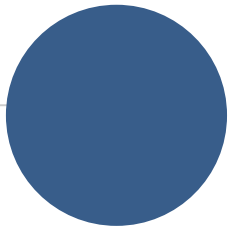




The Alchemy of Data “Be Reasonable”

Ankit Sahni

exante data

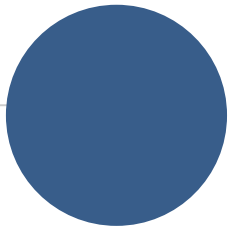


Ankit Sahni: Head of Research, Exante Data

- ❑ Studied Electrical Engineering at IIT Delhi, with a focus on Electronics and Communication. Then studied finance at IIM Bangalore
- ❑ Member of the fixed income research team at Nomura Securities, with a focus on rates and FX markets. #1 ranked currency strategy team *in the Institutional Investor Survey* in 2012 and 2013, as well as the US Governments Strategy team ranked as 'runner-up' in 2012
- ❑ Head of US Strategy at Prologue Capital, a global macro and rates relative value hedge fund; helped manage the firm's major rates and currency portfolios
- ❑ Head of Research at Exante Data, overseeing the development of analytical tools and models, as well as turning model outputs into actionable market insight for clients

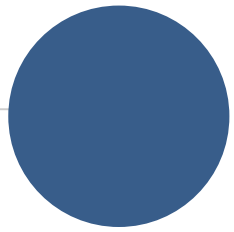
- ┆ **Current ‘state of play’**
- ┆ **Stylized market landscape**
- ┆ **Key issues observed, and solutions**
 1. Dis-economies of scale
 2. Shortage of truly “big” data
 3. Overhype of data quality
 4. Unstructured data is key
- ┆ **Summary**
- ┆ **Example of hybrid approach**

ex ante data

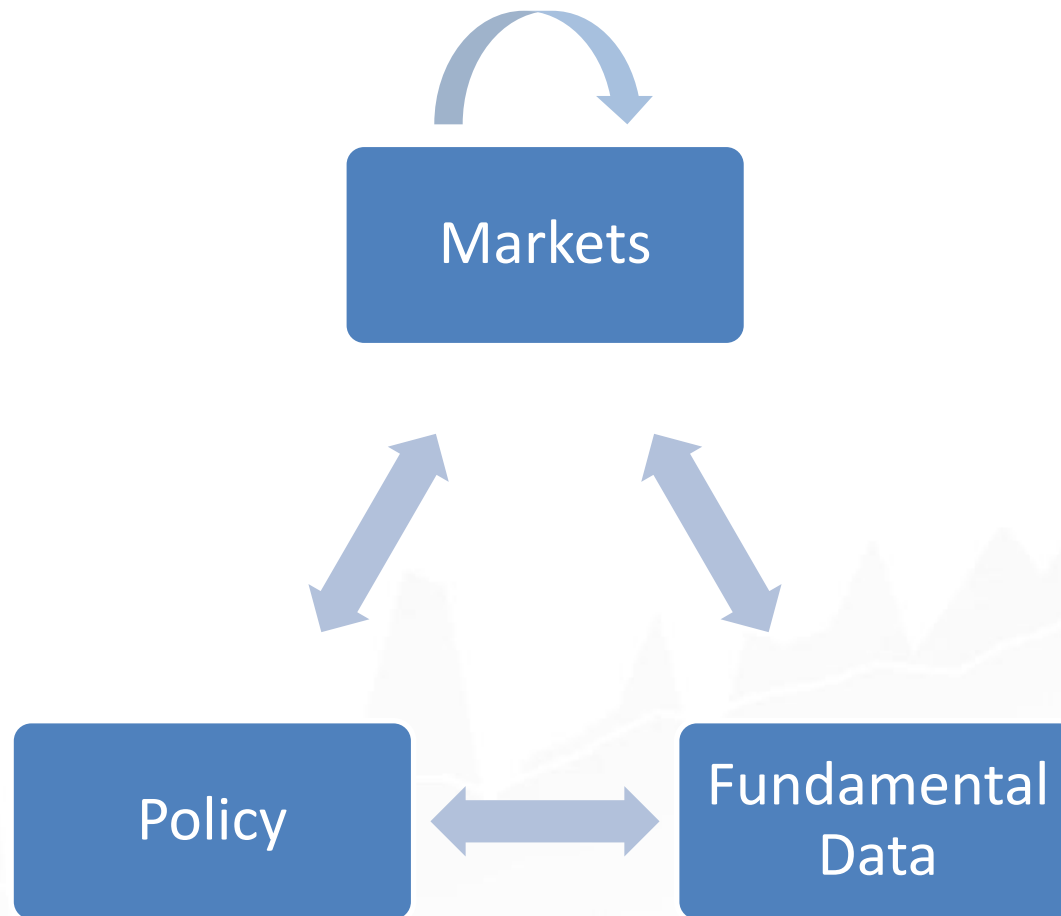


The Context: Data is getting *big* in finance

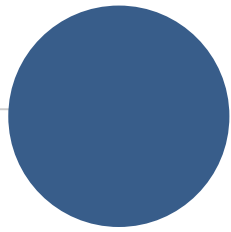
- Oft-quoted statistic, but bears repeating: “90 percent of the data in the world today has been created in the last two years alone”
- Computing power has also increased exponentially, as have access to tools for storing and processing large amounts of data
- In the investment management space, firms increasingly view differentiated tools and data as a source of competitive advantage. Buy side spend on “alternative data sets” has more than doubled in the past 2 years, and growth is accelerating.



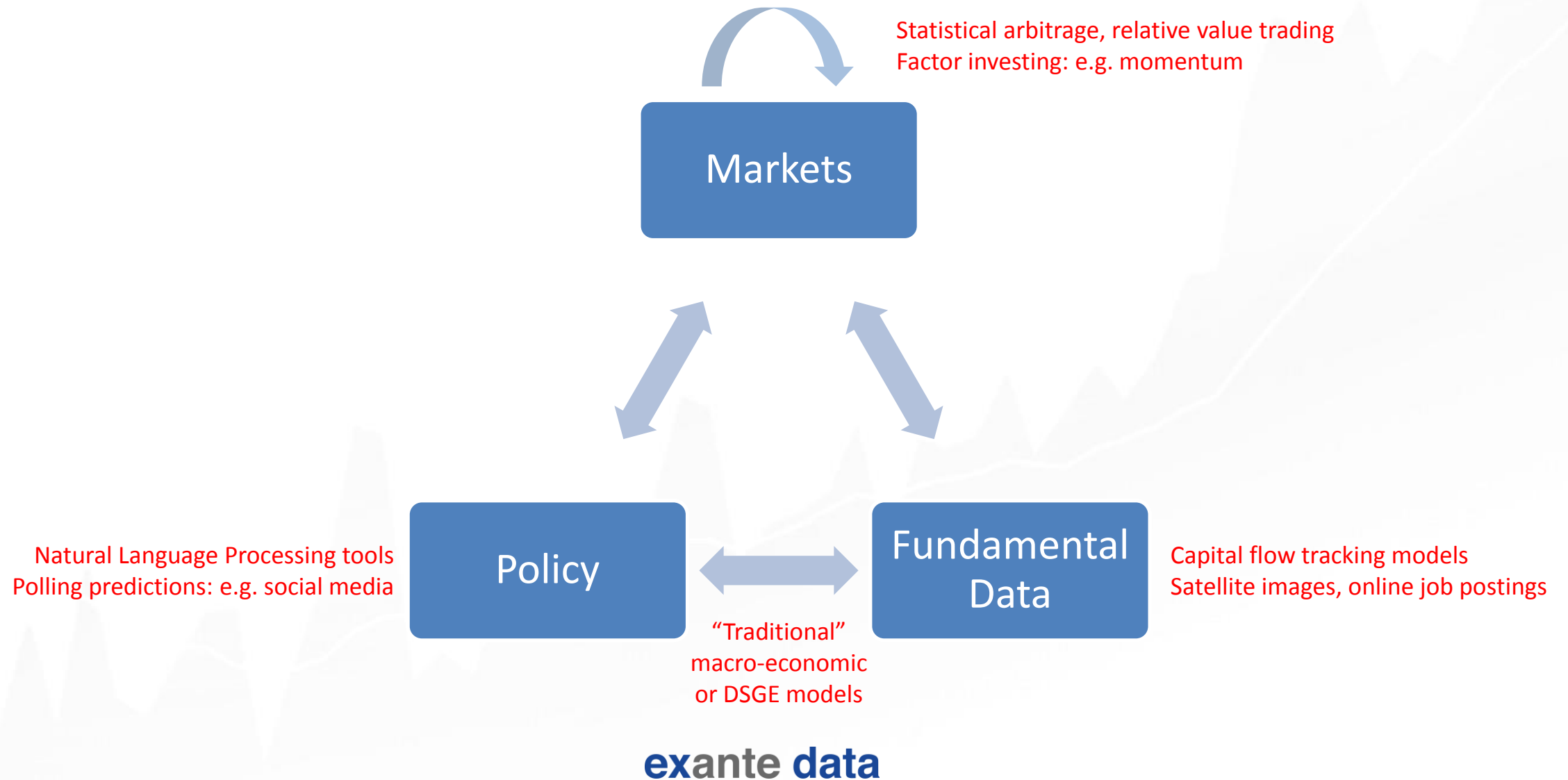
Stylized Market Map

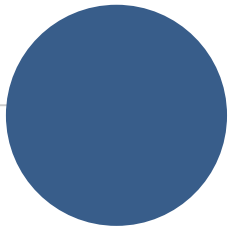


ex ante data



How Better Data Analysis can Help Generate Alpha





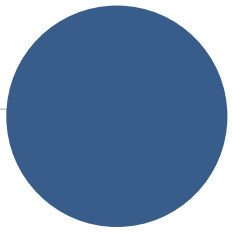
Issue 1: Dis-economies of Scale

Large incumbents in the financial services space suffer from dis-economies of scale, due to:

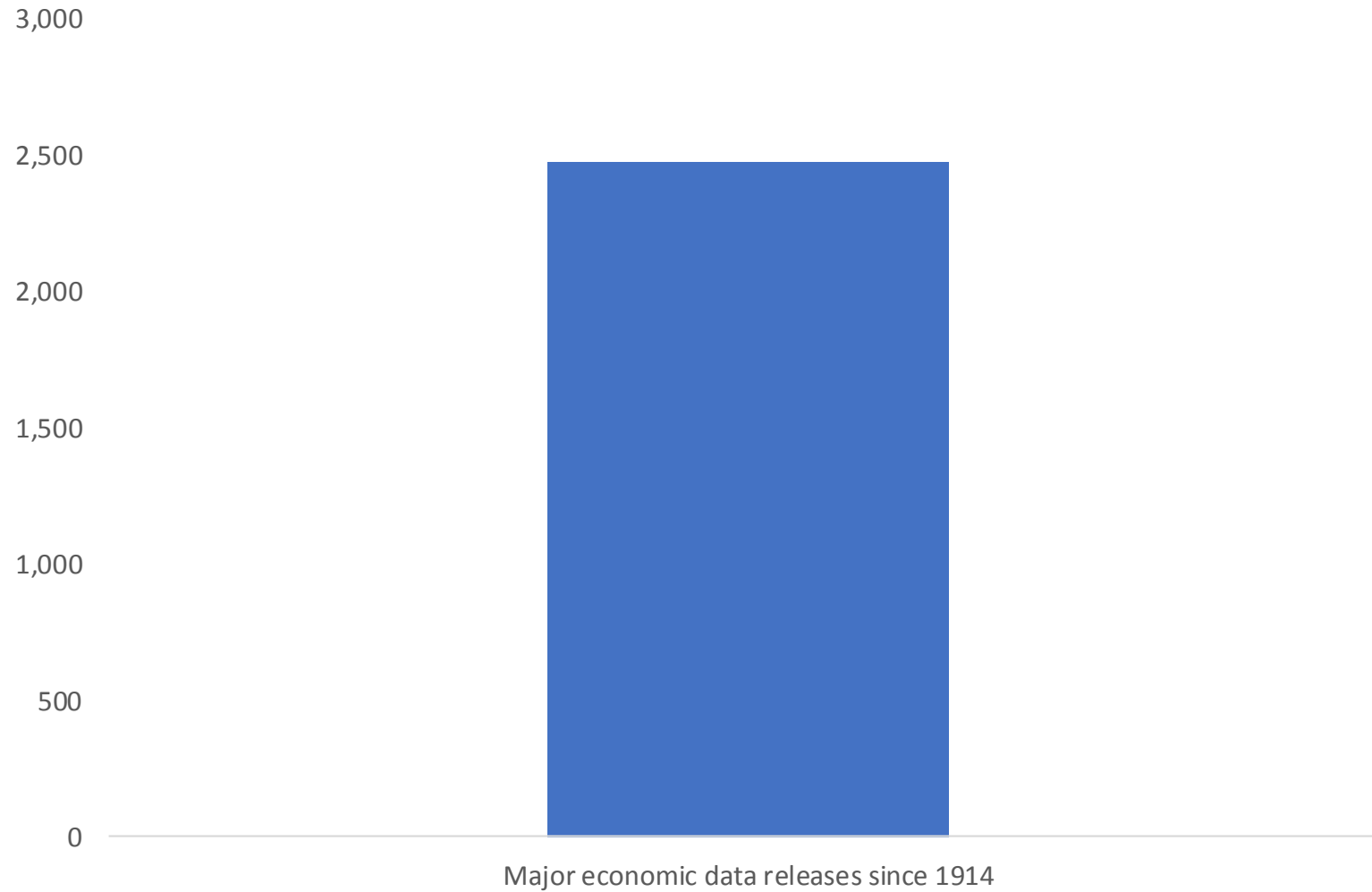
- Regulation
- Inertia
- Cloud computing
- On-demand hiring
- Venture capital

Solution

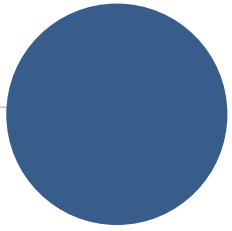
- Keep core functionality in-house, and outsource the rest to *reliable* outside vendors.



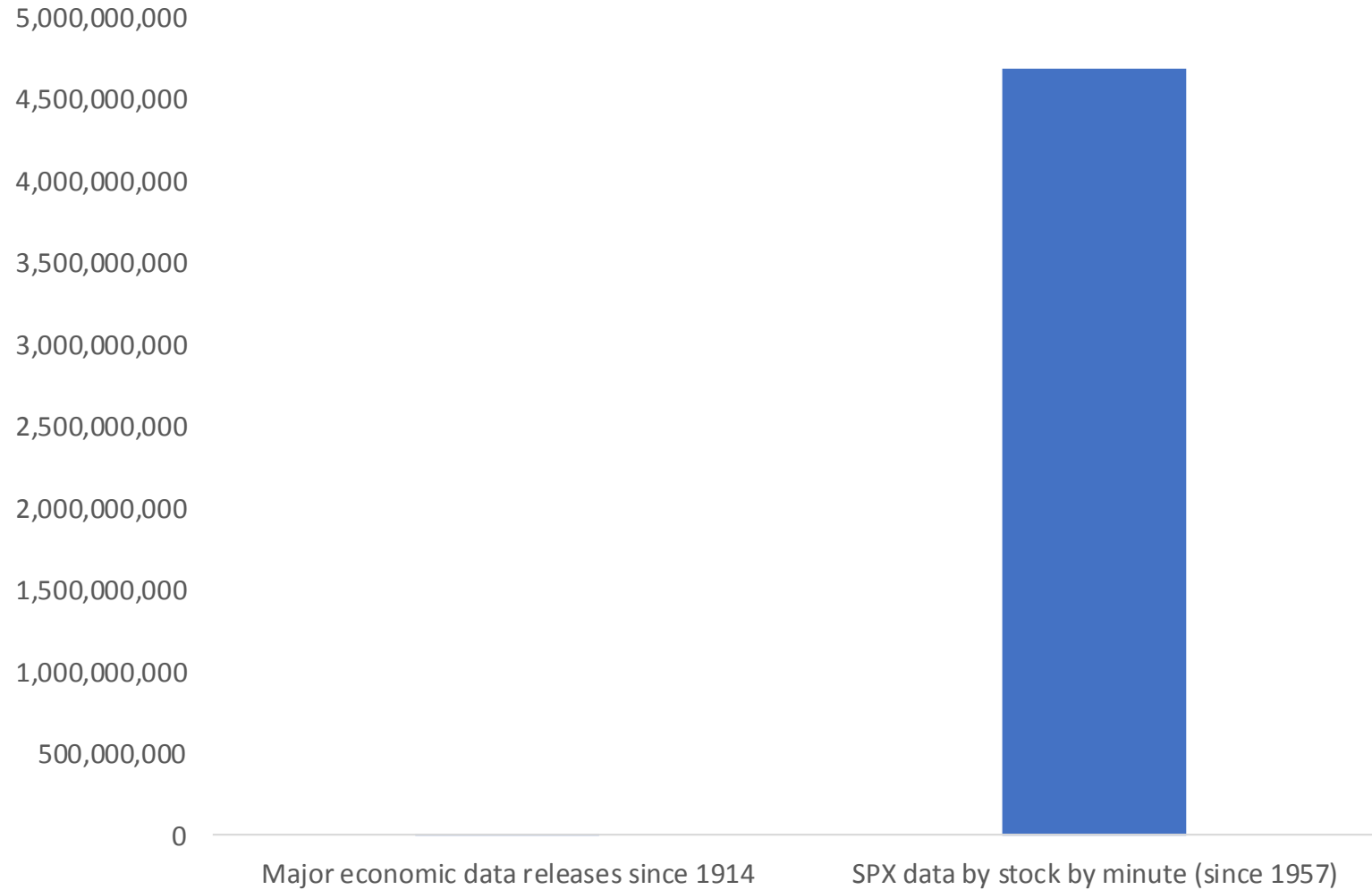
Issue 2: Most alternative data is not “big”

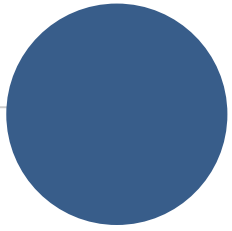


exante data

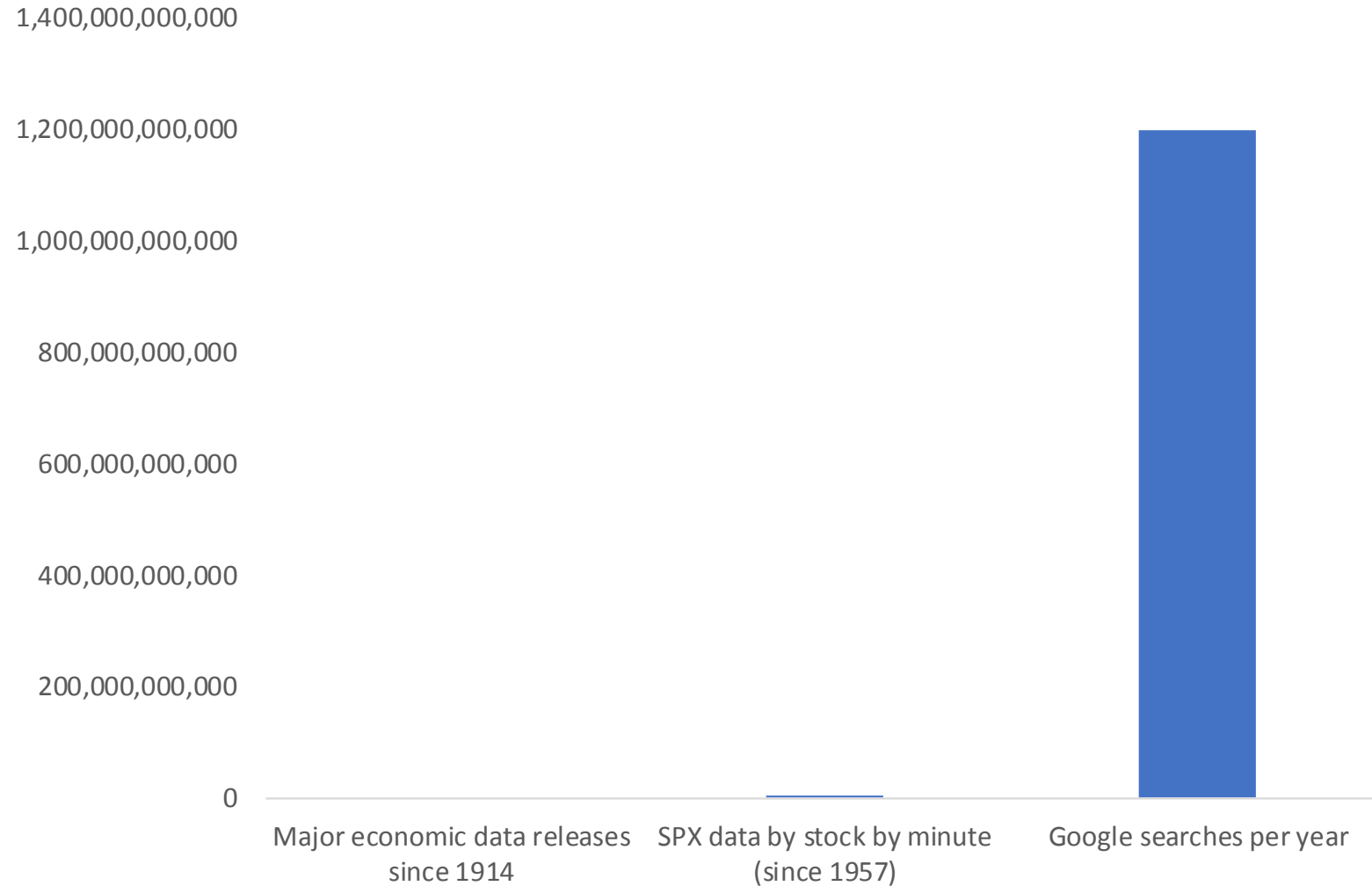


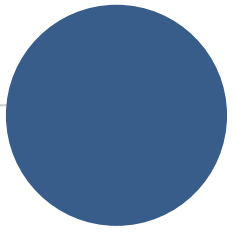
Issue 2: Most alternative data is not “big”





Issue 2: Most alternative data is not “big”



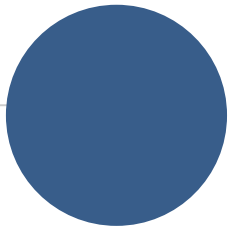


Issue 2: Most alternative data is not “big”

- The shortage of truly “big” fundamental data, combined with frequent “unprecedented” events and regime changes implies that alternative data needs to be handled with care, especially with the availability of highly sophisticated machine learning techniques

Solution

- Use the right tools for the job, which would vary by data set being analyzed and conclusions needed
- Build robustness checks within each data analytics process, including in terms of number of optimization attempts
- Play to the strengths of man and machine

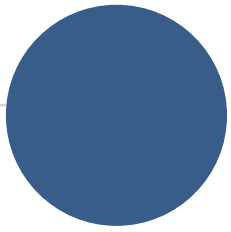


Issue 3: Over-hype of Data Quality

- Alternative data providers have sometimes focused too much on data novelty, and gotten carried away with buzzwords like “big data”, “machine learning” and “artificial intelligence”
- The truth is that the current ‘small’ data sets the bar quite high, such that delivering true insight is difficult and rare

Solution

- For the buy-side: Use current ‘small’ data better, and build in-house analytical expertise to independently verify the quality of new data and insights
- For the data providers: Focus on data quality

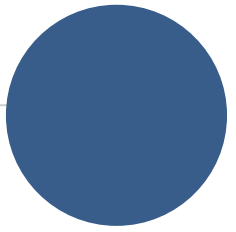


Issue 4: Unstructured data is key

- A significant portion of market-relevant data, especially that related to policy changes and expectations is non-numeric and unstructured, especially in the form of text or sentiment
- Development so far in NLP and social media analytics, as applied to financially relevant data, have not led to very positive results
- Policy communications are also not generally one-dimensional, in terms of output creating issues in mapping to asset prices

Solution

- For now, use hybrid machine + human approaches, as humans remain better at grasping the context
- Over time, look to build tailored tools to answer specific questions



Summary

There is a clear trend towards bigger data sets, and more accessible machine learning tools to analyze them. This is having a major impact in the investing world

To make the most of this revolution, *be reasonable*:

1. Build core competencies in-house, including processes to evaluate alternative data sets and insights thereof
2. Outsource most data creation and collection tasks, given diseconomies of scale
3. Use the right techniques for each data set, to create robust frameworks that are not overfitted
4. Play to the strengths of humans and computers

Disclaimer

This communication is provided for your informational purposes only. In making any investment decision, you must rely on your own examination of the securities and the terms of the offering. The contents of this communication does not constitute legal, tax, investment or other advice, or a recommendation to purchase or sell any particular security. Exante Advisors, LLC, Exante Data, LLC and their affiliates (together, "Exante") do not warrant that information provided herein is correct, accurate, timely, error-free, or otherwise reliable. EXANTE HEREBY DISCLAIMS ANY WARRANTIES, EXPRESS OR IMPLIED. This communication is intended solely for the person to whom it is addressed and is strictly confidential. You may not reproduce its contents in its entirety or in part, or redistribute its content to any party in any form, without the prior written consent of Exante.

The Company

Exante Data provides analytical tools and smart data solutions to institutional investors globally.

- ❑ **Exante Data** was founded in 2016 and has quickly become a trusted provider of analytical solutions to a growing group of the world's most high-profile macro hedge funds and asset managers.
- ❑ The goal of **Exante Data** is to build a bridge between new technology that allows extraction of information from a myriad of new data sources and macro analysis/strategy.
- ❑ **Exante Data** is based in Manhattan, New York and employs a mix of data scientists, programmers and currency strategists.