October 31, 2022

Office of Financial Technology

Federal Housing Finance Agency

400 7th Street SW, 5th Floor

Washington, DC 20219

**RE: Fintech in Housing Finance**

Dear Madam or Sir:

To provide context for our response, we offer a brief introduction. Polygon Research is a mortgage data science company based in Washington, DC. Our mission is to empower stakeholders to imagine, design, and create a better housing industry through open data. We do this through high-value, low-cost products and services focused on a simple but powerful value proposition: any to-any-analysis across people, households, loans, lenders, and geographies without predetermined endpoints, via browser or API, across all open data sources.

**General Comments**

Our response focuses directly on your stated goal of “responsible financial innovation that promotes greater and more cost-efficient access to financial products and services”. We believe the path to this outcome requires the elimination of barriers to existing data, and to the curation and publication of more data.

Today there are structural impediments to innovation that may not be obvious at first glance. Fintechs have been producing a dizzying array of products and services over the past 5+ years, but with a concurrent increase in loan production costs and unanswered questions about risk management of their activities.

We believe the best, most direct, and lowest cost answer to both problems (increased costs and increased risks) is open data. The lowest hanging fruit lies within the reach of the FHFA today: remove the fees Fannie Mae and Freddie Mac currently charge for commercial use of loan-level performance data. These fees are rounding errors within VC-funded budgets, but not so for self-funded companies. This results in a perhaps inadvertent concentration of fintech activity and even of the definition of what it means to be a fintech. With your next steps as you evaluate the responses to this RFI, we encourage and call on you **not to pick winners** by focusing the market on particular solutions which a concentrated set of market participants are financially optimized to pursue, but rather to lower barriers to entry for all participants. Put differently, equity in access to mortgage finance will not come from non-equitable industry arrangements**.** This is not to say that data alone will mitigate risk, improve equity, and lower costs. Innovation is needed among commercial and non-commercial participants alike. But, summarizing, we believe your actions in creating more truly open data is the first and best action you can take to drive these outcomes.

Once you gain equal footing with your peers at e.g. CFPB (HMDA), FFIEC (CRA), Ginnie Mae, and HUD, and the VA in publishing open data, many opportunities for improvement will persist in the form of more data and better granularity. We enumerate some of these opportunities in our answers to specific questions within the RFI.

**A: Fintech and Innovation**

**Question A.2** How could FHFA facilitate adoption of “responsible innovation”?

The three characteristics of responsible innovation are sunshine, openness, and collaboration. To drive responsible innovation, FHFA should advocate for and engage in providing access to more data. Innovation requires the ability for diverse stakeholders to exchange ideas and drive solutions. We participate in an industry that produces a great deal of data about the property, the consumer, the lender, the loan products, and the geography where market participants compete and serve the end-consumer. Democratizing access to Fannie and Freddie’s micro data sets will enable market participants to create cheaper, faster, and better solutions.

The measure of whether these innovations are responsible should be based on both how they were developed and on their impact on consumers.

* *Data transparency:* better information on loan outcomes – what happens after the loan is originated – will inform FHFA whether market participants are innovating responsibly or not. If patterns emerge where consumers are steered into situations that create more costs or prematurely force them out of their home, then FHFA (and its sister agencies) would be able to glean insights around the responsibility of any fintech innovation.
* *Data provenance:* declaration of data sources. Data about the consumer has been treasured by both the primary and secondary market. Data is aggregated by fintechs from multiple sources in order to produce an insight, model a lead, or aid decisions during the mortgage manufacturing process and/or loan administration. When consumers allow their data to be used explicitly, there is a clear provenance of the data. Too frequently, though, consumer data comes from gray area sources with questionable ethics. We believe that asking or requiring fintechs to declare their data sources will be one of the best guardrails to ensure that any innovation will not harm the consumer. If the consumer is unaware that their data is being used to offer an ostensibly innovative way of serving their credit needs or processing their loans could entail unintended impact and place the consumer in financial harm. Declaring the data sources will ensure the consumer and other stakeholders clearly understand and can see the link between data, fintech innovation, and long-term results.

**Question A.3** What factors currently inhibit the adoption of fintech and innovation in the primary and secondary housing finance sector? Are there specific challenges related to privacy laws, industry standards, or current practices?

Frequently there is impedance between new solutions and the readiness of organizations to adopt them. One important aspect of this is data literacy – the readiness and ability of stakeholders to understand data and to derive its full utility.

**B: Identifying Fintech Opportunities in the Housing Finance Ecosystem**

**Question B.4** What are the existing data challenges that most prevent data-driven decision-making in the mortgage lifecycle?

Fannie and Freddie publicly disclose multiple data sets but to use these data sets commercially, industry participants have to pay a fee. The process of obtaining permission and paying the annual fee (upwards of $50k per GSE) is cutting off information access of smaller companies that might contribute to developing solutions. An example of a contrasting approach – the one we are advocating – is that Ginnie Mae, which provides their MBS disclosures free of charge, without terms and conditions.

But data doesn’t just drive solutions, it also helps clarify and define problems. Primary and secondary market data sources have mutually exclusive strengths and gaps. HMDA excels in demographic data. Fannie and Freddie performance data excels in time grain and the inclusion of the credit score. To drive improvement of decisions-making on fairness and effectiveness across the entirety of the mortgage lifecycle – and to clearly define the boundaries of problems that need to be solved – we advocate for the enrichment of each data set with its respective counterpart’s unique strengths: improved time grain and credit score disclosure for HMDA; demographic information for agency data. This level of public end-to-end disclosure of mortgage activity and outcomes will create transparency and will provide a missing accountability feedback loop for all stakeholders in the mortgage industry.

**Question B.5** What are the existing regulatory and policy barriers to adopting and implementing fintech within the mortgage lifecycle?

We recognize that our recommendation in B.4 would have to be worked out through the regulatory framework, and that a chief legacy objection in this context is data privacy. We strongly believe in consumer-level data privacy. But ending the discussion at this objection veils the reality that many fintechs already obtain and incorporate into their products data that is considered private. We believe the path forward is to synthesize the discussion of this barrier with the dialog you are facilitating in this RFI about responsible innovation. Echoing our feedback from A.2., a starting point is for fintechs to declare their data sources, and for all of us to incorporate this new level of transparency into traditional debates on data privacy within government/agency data.

**C: Equitable Access to Mortgage Credit**

**Question C.3** Are there effective ways to identify and reduce the risk of discrimination, whether during development, validation, revision, and/or use fintech models or algorithms? Please provide examples if available.

Our answer to this question ties back to our answers for section B. When fintechs develop models and algorithms, they need consumer and market data either for designing and testing their solutions. In the primary market, we have seen the use of the demographics data both from HMDA and from Census surveys (e.g. ACS) to ensure lending patterns are equitable. For example, our own product HMDAVision is one such solution that lowers the cost of analysis and provides maximum visibility into gaps for all stakeholders in the mortgage banking sector.

However, the secondary market does not produce comparable data so the visibility into potential discriminatory patterns using innovative fintech solutions based on secondary market data is non-existent today. Only when we are able to trace patterns from the application all the way to the end of the loan will FHFA, its peer agencies, and all stakeholders have a measure of understanding of how various fintech solutions impact families.

**F. Office of Financial Technology Activities and Stakeholder Engagement**

**Question F.1** What forms of stakeholder engagement are most effective in facilitating open, timely, and continuous discussion on the challenges and opportunities presented by the application of fintech to housing finance.

We commend your launch of the FHFA Office of Financial Technology, and encourage the ideas you are circulating in the RFI in your introduction to section F. We have 2 suggestions in this context.

First, should you launch a tech sprint, please require that solutions developed during the sprint be published in full detail – i.e. open sourced. This facilitates true collaboration and innovation, rather than the promotion of pre-existing products.

Second, we believe that the forms of stakeholder engagement most effective in facilitating open, timely, and continuous discussion on the challenges and opportunities presented by the application of fintech to housing finance are those based on transparent, well understood insights with clear lineage derived from open data. We encourage your own innovation in this regard, and look forward to the results.

Thank-you for your consideration of our responses. Best regards,

Greg Oliven, CTO

Polygon Research