that was really born out of I think first and foremost a mutual respect between the two firms. I know I first learned of you guys several years ago reading your papers. Adam had read some of my papers and the team there and I think we were approaching this idea of ensembles and thinking about diversification more holistically from different angles.

Corey (07:01):

I know Adam you and your team were very focused on the process based benefits of ensembles and I was really focusing on the timing elements of ensembles and I think we were mutually attracted because we weren't looking at it from that other angle so they were bringing some new thoughts in place. Over time obviously through sharing that research developed a mutual respect. Ultimately we share a lot of similar philosophies and product design and product construction which lead to having a lot of clients in common.

Corey (07:33):

And over the last year really had a lot of clients clamoring for a strategy like this from both of us and so it seemed natural to us that rather than each independently launch a strategy, it seemed natural to work together and create a best of ...type of product where we could both inject some of IP in thinking around this concept of ensembles and then really work together from a business and servicing perspective to help service the advisors and clients who end up using the strategy.

Corey (08:01):

So it was really born out of a long term relationship that we've had. Again that mutual respect and just sort of a unique need we saw in the market for a strategy like this.

Rodrigo Gordillo (08:11):

Yeah I think one of the interesting aspects of it was that we were quite alone in this concept and this belief. There weren't a lot of people out there that truly subscribed to this process. So we almost found refuge in our conversations and it made sense for us to do our best to communicate the benefits of all of this. And it's interesting because in spite of all the papers and communications we've done we continue to get the same type of objections or pushback as to why ensembles aren't necessarily ideal. I want to cover some of those today if it makes sense.

Corey (08:43):

Yeah I mean I've been full on branded the crazy rebalance timing luck guy and I think that one's going with me to my grave.

Rodrigo Gordillo (08:50):

Yeah, well I think we all deserve that moniker. I've been talking about this forever and we keep on coming back to especially the quants and junior quants or individual investors, there's a lot of model testing software out there and services that you can go in, create all types of crazy parameter sets that give you a fantastic back test and some of the most popular ones happen to be trend.

Rodrigo Gordillo (09:14):

So the main objection I get is why would I do anything with you guys? Because I've through your index, downloaded the data. My trend model with my specifications out- performs yours. So you talk a good game but what does it make sense if the data isn't bearing it out? Anybody have some thoughts on that?

Corey (09:33):

Well maybe we can back up first and just sort of make sure we lay some ground work. Set the table, discuss what we actually mean by ensembles here. And maybe I'll just start giving my definition and then Adam you can jump in. So I think probably long-time readers of my research will know I really like to think of diversification holistically across three different axes. What you invest in, how you invest and when do you invest? And for me, ensembles really capture those how you invest and to a lesser degree when you invest decisions.

Corey (10:07):

And so the idea is most investors are used to this idea of diversification across assets. That's what you invest in. But the means by which you make those decisions obviously have a huge impact as well. So to your point Rod, I think momentum and trend are very simple strategies for people to start implementing because it doesn't require a lot of data. Very easy to calculate and I think the gravitational pull is to try to then first, and I think every quant goes through this, try to find that magic set, that holy grail of parametrization, that specification of the model that leads to the best performance.

Corey (10:41):

And the whole idea of an ensemble is well the same way I diversify what I invest in, I'm not going to use just one model, I'm going to diversify how I make that decision. I'm going to use all these different models together in combination and I think that combination is going to bring all the same benefits of diversification that you see when you diversify across assets, and I think when you state it that simply I think it's hard to refute, and yet we still see time and time again in the industry this is not a well-adopted concept. So Adam I'd love to hear how you think and define ensembles.

Adam Butler (11:12):

For us it's about doing our best to avoid the potential for being specifically wrong. We've identified a set of indicators or strategies or signals that seem to have some predictive ability in markets. Trend is one and momentum, it's kissing cousin. There's obviously a myriad of others but depending on how you specify that signal or specify a strategy that uses that signal you can get drastically different results over fairly long, like surprisingly long periods of investment, from what seemed like trivial differences in how you define the model.

Adam Butler (11:55):

So what we just wanted to do, we know that there's something there. We know there's something to trend and momentum or we have strong confidence. So how can we avoid the possibility that we'll just happen to choose a model specification that ends up having really bad luck over the time horizon that's going to matter to our investors? And one of the things I think that gets overlooked with this luck aspect is that everybody seems to perceive that if you have bad luck in your early part of your investment horizon it's inevitable that bad luck will mean revert later on.

Adam Butler (12:36):

And I think there's probably room for some discussion on the difference between the ensembling of cross sectional ensembling and the impact of compounding on good and bad luck. Because the reality is if you have a few months or a few days of really bad luck in the beginning that carries on through the entire duration of your investment experience and

there is no expectation of mean reversion. So to the extent that you can minimize the probability of those outlier bad luck experiences then the likelihood is you're going to get a lot closer to that general basic style premium or basic investment experience that you're going after.

Rodrigo Gordillo (13:23):

Yes. I would even take it a step further. We've been talking about trend momentum. What are those things? What is it that we believe we are extracting? And from the perspective of a repetitive error in terms of trend and momentum, there's many theories but in a nutshell we could say that we're all trying to extract a phenomenon, maybe herding behavior. That humans tend to herd and so somebody some day came up with the 200-day moving average as a pretty good way of extracting that phenomenon. And it does, it works over long periods of time. It works across many asset classes but it doesn't work every time. It works over time and depending on how lucky or unlucky you get in any given year, you may or may not have gotten out of the S&P 500 at the right time at the 200-day. Whereas, the 205-day may have done something completely different.

Rodrigo Gordillo (14:14):

Now, from a theoretical basis the 205-day moving average manager versus the 200-day moving average align perfectly philosophically. But their specification made a massive difference in that particular year that they didn't hit that trigger.

Adam Butler (14:30):

Well the most recent year, the 2018 experience is a really good example.

Rodrigo Gordillo (14:34):

Yeah you have the statistics there Corey, you seem to have a better grasp on where the beeps that a 200-day moving average manager may or may not have gotten in and out of?

Corey (14:43):

Yeah I mean there is a couple of very simple models, like a 200-day moving average that I think it was November 2018 that a lot of those models were about 40 basis points away from triggering going back into the market which would have been a pretty substantial whipsaw. And that is what you see with a lot of these very simple models making these big all in, all out binary decisions. You start looking back through the full history of these models there's all sorts of occasions where they just got really close to the line. Or conversely they just edged over the line and it creates this big whipsaw. So the question ultimately becomes do you really want to be sensitive to that sort of 40 basis point afternoon luck?

Rodrigo Gordillo (15:23):

Right and at the end of the day what we all agree on is there's something there in terms of herding and we want to minimize the chances that we don't capture the signal in a one to three year period. And the way to minimize that chance through ensembles is saying, "Look I actually don't know whether it's a 205-day or the 200-day or the 20-day or the 50-day, or the 300-day." We know generally they work. Any one of those over long periods of time tend to do just as well from a Sharp ratio perspective, but they all blow up also at different times.

Rodrigo Gordillo (15:54):

So the key is in putting an ensemble together where they're all blowing up at different times. Sometimes they coincide but there's enough that are not blowing up that it actually helps reduce the luck risk. The bad luck risk. And I think when we were talking about momentum, I know that's a popular theme but the one that seems to resonate the most is the

value conversation because everybody knows that value is price-to-book. That is the most famous - white paper's written on it. Large multi-billion dollar companies that base their whole business on the price- to- book phenomenon and it has just done poorly.

Rodrigo Gordillo (16:26):

But they don't talk about price- to- book, they talk about the value phenomenon as if that's the thing. But there's other ways to extract value that are just as reasonable. Price-to-sales, even the enterprise value, there's a wide variety of them and they've all done differently over the last couple of decades.

Corey (16:41):

Well not just differently, I think it's dramatic how different, just sticking on value for a moment, I know Adam you just ran some of these numbers a week or two ago and we ran some of these a couple of years ago on our blog where you can go to Follow My French and download their price-to-book talk decile and I think you can download price-to-earnings and you can download price-to-sales or price-to-cashflow whichever one it is, price-to-cashflow. And you plot the differential in rolling one year returns between the performance and it can be thousands of basis points. For two long-term value strategies you put them in as factors and they're pretty repetitive to one another but from a short-term experience perspective.

Rodrigo Gordillo (17:20):

Yeah the underlying thesis is the same.

Corey (17:22):

Absolutely.

Rodrigo Gordillo (17:22):

The implementation is different.

Corey (17:24):

Every value manager has struggled in the last year and a half. Just the degree to which they've struggled has to do with their specification for the most part but that style is gravity ultimately. If the style is gravity they're just trying to escape that gravity or outperform that gravity to a certain degree based on their specific choices.

Rodrigo Gordillo (17:42):

What was interesting as this value conversation evolved is how people weren't going directly to "Hey maybe we don't know we should use the ensembles." But they were saying, "Oh look at enterprise value to EBITDA." I can't remember which metric that had killed it over the last 10 years and then they were like, "Well this is the new value. It's just we should be doing this." And not recognizing that, that's legitimately data mining.

Adam Butler (18:04):

It leads to a really good, we may have wanted to discuss this a little further on but it does beg the question, you believe there's a phenomenon, you're going to have to define a model to harness this phenomenon as an investment strategy. How do you choose? How do you choose the specification? How do you specify the model in advance? Follow My French shows book to price because they had good long-term data on book value, and because book value is relatively stable and difficult to fudge.

Adam Butler (18:42):

But it was mostly just data availability. It's sort of a lamp post problem, why are you looking for the keys over there? Did you lose them over there? No I didn't lose them over there but this is where the light is. So they had light on book value, they didn't have any light on earnings or some of the other value metrics so they lit on book to price. But if you've got other specifications available how do you choose which one? And I think the guys at Alpha Architect, they did a really great study.

Adam Butler (19:09):

They did this horse race back in 2011, published a paper on a horse race of different value strategies. I'm just going to go by memory but I think it was sorts on, enterprise-Value-to-EBITDA, price-to-cashflow or price-to-free cash flow, price-to-book, price-to-earnings. Anyways, there were four or five and they had a simulation of these sorts going back to 1962. And the enterprise-value-to-EBITDA sorts did substantially better than several of the other sorts. That was the dominant strategy both for equal weighted and for valuated portfolios.

Adam Butler (19:45):

And then as a follow-up five years later, they examined the out of sample performance of the same strategy specifications and what they discovered is that the enterprise-value-EBITDA specification, which had done the best in the historical sample had profoundly underperformed in the five years out of sample by, I'm just going my memory, but on the order of 9% annualized over the five year period on the valuated portfolio. And it could be the enterprise-value-EBITDA is legitimately the better specification. But for an investor with a three to five year emotional investment horizon and a 20 to 30 year financial investment horizon five years is an eternity, to underperform by 9% a year is a catastrophic outcome that could have been avoided by simply taking a more humble approach to how you specify your desire to get exposure to value.

Rodrigo Gordillo (20:45):

Well. On that note, it was in fact the research done, I think this was also on that website that showed price-to-book price-to-equity, price-to-cashflow, enterprise-value-to-EBITDA and it showed in the US markets price-to-book being the worst performing. So validating the fact that price-to-book is dead. In the European market same thing, but then price-to-book is the best performing one in Japan.

Corey (21:08):

Japan, it's always Japan.

Rodrigo Gordillo (21:09):

It's always Japan. But moreover they have another equity line there. That's the multi metric one where they just equal weight all of them and guess what? It's the best performing line. It literally ensembles more often than not - ends up being the perfect foresight portfolio.

Corey (21:24):

I want to go back to something you said earlier, Rodrigo, almost back to the original question, which is this issue of, okay, how do you compete against, well, my back test is better than your ensemble? And I know you guys don't really love this line, but I almost like to just say, look, ensembles are average by design. We're going to include models that did worse than a back test. We're going to include models that did better in a back test. The idea is creating that consistency. And

what becomes really hard is trying to demonstrate to people how the ensemble would've done, versus something else. When we only have one reality, there's only one back test.

Corey (21:59):

There's only that evidence and unless someone has a really good grasp of probability and statistics, it can be really hard for them to understand that the precision that they're seeing, that compound annualized growth rate that goes out multiple decimal places, is a false precision. It's really shrouded in this probability distribution that never really gets reported. So we have this annualized return of the strategy and yes, that is how the strategy did on the back test, but our expectation going forward is really a distribution.

Corey (22:29):

And so what you tend to see is actually when you start comparing these distributions to each other, they're not statistically meaningfully different. And that I think can be really hard, especially on a Sharpe ratio perspective. Adam, you've done a ton of work around this, but I think it really amazes people. I see this all the time in papers. Oh, we increase the Sharpe ratio by 0.05 by 0.1 and it's over a 20 year period, so it's got to be statistically significant and you go, well, no, actually the spread there is actually like a 0.2. The standard error around Sharpe ratios is massive and I think that really takes people aback.

Corey (23:02):

And so we've seen two ways to try and combat this. The first is, Hey, we're just going to show you how these are not statistically different. Let me show you the standard errors and that works for some people. The other way, I've done this in the past, and Adam, you just wrote a phenomenal paper on this, is actually just going to history and just adding a little bit of noise. I think you called it a jitter test Adam. And you go back and you take the S&P and you're going to run the 200 day moving average, but every day we're going to either add a little white noise or Adam, I think you shuffled returns in a very local probability type of way. Take tomorrow's return and pretend it was yesterday and shuffle these around a little bit and in theory it should not make a big difference.

Corey (23:40):

But what you find is that when you use a hyper simplistic model, just a 200-day moving average and then you compare that 200-day moving average result over all these different little slightly altered histories, you end up with this massive dispersion and results that I think is really more indicative of the experience you would have going forward versus when you use the ensemble, it ends up being much more consistent across all those little sort of fake histories.

Adam Butler (24:06):

It's pretty remarkable because you can create these slightly noisy histories if you were to see this jittery sample or the S&P with a little bit of noise added to it. If you were to put the S&P with a little bit of noise added against the true S&P visually, you would not be able to tell them apart. They look the same. The profile, the character looks exactly the same. There's just little bit of noise that occurs locally.

Adam Butler (24:35):

So instead of going up 1% today, it went up 0.5% today, but you go up a little bit more tomorrow than how it would actually happen. Over time, that all it cancels out, but locally it makes a big difference and it's astonishing to see that dispersion of results to the exact same strategy specifications. So you apply a 200-day moving average to a market that's had a little bit of noise added to it and the dispersion of outcomes is just as large as you observe from completely

changing the strategy specification. Just a little bit of noise turns a 200-day moving average, gives it the same performance as a 20-day moving average or 300-day moving average. It's crazy. It's a really good way to visualize.

Rodrigo Gordillo (25:22):

And also from a deterministic mind view, which is, okay, cool statistics, Monte-Carlo whatever, jitter test. I want to see how it did in the market. In the Craftsman's Perspective white paper that you wrote, you did a simple experiment to show just that we basically grabbed the data, I think from 1970 to 2012 right?

Adam Butler (25:46):

I think it was 1974 to 2011, that was the original and to NACI global equity momentum paper. He tested from '74 to 2011.

Rodrigo Gordillo (25:56):

And the specification was 12 months. So either you are above zero or below zero and then what you did is you grabbed data going back to 1950 and then walked forward data from whatever, 10 year period after.

Adam Butler (26:07):

Well, yeah, 2012 to 2018. Yup.

Rodrigo Gordillo (26:11):

The original specification for that period was fantastic. It was a top desk style performance?

Adam Butler (26:16):

Yeah, was about the 90th percentile.

Rodrigo Gordillo (26:18):

Of all the ensembles that we'd looked at. And so it's an issue when you're like, okay, well you still haven't proven anything. Your ensembles are still worse than that 12 month specification. But then what happened? Why don't you tell us what happened, what you did, and then what the outcome was.

Adam Butler (26:30):

Well, we just applied the exact same methodology, but to out of sample data. So we had out of sample data from 1950 to 1973 and from 2012 to 2018. And so when we applied all the different model specifications to the out of sample period, what we found was that the 12 month specification was pretty well smack dab in the middle. It had converged to the average, which is exactly what you would expect it to do. Just like every other reasonable specification over time will converge to the average of all specifications. And if that's true, then why don't you just target the average all along and you do that with an ensemble and thereby also minimize the probability that you just happen to specify with a profoundly unlucky specification in the short term, which completely derails your financial outcome.

Corey (27:17):

I think another really important point here, and you alluded to this a little bit earlier Adam. Because a lot of these back tests, they take place over 70, 80 years. And there's your evidence. The reality is investors don't invest over 70 or 80 years. They invest over 20, 30 maybe they invest over 50 or their full life cycle, but this is all. Also another really

important consideration here is that at the end of the day, investors need to eat their dollars. So we always talk about annualized return, but they actually need to withdraw those dollars.

Corey (27:46):

And so from a dollar weighted perspective, investors tend to be contributing more and more and more than they get peak dollars. When they retire, then they get less and less and less. And so when you think about a dollar weighted time horizon, they're investing. It is like a super concentrated period that if they happen to be right around retirement and they get a massive either whipsaw or some specification driven event that causes them to underperform or realize losses, that has a huge impact on their actual realized lifestyle.

Corey (28:20):

And I think it's a little nuanced and it goes massively underappreciated. But in terms of actually taking this back a step and getting out of our ivory tower of asset management and being quantum and putting our feet on the ground and talking about what does this really mean for investors, what I would hope ensembles mean is a more consistent outcome, particularly in the short-term where that short-term is hyper-relevant, especially when they're at those peak contribution and peak withdrawal years, right around retirement.

Rodrigo Gordillo (28:51):

That's good for retirement and withdrawals. But it's also the truth about our stick-to-itiveness as human beings, we cannot handle the pain of doing really poorly for more than three years. Institutions can't handle the pain of doing poorly for more than three years. And so what this ensemble concept allows is to reduce that risk. Not completely, because again, this is what I call the ensembles, is the anti data mining approach to evidence-based investing, but there is one thing we're data mining, which is the phenomenon. If herding behavior just simply doesn't exist, we've got it all wrong. Forget it. Like no amount of ensembles would make a difference.

Rodrigo Gordillo (29:28):

There is the chance that maybe herding didn't work for the three years and there's nothing we can do about that. But if we do think that the phenomenon is pervasive and it's going to continue, then the ensembles allows us to minimize the chances that we're going to be specifically wrong, ie: not getting out of the S&P 500 because of a 40 basis point difference and be able to stick to it. The stick-to-it is key here. And also in our research we find that from a risk perspective in every case, ensembles tends to have the lowest drawdowns or very close to the lowest drawdown characteristic.

Rodrigo Gordillo (30:02):

A lower volatility, higher Sharpe ratio, not necessarily the absolute return that everybody cares about, but on the risk side and the protection side, it tends to do better across whether we look at seasonality or skewness or mean reversion or momentum or trend or value. It's a reality across the board. It transfers.

Adam Butler (30:21):

I wanted to just close the circle that Corey was going on there with the retirement because I think, so, Michael Kitces ran a really neat analysis about 10 years ago that found that the probability of running out of funds before you die is 85% explained by the performance of the portfolio in the first five years after you retire. And 95% explained by the performance of your portfolio in the first 15 years after you retire. So this gets into the duration of cash flows and all kinds of stuff, which is also very interesting.

Adam Butler (30:59):

But just to close the circle mathematically, that time around when you retire has a very outsized impact on the probability that you're going to have a successful, you're going to meet your financial objectives further down the road. So that five year or 10 year horizon is the true horizon. The 30-year horizon, I think even financially is a myth, but certainly emotionally is a completeness.

Corey (31:25):

Well, to put some numbers to that, let's pretend for a second. You retired in 2018 and you were investing in this type of strategy. Depending on the specification you had, you either were down just as much as the market. You rode it all the way down in Q4. Some of them got out right at the end of December 2018, missed all of Q1 2019 and so you got all the loss of the S&P and some of these trend following strategies, and these are public funds that you can look up that implement this type of idea. Were actually negative returns in 2019. They got whipsawed so many times because they're one specification.

Corey (32:02):

Then there were other funds that did protect you in 2019 a little bit tough year for trend-following anyway, but 2019 we're up 20, 25%. You talk about specification versus style Rod, they're all doing the same style but they were so driven by their specifications. So if you're an investor and it's year one of retirement, you rode the S&P and you chose the wrong specification, you got all the downside and then maybe you chose a specification that in 2019 despite the market being up over 30%, in the US you were actually down even though you were just trend-following equities.

Corey (32:35):

And that's a reality that happened. And so the question is, well do you try to protect against that? Oh the answer for us is ensembles, but I want to go back. I want to rewind again to something you said Rod, because you were talking about this institutional behavior. And I think this is really ingrained in the industry. It might just be human nature in general. And one of the ways I've said it in the past is we are an industry of cake bakers that are obsessed with ingredients and don't care at all about the recipe.

Corey (33:05):

Everyone wants to know what's your secret sauce? What's your secret signal? And when you say to people, "I actually don't think I have a secret signal, I think I'm going to use all these great ingredients. And I don't think I have, it's not like I've got a secret source for my salt or my sugar, but I think I've got a really good recipe that creates a really consistent outcome." There's just not a lot of fanfare for that concept. It almost sounds like you've given up in the pursuit of alpha, but in reality, I think that there's a tremendous amount of craftsmanship edge in creating consistent performance, and even harvesting potential alpha in that recipe part of the equation that gets totally ignored.

Rodrigo Gordillo (33:47):

Well, you can't cook it at home. That's the big objection. You can't cook it at home. And it's a massive cookbook that you've just designed with a lot of condiments and table of contents that I can't follow. I'd rather have my own cookbook for dummies that I can follow. I can get it done in 20 minutes and I'm probably going to get roughly 80% of the same result if I do this for the rest of my life. And I think that's another major objection that we get is, it's too complex. Why are you taking all of this to just capture what an extra 20% of the signal? I'm fine with capturing the 80% and keeping it simple for myself.

Rodrigo Gordillo (34:21):

And this is what we fight against even from other asset managers that apply simple recipes because from a marketing perspective it's simple to explain. I actually don't think they're doing it from a negative approach. I don't think they know about ensembles. I truly think that they're doing a bunch of tests and they're showing, "Hey, you see that top line? That's what you're mining by the way." I can explain it in two lines and this is the kind of the junior quantum approach all of us went through that when we first started that overly specified data mine piece. And then you run with it and it does okay because it is kind of capturing that trend factor over time.

Rodrigo Gordillo (34:56):

But again, you can get it like you just described, specifically wrong in a very short period of time. And so the problem is of course, complexity, complexity of understandability, complexity of execution. And so why don't we talk a little bit about that? I mean it is a valid argument that a do-it-yourselfer probably can't implement the ensemble approach to the degree that we have, for what do we have, over 30,000 different specifications in our ensemble.

Corey (35:23):

Yeah, and we can get into that. I like to reframe this discussion because I think when you start talking about the process, people go, Oh, that's complicated or complex and simple is robust and I know we can talk about that. But I like to just start by reframing and saying, look, that simple 200-day moving average model might capture 80, 90% of the style, so why bother doing anything more? But isn't it also true that if you buy an individual stock it captures 80 or 90% of the market return over the long run, but are you ever going to put 100% of your money in a single stock portfolio? Probably not.

Corey (35:58):

If you get the call right, it's a great way to get rich and when you get the call wrong, it's a great way to go broke, and I think it's the exact same-

Rodrigo Gordillo (36:04):

It amounts to flip heads 50 times.

Corey (36:06):

Exactly. It's the exact same conversation around when you choose a specification for your investment style. Yes. If you get it right, if the 200-day moving average is the right model for the next decade, you are going to look like a hero. But if it happens to be, I think it was the 1950s were really bad for intermediate term trend models and short-term trend models did really well. You're going to look like a complete zero and not just look like a zero, you're going to be destructive to your wealth. And so from that perspective, yes it might over the long run, that 70, 80 year period capture the majority of the style. But again just like you want to diversify what you hold, you want to diversify how you're making those investment decisions.

Making Decisions

Adam Butler (36:47):

The other question though is how do you even decide in the first place? So you've got all these potential specifications. They all, the 200-day or the 180-day or the 60 day, they all have approximately the same long-term Sharpe ratio, certainly well within standard error bounds. How do you even decide? What's so interesting is the specification that you decide

will be different depending on when you run the tests. So one of the things that we did in our most recent piece is we said, well, if we're going to have to make a decision about what trend model to use, well let's go back in time and assume that at each point in time we had to make that decision.

Adam Butler (37:34):

So we go back to 2008 to 2005 to 1998 in fact, every single month we're going to assume that we're approaching this brand new or engineering a strategy, a trend strategy, and we've got to decide which specification we're going to deploy going forward. And so it's this walk forward test and we're just going to use the historical performance of the strategy up until that point in time, to decide which strategy to choose. And we examined using the Sharpe ratio using the annualized returns, the Ulcer ratio. We used all kinds of different approaches to select the best model.

Adam Butler (38:15):

And what we found was that any attempt to select the best model using whatever model had performed the best up until that point had worse results than just using the average of all the models at each point of selection. The point being that as a researcher, you do this back testing. You think, okay, this model has performed the best over my analytical horizon, and you think that that means that if it's going to go perform the best out of sample, when the reality is there's no evidence that's the case. There's no evidence that your objective criteria have any signal in helping you to decide which specification is going to deliver the best performance going forward.

Rodrigo Gordillo (38:57):

So you just, I don't know, pick one randomly. It's one that I only have to rebalance once or twice a year. You're right. I don't know. I'll pick the 176.5-day moving average. I'll trade at 12 and just call it quits.

Adam Butler (39:09):

Sure. So you could do that, but just taking that one extra step and examining the performance of the ensemble of all of them relative to a random selection of strategies, you will see that the ensemble delivers the same expected performance, but with a vastly reduced potential for bad luck. And so you can get the exact same performance but vastly reduce the potential that you're going to have a really adverse experience at your particular forward timeline. Why would you choose anything else?

Corey (39:45):

Yeah, I think the really amazing thing and Adam, you show these graphs all the time in your papers, you always take the performance of all these different underlying models and you sort them and you sort them by drawdown and you sort them by return and you sort them by volatility. And what you find so consistently is that it's not so much that the ensemble improves returns, and it certainly does improve returns versus those specifications that have a substantial blow up, that over that in sample period had a very bad experience. But what's amazing is that the ensemble consistently ends up top quintile on drawdowns.

Adam Butler (40:22):

Drawdowns worst five-year returns, worst calendar year returns, all of the adverse potentiality, the ensemble just completely dominates and it preserves the same expected return. So why I always say, I honestly don't understand when people say to me, what's the benefit of ensembles? It's like, well, here's a list of 50 different reasons why ensembles are better. Give me one reason why your single specification is better.

Rodrigo Gordillo (40:50):

Well, no, they're not implementation to ease and we'll get to transaction costs as well. I think the biggest thing is, well, I can't do-

Adam Butler (40:56):

But that implementation ease is another red herring. I don't understand why everyone thinks that they could just run their own strategies. Do you have a bunch of individuals running around performing their own appendectomies? Are they defending themselves in court? Why does everybody feel that they're an expert in investments when they're not an expert in any of the domains? It's just ridiculous.

Rodrigo Gordillo (41:15):

Okay. Well, a lot of people do find finance to be a good hobby, hopefully for a very small portion of their wealth.

Adam Butler (41:24):

Lots of people play chess, but they don't pretend that they're a world grand wizard chess champion.

Rodrigo Gordillo (41:30):

No. But I would say that if you hire us we'll be the grand wizard chess champion for you for a fee. And they're thinking, well, I just want to play chess once or twice a year because I like it, but I don't want to play it every day like you guys and I don't want to pay nothing for that. So the first thing is, I think we've addressed it. Why not do it yourself? Because if you're doing it for your own wealth, really for 100% of your wealth, the big issue is not about necessarily the returns, but also the fact that demonstrably we reduce the risk of you getting it wrong in your lifetime when you're in retirement, that should be enough.

Rodrigo Gordillo (42:06):

The fact that the drawdowns are lowered, the fact that volatility is the risks the stick-to-itiveness, and it's all of that is important for you to be able to survive during even your savings and your retirement years. So that alone should be worth the extra effort to try to create more ensembles of you doing it on your own. But of course there's difficulty in implementation if you're a do-it-yourselfer or trying to stay on top of those traits. And so that's why you probably should get a professional to help you out with all of that.

Corey (42:32):

I think it's always worth considering that a lot of this wasn't really possible 40, 50 years ago. Rod, you mentioned we use 30,000 different models. For quants that can program, that is so simple to implement a number of different trend models, a number of different momentum models, different lookback specifications, different re-sampling specifications and the combinations explode. And is the 30,000th model adding that much diversification benefit? Probably not to very marginal, but it also doesn't hurt to run it. It's just computational time and it's a couple milliseconds. So can you do it in a spreadsheet?

Rodrigo Gordillo (43:09):

If you can wrap the signal around completely rather than be parsimonious about it and only capture a portion of those signals, why not? It costs us nothing. It costs us nothing to do so.

Corey (43:20):

Could we have done this 20 years ago with processors the way they were? Maybe it would have taken a while. It depends how long you want to run things. Could we have done it 50 years ago? Absolutely not. So I think it's worth considering. Adam, you mentioned, price-to-book Follow My French. The way they built those portfolios, if someone was starting completely fresh today and wanted to build factor models, they probably would not do it the way Follow My French did it. They made a bunch of choices that made sense at the time with the data they had, the computational power they had.

Corey (43:54):

All that sort of stuff is important in considering and yet it becomes the standard and it remained the standard and it becomes a benchmark and you sit there and go, Oh, this is a little crazy when you think about it, but this is just no questions at anymore. And I think we're getting to a point of evolution in the industry where people start going, no, hold on. this actually doesn't make a whole lot of sense for a whole number of reasons.

Adam Butler (44:18):

I would love to think so. I'm not seeing much evidence of that. But maybe you're seeing some things that I'm not seeing.

Rodrigo Gordillo (44:26):

Let's assume that we got somebody across the line on, okay, ensembles good. I'll take the added complexity of trading for this. The next thing that we've got is trading costs. How much more expensive is it going to be than me just doing a couple of trades a year, and I know you guys put a lot of thought into that.

Corey (44:45):

Trading costs are one that I hear all the time because with an ensemble almost by definition you will be trading more, but we have, I personally have five, six years of institutional trading experience of doing this type of trading and what I can say is there's two points. One, yes you do tend to trade more frequently, but you tend to trade in smaller size. So to our point that the conversation we have a lot is this idea of a light switch versus a dimmer switch. When you use a simple model, you tend to be all in- all out versus when you're using ensembles, some models stay on some turn off. It tends to be sort of a gradual change.

Corey (45:23):

And so when you go to market to trade, maybe the light switch stayed on for three months in a row while the dimmer switch was sort of making some small tweaks. And so yeah, you were trading more frequently, but when that light switch turns off and all of a sudden you've got to go to market with hundreds of millions if not billions of dollars to execute that trade. Now you're creating an impact in the market. Now that may not be true for an individual, but the individual isn't necessarily doing the institutional caliber trading that we're getting.

Corey (45:52):

So from an actual implementation perspective, trading more frequently, you can have a smaller market impact because you're trading in less size. Now the flip side of that coin is you are often therefore paying more commission potentially. But I can say from all of our best execution review, when you are working with institutional traders and you're placing these trades, commission tends to be a very, very small piece of the equation. And so what I go back to is these are known costs, but in the ultimate Sharpe ratio calculus we need to do, we do need to subtract those known costs from our expected return.

Corey (46:28):

But how much does it necessarily improve our expectation of our outcomes? What I want to say, even if I knew that the trading costs were going to be more, I don't think they are, but even if I knew the trading costs are going to be more with the ensemble approach, are those costs worth it to reduce the huge lurking latent specification risks of just choosing one model. I don't know how you can defend that. Same with, and we talked about this, I'm just going to go on a tangent here. Turn of month effects same sort of thing. I think there's ample evidence there's seasonality out there. Turn of month does seem to at least historically have provided a benefit to all sorts of different models.

Corey (47:04):

But it's such a slight edge that the choice to rebalance just at the end of month instead of sort of continuously throughout the month ends up being this very tiny expected alpha for this huge specification risk that can come out and wipe out decades of alpha you created from that turn of month effect in a single month. So to me it's just one of these trade-offs that does not make any sense. Yeah.

Adam Butler (47:28):

Yeah. The exact same point applies to the trading costs. This vanishingly small incremental trading cost protects you against the potential for a catastrophic tail event from whatever individual specification you might choose. That would inflict orders of magnitude more wealth damage on your portfolio than decades of compounded trading costs.

Rodrigo Gordillo (47:56):

It might not be true in a world of 799 brokerage trades in a \$20,000 account, but that's not what we're talking about here.

Adam Butler (48:03):

Well, no, but we're talking about for individual investors that are choosing an individual specification, the potential size of wealth impairment from bad luck from that single specification, they might be able to trade cheaply. Although I'd be surprised if many individual investors were able to trade their lower trade strategies for cheaper than we trade our slightly higher turnover strategies because we trade at institutional rates. There's that factor too, but I mean the big overwhelming point here is that the risk of a major adverse outcome swamps whatever the incremental potential trading cost differences are.

Rodrigo Gordillo (48:44):

Excellent. Those are very good points. I think we covered the major objections, at least that I'm getting. Is there anything that you've faced Corey, beyond those that we want to chat about?

Corey (48:53):

No, those tend to be the big ones. This idea of simplicity versus complexity is the one that is underneath all of this. We live in a world where it's very obvious that hyper complex strategies tend to fail for all these edge cases and they're hard to reason about. People tend to think of simplicity as being more robust. The gap we need to bridge is getting people to understand that an ensemble is not complex, and ensemble is just taking the simple idea and diversifying it. That to me is the real core concept.

Rodrigo Gordillo (49:28):

It's important to parse out the language here because the original, "Hey, you don't want complexity, you want simplicity" comes from the fact that a 200-day moving average is probably more robust because we've seen a lot of data points

where that has triggered across many global markets. Don't get into the other complex strategy which requires you to cross it to one a day when volume has gone up by 20% and the RSI and this and that. That gives you a fantastic back test that may have only triggered three times in 100 years, and so there's not enough statistical evidence for that complexity to play out in any sort of alpha or whether it's just you're capturing pure noise. That is data mining, that is complexity.

Rodrigo Gordillo (50:11):

And the other 200-day moving average, we have a lot of data points. It's enough for us to have a good idea that it's probably going to work decently well over time. That is not what we're doing here. When we talk about 30,000 different specifications, ensemble shouldn't be about creating complex ensembles. It should be creating very simple versions, but aggregating them 200-day, 150-day, two months versus 18 month moving average, the six months and nine months moving average. Simple. Lots of data points independently. They're robust over time. They're statistically insignificantly different from each other. The ensemble comes insane. I don't know which one it is.

Rodrigo Gordillo (50:50):

Do I have the hubris to say, I actually think it's a nine month look back? Do I have any confidence in that? And so really comes down to approaching this from a place of humility and creating robustness, not complexity, robustness. It's not simple versus complex. It's simple versus robust. That's a big takeaway from all of this stuff. Do we have a clue as to whether the ensemble actually has a bigger edge over time than the individual specifications?

Corey (51:18):

So I think there's something really interesting to the whole, if they have the same expected return, the whole rebalancing premium. If you keep taking the capital and reallocating equally, if you truly believe they all have the same expected return and they're just going to have random noise that you can actually harvest.

Rodrigo Gordillo (51:34):

Some rebalancing.

Corey (51:35):

I haven't worked out the math there.

Rodrigo Gordillo (51:37):

Yeah, neither have I. My intuition tells me that the ensembles probably have a slight edge.

Corey (51:44):

Well, I think the argument against it would be if you just let the ensembles drift, you end up overly allocated to a specification and if you think that the optimal allocation is the equal weight, then you have to keep rebalancing to the equal weight. And so that would make sense as to like there's going to be a rebalanced premium there.

Adam Butler (52:02):

Well, the diversification on its own reduces expected portfolio variance and therefore increases expected geometric returns. The magnitude of that is probably marginal, but in an industry where people move tens of billions of dollars from one ETF to the next because the one ETF is one basis point cheaper, these incremental edges add up.

Corey (52:28):

Yeah, I think that the interesting thing is it doesn't really show up in traditional volatility measures. I don't see it show up as much. When you do an ensemble and it's the same concept of you take five value managers and you mix them together, the volatility of that portfolio doesn't drop a huge amount.

Adam Butler (52:44):

I agree on value, but I mean I pulled up my paper just because I wanted to make sure that we didn't miss anything critical. And one of the things I noticed is that the Sharpe ratio of the ensemble is at the 93rd percentile.

Corey (52:59):

Yup. So what you tend to find is like in more homogenous categories, it doesn't make a huge difference in heterogeneous categories, like especially the hedge fund space, you definitely see that drop in vol. But if you reframe it and say, okay, I'm not going to look at their vol, what I'm going to do is I'm going to look at their excess returns and the vol of the excess return of each strategy say versus the average. Then taking that ensemble, all of a sudden you see a huge collapse in that excess of all which you can think of as idiosyncratic vol relative to the style.

Corey (53:29):

And so in that sense, you can start to say, okay, there's some rebalancing. You're decreasing that geometric towards the arithmetic by decreasing the idiosyncratic, but it's super weirdly theoretical.

Adam Butler (53:41):

Yeah. Well, the total variance as a function of, as you've mentioned, the what, the when, the what is it, how all of those dimensions add idiosyncratic variance. You add them all up that's the total variance. And so if you can reduce the variance by diversifying across each of those different dimensions, you are incrementally causing the geometric returns to converge towards the arithmetic. Your expected value is higher and the dispersion of potential terminal wealth is lower, so it's a win-win.

Corey (54:21):

One thing we haven't mentioned - the whole uncompensated source of risk concept.

Uncompensated Sources of Risk

Adam Butler (54:25):

Yeah, exactly. Why would anybody assume this extra risk if there's no expectation for getting compensated and so to me that's another kind of no brainer assertion on ensemble.

Rodrigo Gordillo (54:37):

Oh my God, of course that is the argument. It's used broadly across portfolio construction experts.

Corey (54:44):

Scrap the podcast, just record that one sentence.

Rodrigo Gordillo (54:47):

That's right. Well I mean it's used broadly on portfolio construction concepts, but this is just another way of constructing portfolios. It just happens to be that we're constructing an ensemble of signal portfolios. We are minimizing the chances that we are taking uncompensated risk.

Corey (55:05):

You go back to my framework of what, how and when. We already know no one's just buying one stock because there's so much idiosyncratic risk. What, how and when no one's buying one stock because of the idiosyncratic risk argument. You want to diversify its whole basis of modern portfolio theory. Why does that not apply in your decision how you're investing? Yeah. Obviously you want to find a style that you think is going to be value additive, whether it's adding alpha or reducing risk, but that's specification. If you don't think the 200-day is necessarily inherently better than 199-day, why are you bearing that risk? Why are you bearing that idiosyncratic risk of that specification? Same thing as when you're rebalancing.

Rodrigo Gordillo (55:46):

You know what answer I always get? Don't you just get the benchmark then? This concept has been used for the S&P. I want my 15 stock active manager because I want active risk and I don't want the benchmark. I don't want the S&P but it turns out that when it comes to factors, you are actually trying to capture that benchmark. You do run the benchmark. The benchmark is a period of this abstract concept, this hurting behavior signal. And precisely what we want to do is be the benchmark. Otherwise, you are taking idiosyncratic risk that you're not compensated for and you're just capturing a portion of that benchmark. We actually want to be the benchmark.

Corey (56:22):

Another way to flip this, and I think Jake, our joint friend Jake always does a good job of this on Twitter. You just flip the conversation and you say, well, what would it take for you to go from generically capturing a style to say, I want to pick that one specific way of doing it. Because we're trying to get people off of one specific way to capture a style and go broad. Let's assume broad, same way what would it take for you to go from buying equities as a whole to having the certainty to buy that one specific stock? Well, you would need a whole lot of confidence.

Rodrigo Gordillo (56:55):

Yeah. You would need for the 166-day moving average that I talked about to be working as the best specification for every market you test it on. If for some reason you test it on like 100 global markets and it just so happens that the best look back is that then there's a lot of evidence. What people tend to do is they tend to have, any of you saw this from Blackrock, it was pitched to me on CTA when they're like, "Here's how we're doing things differently." The lookback for equities is going to be 12 months. The lookback for commodities is going to be five months. Because we've found that they work better in those specifications.

Rodrigo Gordillo (57:30):

They literally data mine the best parameter set and then created this awesome back test. And my eyebrows just went all the way to the back of my head while these guys were trying to pitch this to me.

Corey (57:42):

So they moved on you like a.....

Rodrigo Gordillo (57:44):

Good thing the cameras aren't on baby. So the interesting experiment would be, let's find, we kind of talked about it. We'd done all these types of experiments, but grab all markets you can from as far back as you can, get them to the 1970s find their best specifications, assume that, that's the best parameter set for each market, and then walk it forward another 30, 50 years to show how it's all nonsense. You just can't, you can't pick,

Adam Butler (58:11):

Well you can't, but you know what I think even more fundamentally shocking is that let's assume that there is evidence that the 166-day moving average crossover strategy is the best across every market you test and if the difference is economically significant, this is over a 50 year period. If you've got a five year horizon, you still don't want to have all your eggs in that single basket.

Rodrigo Gordillo (58:42):

You're not even talking about the risk component.

Adam Butler (58:44):

Exactly. You still want to diversify because you have a finite time horizon and because of the duration of your cash flows.

Rodrigo Gordillo (58:53):

Well, hopefully we've given a lot of content and information for further discussion. I look forward to seeing how this plays out on Twitter, but thanks everyone. Thanks for joining us, Adam. Corey.

Corey (59:04):

It's been fun gentlemen.

Rodrigo Gordillo (59:06):

Yeah, we'll pick another topic next time and riff on it for a couple more hours.

Adam Butler (59:09):

Sounds great. Thanks guys.

Rodrigo Gordillo (59:10):

All right guys.

Corey (59:11):

Thanks.

Speaker 1 (59:13):

Thank you for listening to the Gestalt University podcast. You will find all the information we highlighted in this episode in the show notes at investresolve.com/blog you can also learn more about ReSolve's approach to investing by going to our website and research blog at investresolve.com where you will find over 200 articles that cover a wide array of important topics, the area of investing. We also encourage you to engage with the whole team on Twitter by searching the handle [@investresolve](https://twitter.com/investresolve) and hitting the follow button.



Speaker 1 (59:43):

If you're enjoying the series, please take the time to share us with your friends through email, social media, and if you really learn something new and believe that our podcast will be helpful to others, we would be incredibly grateful if you could leave us a review on iTunes. Thanks again and see you next time.