

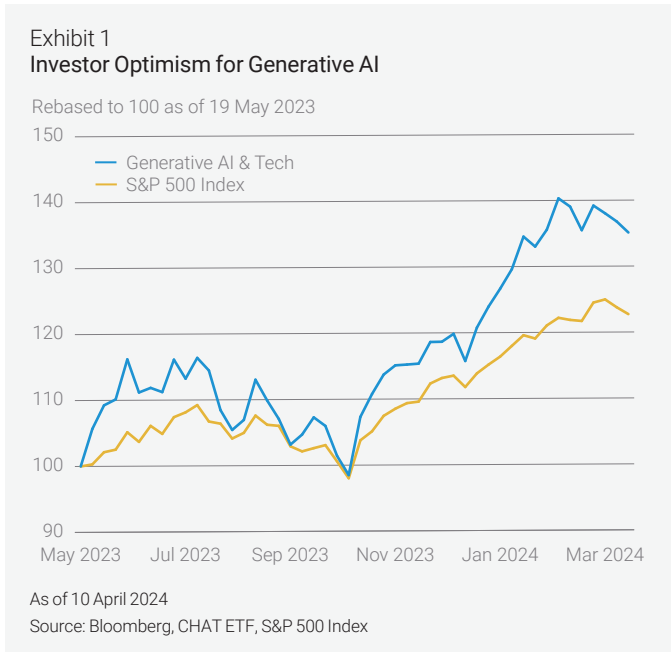


In Praise of Small Language Models

- Advances in large language models (LLMs) have generated excitement among investors, but their application as a reliable investment tool brings challenges.
- Small language models (SLMs) avoid the problems associated with LLMs and can deliver consistent results especially suited to specific questions.
- Our case study on the use of an SLM to identify Chinese companies' exposure to the country's struggling housing market provides a real-world example of SLMs at work.

Large Language Models are miraculous, but ...

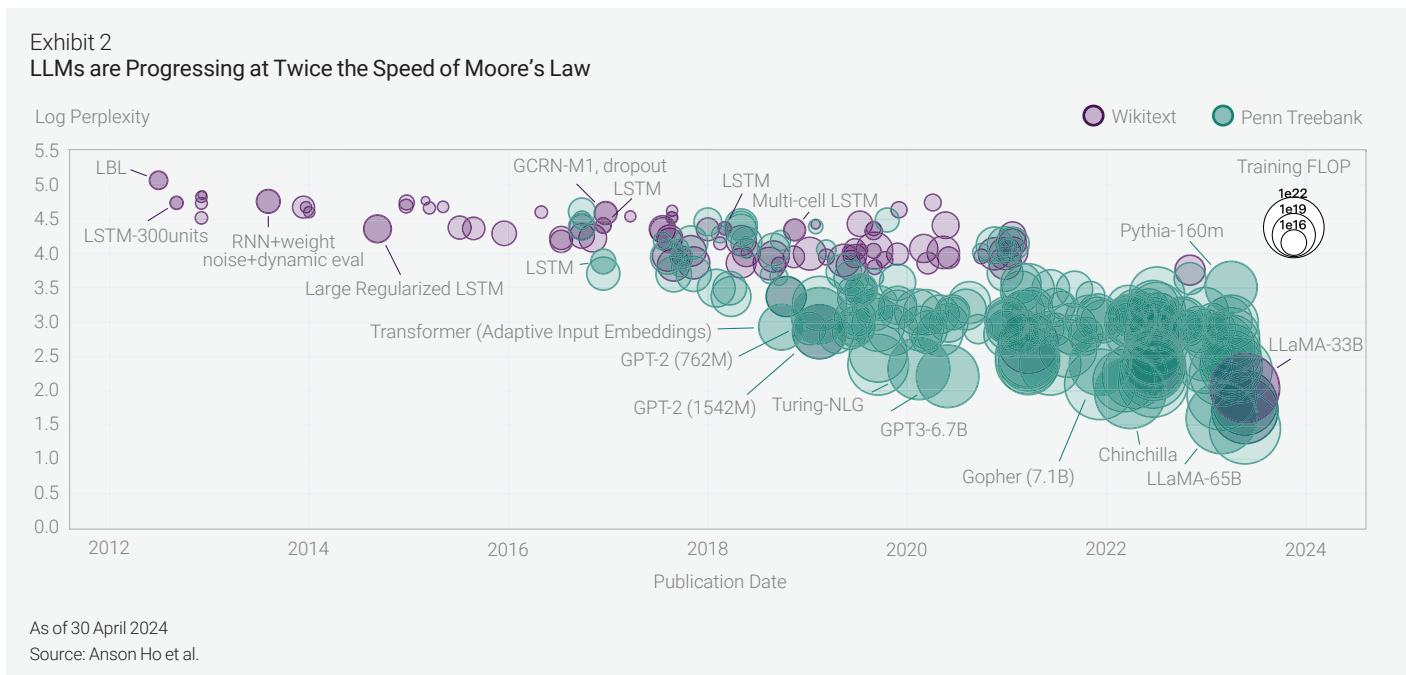
A year on from the release of ChatGPT-4, an artificial intelligence (AI) boom has captivated investors and helped propel global stock markets, led by the US. The Roundhill Generative AI & Technology ETF (CHAT), comprising 25-50 companies involved in generative AI and related technologies outperformed a buoyant S&P 500 index by a substantial margin (13%) in the six months to March 2024 (Exhibit 1).



LLMs, the technology powering this boom, are advancing rapidly. A recent paper highlights that the computational power needed to achieve a given performance threshold is halving every five to 14 months, which is roughly twice the speed of hardware improvements under Moore’s Law (Exhibit 2).¹

And yet two things make it hard to rely on LLMs in the realm of investing: hallucinations and opacity. Hallucination refers to the tendency of LLMs to confidently present made-up information as fact. The problem is part of a broader “alignment” challenge, which aims to make AI reliably follow human intentions. Many researchers think larger, better-quality datasets will fix the issue, but some argue that the inherent generative and probabilistic nature of LLMs makes some hallucinations inevitable. Regardless of which perspective proves correct, the rapid advancement of LLMs suggests that the often embarrassing and costly instances of hallucinations are likely to decrease significantly over time.^{2,3}

The second challenge, opacity, refers to the fact that the models operate as “black boxes.” With billions (sometimes trillions) of parameters, their enormous scale makes it nearly impossible to fully comprehend their internal logic. And because of the labyrinthine network of interconnected parameters in neural networks, it is impractical to distill the system’s overall behavior into a clear set of rules. This inherent difficulty in interpreting how LLMs process language and arrive at conclusions persists even for their creators, who cannot, for example, explain why LLMs respond to emotional pleas.⁴ Reducing scale could improve interpretability but most likely at the cost of capability, as scale enables unique emergent behaviors. And so there are intrinsic trade-offs—between scale, capability, and interpretability—meaning opacity is a structural feature, not a bug, of LLMs.



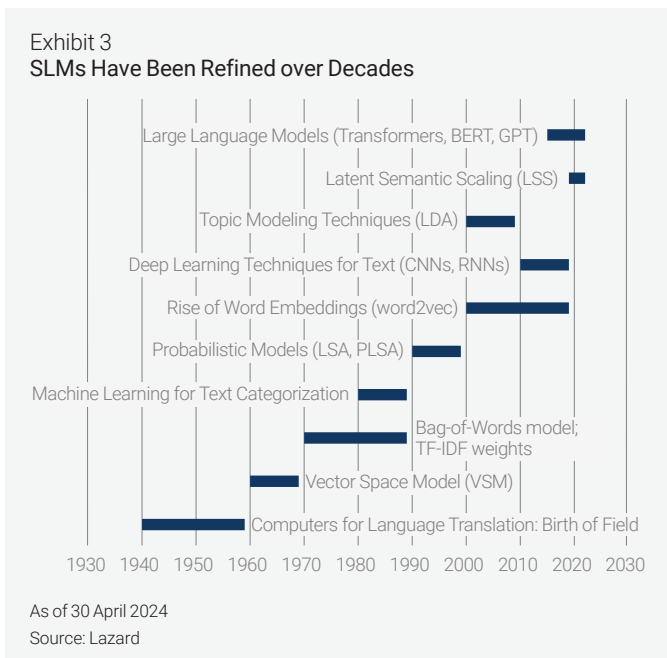
Small Language Models

"I have wondered if it were unthinkable to design a computer which would translate. Even if it would translate only scientific material (where the semantic difficulties are very notably less), and even if it did produce an inelegant (but intelligible) result, it would seem to me worthwhile."

– Warren Weaver, "Translation," 1949.⁵

Three-quarters of a century ago, Warren Weaver's seminal 1949 memorandum laid the groundwork for machine translation, statistical language models, and modern natural language processing (NLP) techniques. Comparing translation to code-breaking, Weaver posited that the "semantic difficulties" of translation, such as multiple meanings, could be solved computationally.

The 1960s saw the birth of dictionary-based quantitative text analysis with the creation of *The General Inquirer* by Stone et al. in 1966. This program searched documents for words registered in a pre-defined keyword list or "dictionary." From these humble beginnings, the following decades witnessed the development of an extensive suite of text processing techniques, which we have collectively named Small Language Models (Exhibit 3).



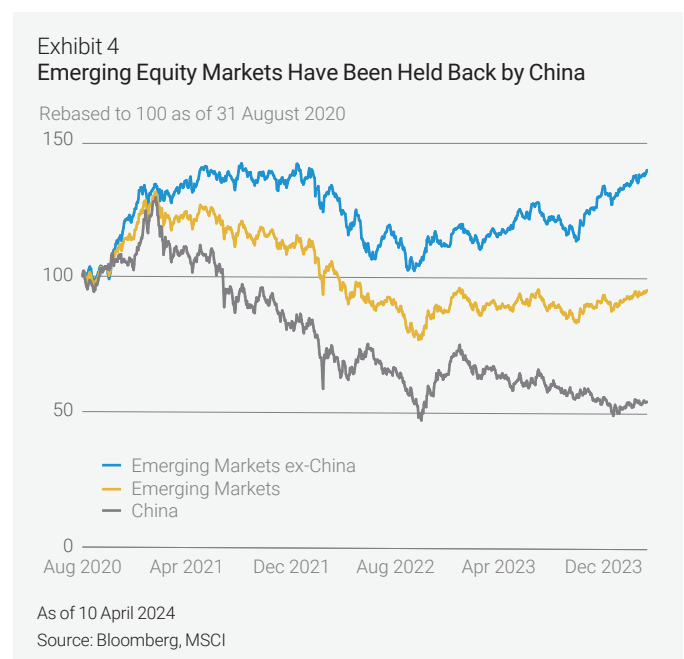
Unlike their large cousins, SLMs do not hallucinate, consistently producing the same outcomes given the same input. They also operate as "glass boxes," with a much smaller number of parameters. One such SLM, developed by Kohei Watanabe, a member of the Lazard Quantitative Equity team, is Latent Semantic Scaling (LSS). LSS is a semi-supervised document scaling technique that locates documents in any language on user-defined dimensions.⁶

To achieve this, LSS combines two elements: the unsupervised Latent Semantic Analysis (LSA) technique and user-provided "seed words." LSA estimates the semantic proximity between words in the data, while the seed words define the dimensions of interest. By measuring the semantic proximity of words to the seed words instead of relying solely on word frequencies, LSS provides a highly interpretable, theoretically grounded, and precise answer to a given question in any language. In the next section, we demonstrate its use in practice via a case study on managing risk emanating from the Chinese housing market.

Case Study: China's Housing Market

In August 2020, the Chinese Government announced its "Three Red Lines" policy, imposing stricter financial requirements on real estate developers to deleverage the sector, tame bubbly property prices, and mitigate systemic financial risks.

China's housing sector has been mired in a prolonged slowdown ever since, with numerous property developers, including Evergrande, Country Garden, and Shimao Group, defaulting on their debts or facing liquidation. This distress threatens household wealth and has been a key driver of the country's economic slowdown, weighing on the broader emerging markets index for the past three years (Exhibit 4).



To help us navigate this “slow-motion financial crisis,”⁷ we applied LSS to annual reports filed in Mandarin (Exhibit 5). By analyzing text in its original form, we can gain a deeper understanding of the nuances, cultural references, and contextual meanings that are often lost in translation. The analysis covered 25,000 documents comprising of 6.5 million pages. Notably, this was achieved using only a handful of parameters and without requiring fluency in Mandarin, because semantic context is extracted statistically.



Exhibit 6 shows the results. Unsurprisingly, the topic is front and center for property developers and the banks that provide them with loans. Industrial companies also feature prominently, with almost twice as many as financial firms having medium-to-high exposure. These include companies involved in engineering, procurement, and construction (EPC) projects, machinery manufacturers such as excavator makers, and firms providing housing-related services like waste management. Facing slowing domestic sales, many are increasingly turning to export markets, leading to anti-dumping probes from both the UK and European Commission.^{8,9}

Crucially, the stock returns of companies outside the real estate and financial sectors have aligned with their degree of exposure to the housing market, as estimated using LSS. Those less exposed have seen better stock performance compared to their more exposed counterparts and the benchmark (Exhibit 7). In addition to being transparent—we can completely unpack and fully understand LSS outputs—this level of precision comes at a fraction of the cost if we were to attempt the same exercise using an LLM (Exhibit 8).

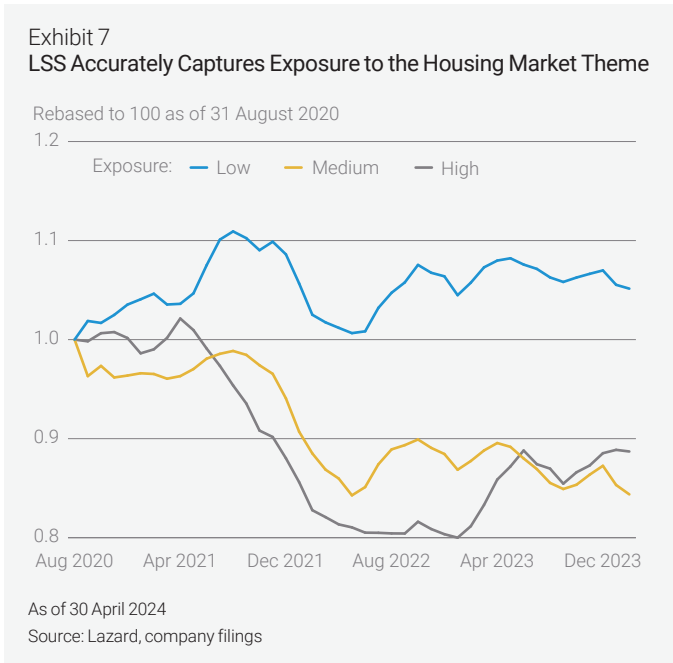
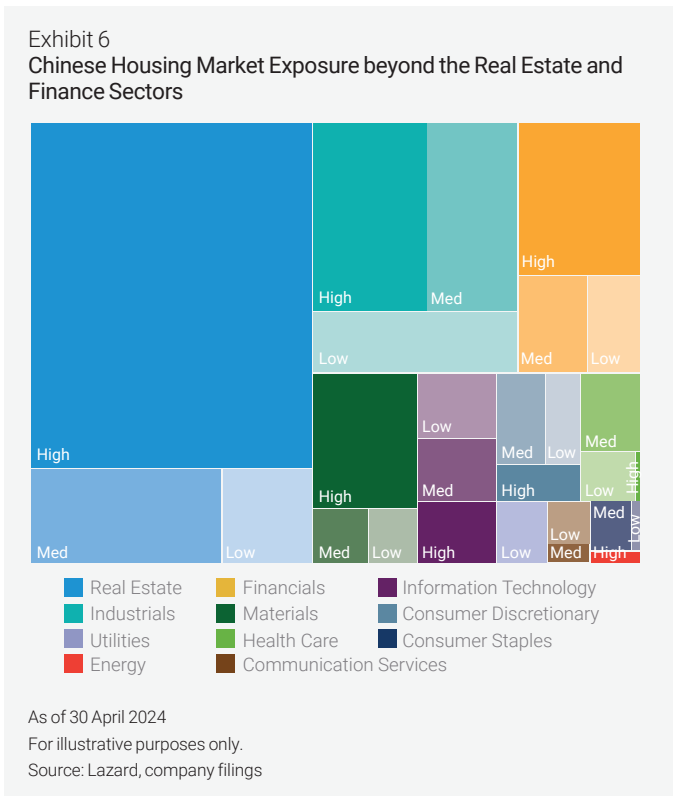
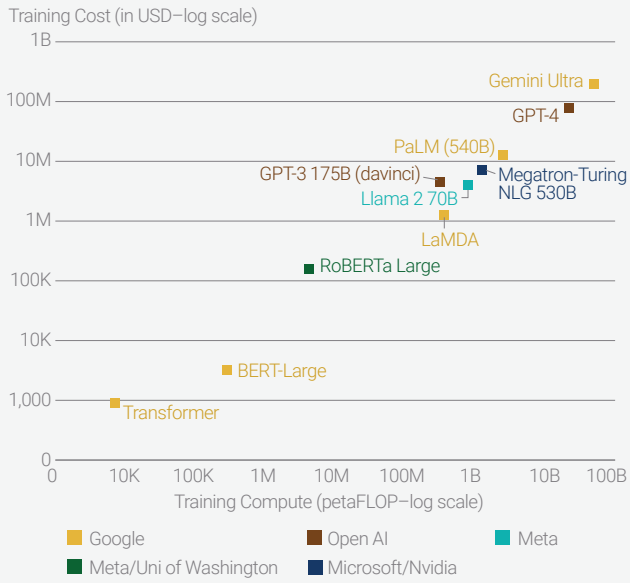


Exhibit 8
Estimated Training Cost and Compute of Select AI Models



As of 30 April 2024

Source: Lazard, company filings

Conclusion

In a world captivated by large language models, it is easy to overlook the value of their smaller cousins. Small language models drawing on techniques refined over 75 years, offer reliability and interpretability—qualities often lacking in LLMs. While LLMs dazzle with their scale and scope, their opacity and tendency to hallucinate limit their utility in the realm of investing. SLMs, in contrast, deliver consistent, tractable results that are especially suited to highly specific questions, as our case study on China's housing market has shown. And so, as the AI race accelerates, investors and researchers should remember that big is not always beautiful.

This content represents the views of the author(s), and its conclusions may vary from those held elsewhere within Lazard Asset Management. Lazard is committed to giving our investment professionals the autonomy to develop their own investment views, which are informed by a robust exchange of ideas throughout the firm.

Notes

- 1 Anson Ho et al. (2024). Algorithmic progress in language models. Retrieved from <https://doi.org/10.48550/arXiv.2403.05812>
- 2 The Guardian (2024). Air Canada ordered to pay customer who was misled by airline's chatbot. Retrieved from <https://www.theguardian.com/world/2024/feb/16/air-canada-chatbot-lawsuit>
- 3 Reuters (2023). New York lawyers sanctioned for using fake ChatGPT cases in legal brief. Retrieved from <https://www.reuters.com/legal/new-york-lawyers-sanctioned-using-fake-chatgpt-cases-legal-brief-2023-06-22/>
- 4 Cheng Li et al. (2023). Large Language Models Understand and Can be Enhanced by Emotional Stimuli. Retrieved from <https://doi.org/10.48550/arXiv.2307.11760>
- 5 Warren Weaver (1955). Translation in Locke, W.N. and Booth, A.D. (Ed.), "Machine translation of languages: fourteen essays" (pp. 15-23). Technology Press of the Massachusetts Institute of Technology, Cambridge, Mass.
- 6 Kohei Watanabe (2021). Latent Semantic Scaling: A Semisupervised Text Analysis Technique for New Domains and Languages. "Communication Methods and Measures," 15(2), 81–102.
- 7 Financial Times (2022). China's property crash: 'a slow-motion financial crisis'. Retrieved from <https://on.ft.com/3UjFAEF>
- 8 Trade Remedies Authority (2023). TRA opens new investigations into excavators from China. Retrieved from <https://www.gov.uk/government/news/tra-opens-new-investigations-into-excavators-from-china>
- 9 Council of the European Union (2023). Notice of initiation of an anti-dumping proceeding concerning imports of mobile access equipment (MAE), originating in the People's Republic of China. Retrieved from <https://eur-lex.europa.eu/eli/C/2023/783/oj>

Important Information

Published on 28 May 2024.

Information and opinions presented have been obtained or derived from sources believed by Lazard Asset Management LLC or its affiliates ("Lazard") to be reliable. Lazard makes no representation as to their accuracy or completeness. All opinions expressed herein are as of the published date and are subject to change.

The performance quoted represents past performance. Past performance does not guarantee future results.

Allocations and security selection are subject to change.

Lazard Asset Management ("Lazard" or the "firm") actively manages client portfolios with the objective of delivering positive investment performance and maximizing long-term shareholder value. Portfolio managers at Lazard have discretion to incorporate ESG considerations into their investment processes, and to what degree. Information concerning a particular investment strategy's utilization of ESG considerations (including its status as ESG Integrated or Sustainability Focused under our procedures) are set forth in Lazard's description of that strategy, or are available upon request. Other disclosures herein may describe sustainable investment views and resources that Lazard's ESG professionals have developed to assist our clients and portfolio management teams. However, unless expressly disclosed, readers should not assume that these views and resources are incorporated in a portfolio management team's investment process. Equity securities will fluctuate in price; the value of your investment will thus fluctuate, and this may result in a loss. Securities in certain non-domestic countries may be less liquid, more volatile, and less subject to governmental supervision than in one's home market. The values of these securities may be affected by changes in currency rates, application of a country's specific tax laws, changes in government administration, and economic and monetary policy. Small- and mid-capitalization stocks may be subject to higher degrees of risk, their earnings may be less predictable, their prices more volatile, and their liquidity less than that of large-capitalization or more established companies' securities. Emerging markets securities carry special risks, such as less developed or less efficient trading markets, a lack of company information, and differing auditing and legal standards. The securities markets of emerging markets countries can be extremely volatile; performance can also be influenced by political, social, and economic factors affecting companies in emerging markets countries.

The S&P 500 Index is a market capitalization-weighted index of 500 companies in leading industries of the US economy. The index is unmanaged and has no fees. One cannot invest directly in an index.

The indices referenced in this document are included merely to show general trends in the market during the periods indicated and are not intended to imply that investments made pursuant to the strategy are or will be comparable to any index in either composition or element of risk. The strategy is not restricted to securities comprising any index. The strategy may use various investment techniques not reflected in an index. The indices referenced herein are unmanaged and have no fees. One cannot invest directly in an index. There is no guarantee that the strategy's performance will meet or exceed any index.

The securities mentioned are not necessarily held by Lazard for all client portfolios, and their mention should not be considered a recommendation or solicitation to purchase or sell these securities. It should not be assumed that any investment in these securities was, or will prove to be, profitable, or that the investment decisions we make in the future will be profitable or equal to the investment performance of securities referenced herein. There is no assurance that any securities referenced herein are currently held in the portfolio or that securities sold have not been repurchased. The securities mentioned may not represent the entire portfolio.

Certain information contained herein constitutes "forward-looking statements" which can be identified by the use of forward-looking terminology such as "may," "will," "should," "expect," "anticipate," "target," "intent," "continue," or "believe," or the negatives thereof or other variations thereon or comparable terminology. Due to various risks and uncertainties, actual events may differ materially from those reflected or contemplated in such forward-looking statements.

This document is provided by Lazard Asset Management LLC or its affiliates ("Lazard") for informational purposes only. Nothing herein constitutes investment advice or a recommendation relating to any security, commodity, derivative, investment management service, or investment product. Investments in securities, derivatives, and commodities involve risk, will fluctuate in price, and may result in losses. Certain assets held in Lazard's investment portfolios, in particular alternative investment portfolios, can involve high degrees of risk and volatility when compared to other assets. Similarly, certain assets held in Lazard's investment portfolios may trade in less liquid or efficient markets, which can affect investment performance. Past performance does not guarantee future results.

This document is intended only for persons residing in jurisdictions where its distribution or availability is consistent with local laws and Lazard's local regulatory authorizations. The Lazard entities that have issued this document are listed below, along with important limitations on their authorized activities.

Australia: Issued by Lazard Asset Management Pacific Co., ABN 13 064 523 619, AFS License 238432, Level 12 Gateway, 1 Macquarie Place, Sydney NSW 2000, which is licensed by the Australian Securities and Investments Commission to carry on a financial services business. This document is intended for wholesale investors only. **Canada:** Issued by Lazard Asset Management (Canada) Inc., 30 Rockefeller Plaza, New York, NY 10112 and 130 King Street West, Suite 1800, Toronto, Ontario M5X 1E3, a registered portfolio manager providing services to non-individual permitted clients. **Dubai:** Issued and approved by Lazard Gulf Limited, Index Tower, Floor 10 Office 1008 DIFC, PO Box 506644, Dubai, United Arab Emirates. Registered in Dubai International Financial Centre 0467. Authorised and regulated by the Dubai Financial Services Authority to deal with Professional Clients only. **EU Member States:** Issued by Lazard Asset Management (Deutschland) GmbH, Neue Mainzer Strasse 75, D-60311 Frankfurt am Main. **Hong Kong:** Issued by Lazard Asset Management (Hong Kong) Limited (AQZ743), Suite 1101, Level 11, Chater House, 8 Connaught Road Central, Central, Hong Kong. Lazard Asset Management (Hong Kong) Limited is a corporation licensed by the Hong Kong Securities and Futures Commission to conduct Type 1 (dealing in securities) and Type 4 (advising on securities) regulated activities only on behalf of "professional investors" as defined under the Hong Kong Securities and Futures Ordinance (Cap. 571 of the Laws of Hong Kong) and its subsidiary legislation. **Korea:** Issued by Lazard Korea Investment Management Co., Ltd., 10F Seoul Finance Center, 136 Sejong-daero, Jung-gu, Seoul, 100-768. **People's Republic of China:** Issued by Lazard Asset Management. Lazard Asset Management does not carry out business in the P.R.C. and is not a licensed investment adviser with the China Securities Regulatory Commission or the China Banking Regulatory Commission. This document is for reference only and for intended recipients only. The information in this document does not constitute any specific investment advice on China capital markets or an offer of securities or investment, tax, legal or other advice or recommendation or, an offer to sell or an invitation to apply for any product or service of Lazard Asset Management. **Saudi Arabia:** Issued and approved by Lazard Saudi Arabia, 15th Floor, Nakheel Tower, King Fahad Road, Nakheel District, Riyadh 12381, Saudi Arabia. Commercial Registration No: 1010297437 with a unified number: 7001642417, License of the Ministry of Investment No: 102031098747. Regulated and authorised by the Capital Market Authority (CMA). **Singapore:** Issued by Lazard Asset Management (Singapore) Pte. Ltd., Unit 15-03 Republic Plaza, 9 Raffles Place, Singapore 048619. Company Registration Number 201135005W, which provides services only to "institutional investors" or "accredited investors" as defined under the Securities and Futures Act, Chapter 289 of Singapore. **Switzerland:** Issued by Lazard Asset Management Schweiz AG, Uraniastrasse 12, CH-8001 Zurich. **United Kingdom:** Issued or approved by Lazard d Asset Management Ltd., 50 Stratton Street, London W1J 8LL. Registered in England Number 525667. Authorised and regulated by the Financial Conduct Authority (FCA), providing services only to persons classified as eligible counterparties or professional clients under the rules of the FCA. **United States:** Issued by Lazard Asset Management LLC, 30 Rockefeller Plaza, New York, NY 10112.