

Reproduced with permission from Electronic Commerce & Law Report, 21 ECLR, 6/8/16. Copyright © 2016 by The Bureau of National Affairs, Inc. (800-372-1033) <http://www.bna.com>

BIG DATA

Big data analytics can provide numerous benefits to businesses and consumers. But inaccuracies and biases may put businesses at risk of “digital redlining”—potentially discriminating against certain groups of people. Attorneys from Wiley Rein LLP discuss how the private sector can continue to use big data while reducing the risk of federal investigations and liability over disparate impact discrimination.

Consumer Protection

‘Digital Redlining’: Increasing Risk of ‘Soft Regulation’ and Litigation



BY STEPHEN J. OBERMEIER, MEGAN L. BROWN,
MATTHEW J. GARDNER AND STEPHEN J. KENNY

Stephen J. Obermeier, Megan L. Brown, Matthew J. Gardner and Stephen J. Kenny are attorneys at Wiley Rein LLP in Washington. Mr. Obermeier is a partner in the Appellate, Litigation, and Telecom, Media & Technology practices. Ms. Brown is a partner in the Cybersecurity, Appellate, and TMT practices. Mr. Gardner is of counsel in the White Collar Defense & Government Investigations and Cybersecurity practices. Mr. Kenny is an associate in the Election Law & Government Ethics and Litigation practices.

Several recent developments suggest that private sector users and purveyors of “big data” and related services may be at risk of increased scrutiny over its downstream effects. Commentators and researchers have begun to scrutinize the risks and societal effects of big data, and the federal government is raising concerns about potential harms. The private sector should expect increased attention to practices, as regulators try to nudge business practices in preferred directions and potential litigants look for cases to bring.

As the use of big data has become ubiquitous, commentators are increasingly discussing potential risks and downsides, including the possibility of “digital redlining”—any use of big data that can have a disparate impact on or result in unfair treatment of certain groups. These commentators can point to reports by the Federal Trade Commission and the White House to par-

tially justify their concerns. In January 2016, the FTC continued its trend of monitoring and analyzing innovations in big data that impact consumers. In a report, *Big Data: A Tool of Inclusion or Exclusion*, the FTC identified benefits from big data, but also highlighted the fear that “potential inaccuracies and biases might lead to detrimental effects for low income and underserved populations” (21 ECLR 57, 1/13/16)

The concept has been on the White House’s radar for at least two years. A 2014 White House report summarizing a “90-day study to examine how big data will transform the way we live and work and alter the relationships between government, citizens, businesses, and consumers” explained that big data can be beneficial in generating customized search results and advertisements, but also noted that “perfect personalization . . . leaves room for subtle and not-so-subtle forms of discrimination in pricing, services, and opportunities.”

To the extent data personalization results in a claimed disparate impact on protected classes, a variety of players—from banks to e-commerce vendors to real estate companies—could without any intent to discriminate be exposed to discrimination lawsuits by both private plaintiffs and the government. Fortunately, there are steps businesses can take—and defenses they can raise—to help fend off these types of suits.

The White House Signaled Concern About Digital Redlining in 2014

The 2014 White House Big Data report provides a high-level analysis of the potential benefits and threats to big data use in both the government and private sector. In its analysis of the private sector, the report notes that “the civil rights community is concerned that . . . algorithmic decisions [resulting in consumer personalization] raise the specter of ‘redlining’ in the digital economy—the potential to discriminate against the most vulnerable classes of our society under the guise of neutral algorithms.” The report continues:

Recently, some offline retailers were found to be using an algorithm that generated different discounts for the same product to people based on where they believed the customer was located. While it may be that the price differences were driven by the lack of competition in certain neighborhoods, in practice, people in higher-income areas received higher discounts than people in lower-income areas.

Perhaps the most worrisome aspect of the government’s expanding interest in digital redlining is that at least in some circumstances, a company may be subject to liability regardless of whether there was any intent to discriminate.

As a result, the report found it important “to examine how algorithmically-driven decisions might exacerbate existing socio-economic disparities” including in workplace and educational contexts, “especially when it comes to the practice of differential pricing and other

potentially discriminatory practices.” It called for the federal government to expand its technical expertise “to be able to identify practices and outcomes facilitated by Big Data analytics that have a discriminatory impact on protected classes.”

The FTC Recently Expanded Upon the White House’s Analysis

In its most recent publication concerning big data, the FTC expanded upon the White House’s warnings about digital redlining. The FTC’s report describes in more detail the potential for digital redlining and identifies potentially applicable legal regimes, including the Equal Credit Opportunity Act (ECOA), the Fair Credit Reporting Act (FCRA), Title VII of the Civil Rights Act of 1964, the Fair Housing Act (FHA), the Americans with Disabilities Act, the Age Discrimination in Employment Act, and the Genetic Information Nondiscrimination Act. It also identifies the FTC’s general authority over unfair and deceptive practices under Section 5 of the Federal Trade Commission Act, explaining that “Section 5 may . . . apply . . . if products are sold to customers that use the products for discriminatory purposes.”

Although nothing in the report is *per se* binding on private entities, the FTC offers “considerations” for companies using big data, and poses “questions” for “compliance.” It targets downstream users of data, the aggregators of data, and the creators of the analytics. Some of the FTC’s suggestions seem tailor-made for future investigations, inquiries or creative litigants.

Because the FTC has broad jurisdiction and has signaled interest, it will expect companies to be cognizant of its expectations. The FTC explained to a court of appeals in the context of data security expectations, “any careful general counsel would be looking at what the FTC is doing, [as the FTC] has broad ranging jurisdiction [over the private sector] and undertakes frequent actions against all manner of practices and all manner of businesses.” Such admonitions may have consequences for innovation; as companies manage concerns about liability or the burdens of an investigation into their technical and business decisions.

One potential chilling effect of these government warnings animated the separate statement of FTC Commissioner Maureen Olhausen, in which she reiterated the discipline imposed by the market, to guard against inaccuracies or misuse, and cautioned against giving “undue credence to hypothetical harms” which might discourage the development of innovative tools using big data.

Lawsuits and Government Investigators May Try to Hold Private Entities Liable for Unintended Discriminatory Impacts Resulting From Big Data Analytics

Perhaps the most worrisome aspect of the government’s expanding interest in digital redlining is that at least in some circumstances, a company may be subject to liability regardless of whether there was any intent to discriminate under the so-called “disparate impact” theory of discrimination. That is, facially neutral practices that result in a disparate impact on minorities can

be deemed to be unlawful regardless of discriminatory intent. Discrimination claims based on disparate impact have been available in the employment context even in the absence of express statutory language since the U.S. Supreme Court's seminal decision in *Griggs v. Duke Power Co.*, 401 U.S. 424 (U.S. 1971). And just this last term, in *Texas Dep't of Hous. & Cmty Aff's v. Inclusive Cmty's Project Inc.*, 135 S.Ct. 2507, 2015 BL 203075 (U.S. 2015), the Supreme Court held that discrimination claims based on disparate impact are available under the Fair Housing Act notwithstanding the lack of any express statutory liability.

There is no reason to believe that plaintiffs' and governmental application of disparate impact liability will stop with the FHA. For example, although the ECOA contains even less of a textual basis for disparate impact liability than does the FHA, several administrative agencies have taken the position that the ECOA allows for such liability.

The Consumer Financial Protection Bureau (CFPB), relying on the ECOA's legislative history, has stated: "The Act and regulation may prohibit a creditor practice that is discriminatory in effect because it has a disproportionately negative impact on a prohibited basis, even though the creditor has no intent to discriminate and the practice appears neutral on its face, unless the creditor practice meets a legitimate business need that cannot reasonably be achieved as well by means that are less disparate in their impact."

Similarly, the Board of Governors of the Federal Reserve System also interprets the Act to allow disparate impact claims: "The legislative history of the Act indicates that the Congress intended an 'effects test' concept, as outlined in the employment field by the Supreme Court in the cases of *Griggs v. Duke Power Co.*, 401 U.S. 424 (1971), and *Albemarle Paper Co. v. Moody*, 422 U.S. 405 (1975), to be applicable to a creditor's determination of creditworthiness." And in its big data report, the FTC took the position that big data analytics causing a disparate impact could result in liability under the ECOA:

Disparate impact analysis has important implications for big data. Under such an analysis, a company that avoids, for example, expressly screening job applicants based on gender and instead uses big data analytics to screen job applicants in a way that has a disparate impact on women may still be subject to certain equal employment opportunity laws, if the screening does not serve a legitimate business need or if the need can reasonably be achieved by another means with a smaller disparate impact. Likewise, if a company makes credit decisions based on zip codes, it may be violating ECOA if the decisions have a disparate impact on a protected class and are not justified by a legitimate business necessity. Even if evidence shows the decisions are justified by a business necessity, if there is a less discriminatory alternative, the decisions may still violate ECOA.

It isn't only the government that is tracking the impacts of big data on minorities, and even seemingly benign uses of data personalization can result in negative scrutiny

The CFPB's recent enforcement actions against auto-lending finance companies, although not involving big data digital redlining, may provide a glimpse into how the federal government might pursue a private entity whose data analytics arguably caused a disparate impact on consumers. The CFPB determined after an investigation that certain auto-lending companies had engaged in a pattern or practice of lending discrimination in violation of the ECOA because minorities paid on average a higher interest rate for their auto loans than did white borrowers.

Even though the lenders didn't track concrete data on the race and gender of borrowers, the CFPB used so-called "proxy data" such as surnames and geographic location to determine race and gender and calculate disparities in interest rates. The CFPB found no intent to discriminate by the auto-lending companies, and in fact, the companies did not even loan directly to the customers. Rather, the companies underwrote the loans, and the auto dealers had discretion to increase the rates on the loans up to a certain percentage.

However, because this neutral policy had a disparate impact on the interest rates of minorities, the CFPB referred the cases to the Civil Rights Section of the Department of Justice, which ultimately entered into consent decrees with the auto-lending companies resulting in the payment of millions of dollars in penalties and restitution. One of the conditions of these settlements with the lenders is that they must take "proactive steps . . . that directly address fair lending risk by substantially reducing or eliminating discretionary pricing and compensation systems."

Moreover, it is not just the federal government that could pursue big data users for theories of disparate impact discrimination. ECOA, for example, allows for private rights of action, including class actions. And state antidiscrimination laws in places like New Jersey could also form the basis for discrimination suits both for state attorneys general or class action plaintiffs. Indeed, an aggressive plaintiffs' bar can be expected to pursue every possible avenue for large disparate impact cases based on big data analytics.

Pro Publica's Princeton Review Findings Demonstrate Potential for Big Data Analytics to Result in More Than Just Legal Liability

A recent report by the online newsroom Pro Publica demonstrates that it is not only the government that is tracking the impacts of big data on minorities, and even seemingly benign uses of data personalization can result in negative scrutiny. Three college students published a report on Technology Science purporting to demonstrate that The Princeton Review offered higher

prices for its online SAT tutoring courses to residents of higher-income geographical areas.

Using the data collected in the Technology Science report as well as some additional data, the authors of the Pro Publica report “tested whether The Princeton Review prices were tied to different characteristics of each ZIP code, including income, race and education level.” It found that “[w]hen it came to getting the highest prices, living in a ZIP code with a high median income or a large Asian population seemed to make the greatest difference.” More specifically, the report concluded:

Customers in areas with a high density of Asian residents were 1.8 times as likely to be offered higher prices, regardless of income. For instance, residents of the gritty industrial city of Westminster, California, which is half Asian with a median income below most, were charged the second-highest price for the Premier tutoring service.

The Princeton Review didn’t dispute the findings outright, but it responded in a statement that “[t]o equate the incidental differences in impact that occur from this type of geographic based pricing that pervades all American commerce with discrimination misconstrues both the literal, legal and moral meaning of the word.”

Businesses Using Big Data Analytics Can Reduce Risk of Being Subject to Disparate Impact Discrimination Inquiries and Liability

In light of this precedent and the statements made in the White House and FTC big data reports, it would not be surprising to see the FTC, the DOJ, the CFPB, or even a state attorney general initiate some investigatory activity into facially neutral data analytics practices that result in disparate impact of price and other steering on protected classes. Moreover, there exists the potential for private class action suits based on disparate impact discrimination claims resulting from alleged digital redlining. Fortunately, there are steps big data users can take to help avoid a government investigation or litigation.

Foremost, training of engineers and other data analytics staff on the potential for algorithms to unintentionally harm protected classes could help the staff closest to the issue craft algorithms that attempt to avoid unintended effects and recognize any problems should they arise unexpectedly. Moreover, institution of a policy for responding to identified problems will help ensure that any possible harm from unintended disparate impacts can be minimized.

Such a policy might establish steps for identifying the cause of the disparate impact and provide for a determination of whether the analytics can be revised to eliminate the result without undermining the algorithm itself. While these steps would not constitute safe harbors *per se*, they would demonstrate a sensitivity to the issue, which could go a long way toward reaching a favorable result with government investigators should they identify a disparate impact resulting from the company’s practices.

Taking these steps would also provide for a swift and strong response should a company find itself in the crosshairs of a media investigation like The Princeton Review.

In addition, it would be advisable for big data businesses to include in their privacy policies an arbitration

clause establishing that any dispute would be adjudicated in individual arbitration (as opposed to class litigation or arbitration). The Supreme Court has held that such clauses are enforceable, and individual arbitration can be a good way for both parties to resolve small-value disputes without large and costly class litigation. Such an arbitration clause could also make the business less of a target from an aggressive plaintiff’s bar looking for high-value cases.

Big data companies should also consider the availability of professional liability insurance to cover the risk of unintentional disparate impacts resulting from big data algorithms. Depending on the industry involved, such policies may be available in the marketplace.

Should a business find itself the subject of a disparate impact lawsuit based on its big data analytics, it would likely have strong defenses on the merits as well. Although the Supreme Court’s willingness to read a disparate impact cause of action into the Fair Housing Act is concerning, the Court in *Texas Department of Housing* did reaffirm the limits on the reach of disparate impact liability.

Importantly, the Court explained that, consistent with analysis of disparate impact claims under Title VII, the Department of Housing and Urban Development regulations properly provide that “[a]fter a plaintiff does establish a *prima facie* showing of disparate impact, the burden shifts to the defendant to ‘prov[e] that the challenged practice is necessary to achieve one or more substantial, legitimate, nondiscriminatory interests.’” 135 S. Ct. at 2514-15 (quoting 24 C.F.R. § 100.500(c)(2)). “Once a defendant has satisfied its burden at step two, a plaintiff may ‘prevail upon proving that the substantial, legitimate, nondiscriminatory interests supporting the challenged practice could be served by another practice that has a less discriminatory effect.’” *Id.* at 2515 (quoting 24 C.F.R. § 100.500(c)(3)).

Moreover, at least in the employment context, the Supreme Court has found that “business necessity” is a defense to disparate impact liability. In the big data context, even if a company finds itself on the receiving end of allegations that its big data analytics are resulting in a disparate impact, the company could demonstrate that the impact resulted from legitimate, nondiscriminatory reasons necessary to the functioning of the algorithm. Some useful data algorithms may necessarily cause a disparate impact, and if there is no intent to discriminate, a company applying such algorithms would have strong arguments that it should not be held liable for discrimination under the “business necessity” defense. And following the prophylactic steps described above will further assist in making this defense.

In addition, depending on the statutory regime invoked by a plaintiff or the government, a big data business defendant could have strong arguments that disparate impact liability is impermissible. In *Texas Department of Housing*, the Supreme Court relied on very specific statutory language in expanding the Fair Housing Act to include a disparate impact cause of action. ECOA, for example, has no analogous language. Thus, both the CFPB and Federal Reserve’s interpretation of that statute to provide for disparate impact liability may be improper and subject to challenge under the Administrative Procedure Act. In other words, a defendant may be able to challenge the availability of a disparate

impact cause of action before even reaching the merits of such a claim.

Finally, to the extent a big data business faces a disparate impact class action, there would likely exist good defenses against class certification which, for all practical purposes, would likely be dispositive in any case involving small-value individual claims. For example, in cases involving price or interest rate differentials, the merits and/or damages may be sufficiently individualized to render class certification unachievable. In sev-

eral data cases, courts have refused to certify a class when there are numerous individualized inquiries.

In sum, companies using, selling or managing big data should be aware of growing government interest in digital redlining and disparate impact. Such claims would be novel and face hurdles to success. But, nonetheless, many policy makers and lawyers assessing the utility and risks related to big data innovations are eager to point out downsides. Luckily, companies can take some reasonable steps to avoid and mitigate concerns.