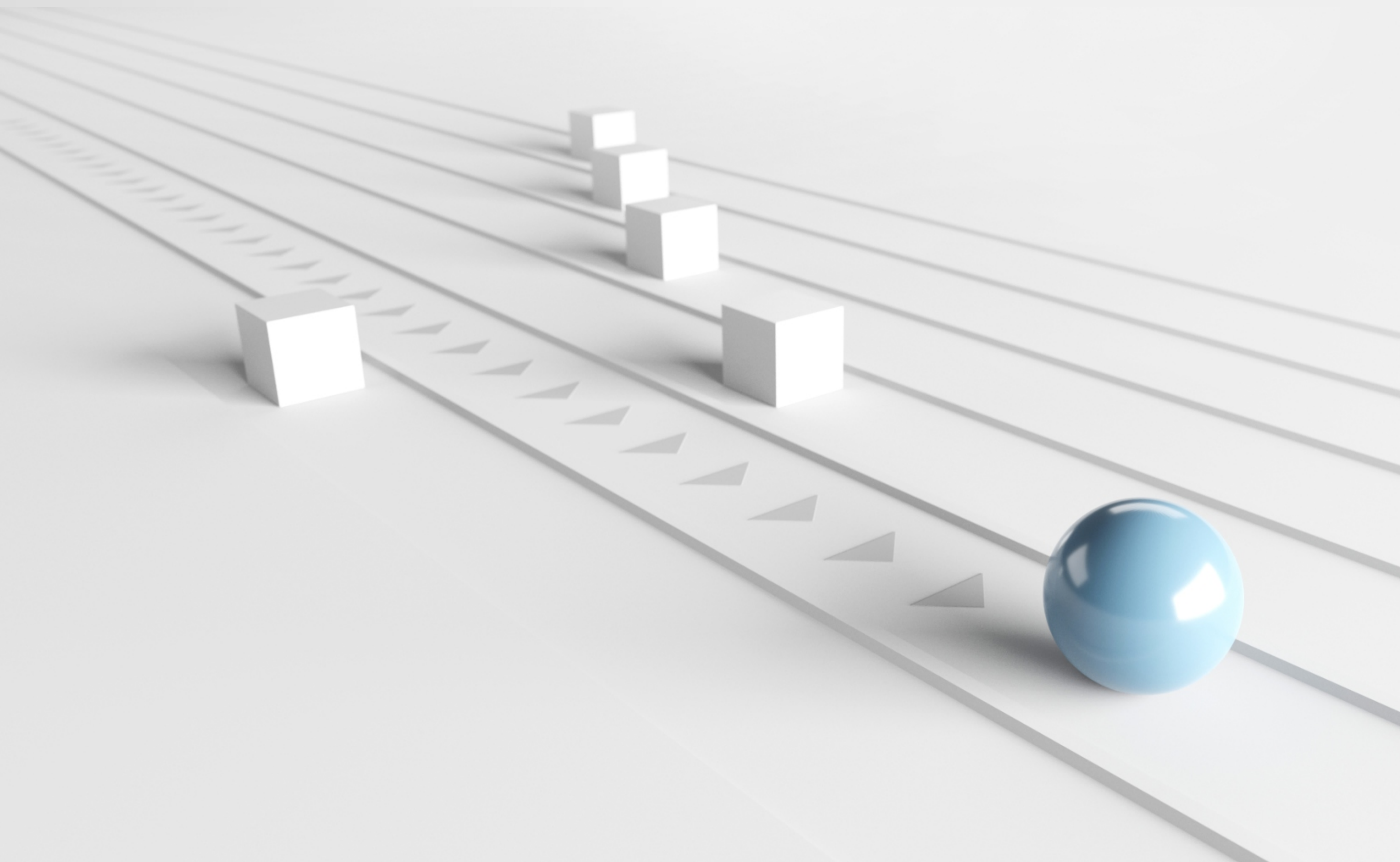
DVARA
Research

INSOLVENCY
LAW ACADEMY
An Indian Institute of Excellence in Insolvency

Assessing the Borrower- Level Impact of the Insolvency and Bankruptcy Code 2016:

A STUDY OF THE FRESH START PROCESS



ABOUT

DVARA RESEARCH FOUNDATION

Dvara Research Foundation (Dvara) is a not-for-profit company policy research institution based in India and incorporated under Section 25 of the Companies Act 1956. The main mission of Dvara is to ensure that every individual and every enterprise has complete access to financial services and occupy a unique position among the policy advocacy institutions in the country for research and policy advocacy activities in financial services and financial inclusion. Since 2008, Dvara has made several contributions to the Indian financial system, specifically in the development of high-quality origination in India through research, and by assessing the impact of various financial products. Dvara has engaged with key policymaking institutions such as the Reserve Bank of India, Securities and Exchange Board of India, Pension Fund Regulatory and Development Authority, etc., through its research contributions.

www.dvara.com

INSOLVENCY LAW ACADEMY

ILA is an independent institute of excellence in insolvency. As a think tank and research institute, ILA seeks to contribute to robust policy making and enhancement of standards in the insolvency industry through cutting-edge research, innovation and development of best practices. ILA aims to develop a community in pursuit of education, research and scholarship in the field of insolvency in India.

www.insolvencylawacademy.com

In more than seven years, the Insolvency and Bankruptcy Code, 2016, remains inoperative for all natural persons, except the guarantors of corporate debt. This essentially leaves most individuals, proprietorships (which constitute the majority of Indian micro, small, & medium enterprises,) and partnerships which are not limited liability entities, at the mercy of colonial legislations, which were products of the then political economy. As one may expect for a country of nearly 1.5 billion, demographic heterogeneity is inevitable. On one hand, there are individuals higher up in the economic pyramid who can endure or often instigate long drawn and arduous litigations to optimize their relief under any insolvency and bankruptcy (hereafter bankruptcy) regime. On the other hand, there are low income households and enterprises who may need additional protections under the bankruptcy regime. Thus, looking at the present status, of the personal insolvency matters, there is huge opportunity for performing extensive research which in turn will help strengthen India's insolvency and bankruptcy framework for natural persons.

Insolvency Law Academy and Dvara Research Foundation have jointly established a Chair for Personal Insolvency (Chair) to serve as a home for various research projects relating to personal insolvency, and micro, small and medium enterprises. The Chair focuses on the opportunity for research to help strengthen India's insolvency and bankruptcy framework for individuals.

We are pleased to present this paper written under the auspices of the Chair. Authored by **Natasha Agnes D'cruze**, **Shree Harini V**, **Dwijaraj Bhattacharya**, and **Indradeep Ghosh** (all affiliated with Dvara Research), the paper was first presented at research conference of Centre for Advanced Financial Research and Learning (CAFRAL), an independent body set up by the Reserve Bank of India, held in Mumbai, India, in December, 2023, and at ILA Annual Conference held in Goa, India, in February, 2024. The paper has been finalized after considering feedback from both the conferences.

Sumant Batra
President
Insolvency Law Academy

Indradeep Ghosh
Executive Director
Dvara Research Foundation

ABSTRACT:

The 2016 Insolvency and Bankruptcy Code (IBC) is a landmark legislation with the potential to impact every borrower. This paper focuses on Part III of the IBC, which deals with natural persons, proprietorships, and personal guarantors for corporate debt. Through the paper, we attempt to estimate the potential consequences of the Fresh Start Process (FSP) defined under this Part. The IBC lays out economic criteria that can qualify (or disqualify) an applicant for FSP. Under FSP, a borrower must be asset-lite, have a low income, and hold minimal outstanding debt to qualify. These thresholds determine the applicability of the process once the IBC is fully notified. Thus, empirical estimates regarding the effects of the provisions on the Indian credit market are crucial to deciphering the impact of the IBC, more specifically, the FSP.

We start by comparing the contemplated processes and outcomes of IBC with other similar legislations, like the Securitisation and Reconstruction of Financial Assets and Enforcement of Security Interest Act (2002), Provincial Insolvency Act (1920), and Presidency Town's Insolvency Acts (1909). We then proceed to estimate how many borrowers are likely to qualify under the FSP. We use the Centre for Monitoring Indian Economy's (CMIE) Consumer Pyramids Household Survey (CPHS) conjoined (using a nearest neighbour model and the Hungarian Algorithm) with the All-India Debts and Investments Survey (AIDIS) for 2019 to estimate how many households qualify under FSP. We perform the analysis for the entire country, except a few states and union territories with relatively sparse population.

Thus, our research is intended as a methodological contribution through which the impact of the IBC across borrower groups can be measured.


DEBT

I. Introduction

The Insolvency and Bankruptcy Code, 2016 (IBC) was introduced in an environment where formal sector lenders, especially banks, struggled with low asset quality. The IBC was intended "to consolidate and amend the laws relating to reorganisation and insolvency resolution of corporate persons, partnership firms and individuals in a time bound manner for maximisation of value of assets of such persons, to promote entrepreneurship, availability of credit and balance the interests of all the stakeholders"[1]. It has been almost 8 years, but not all stakeholders are still covered. Currently, the act is operational (i.e., notified by the government) for corporate debtors and individuals (natural persons) who are guarantors of corporate debtors. Non-limited liability entities like partnerships, proprietorships, etc., are still outside the scope of the remedies proposed by the IBC, since Part III of the IBC, which deals with such debtors, is not notified in its entirety.

Though there is no official declaration regarding why some sections of the code are not notified, it is possible to conjecture that they have to do with the rather complex issue of natural persons. In the case of Part III of the IBC, a human subject in distress becomes a key consideration. Policymakers must, therefore, contend not only with how Part III will impact credit markets but also with the ethical question of whether a natural person deserves relief in some form and, if so, why. The three processes outlined in the IBC provide us a glimpse into the minds of the policymakers, especially highlighting how they envision answering this ethical question.

The three processes under which a natural person (or her creditor) may seek shelter are: a) The Insolvency Resolution Process (IRP) b) The Bankruptcy Process, and c) The Fresh Start Process (FSP). The first two processes form part of a continuum, whereby any debtor (or their creditor) can file for an IRP and apply for bankruptcy if such an IRP fails. The third process, FSP, is unique. It is targeted towards low-income borrowers who are asset-light and have minimal outstanding debt, i.e., the most vulnerable borrowers. For such qualifying[2] individuals, the FSP proposes a scenario where their debts can be wiped clean, i.e., "discharged". In this paper, we situate the FSP in the historical arc of insolvency and bankruptcy regimes and processes, and then present a methodology (and insights therefrom) through which the impact of the FSP can be measured at a borrower level.

[1] Per the long title of the IBC, 2016
[2] Specific qualification criteria are discussed in the next section.

II. Situating the Fresh Start Process in a Historical Context

For as long as credit has existed, there have been borrowers unable to repay their monetary debts, and an attempt to recover the debts has always led to acrimony. In classical antiquity, creditors could repossess the debtor's person, i.e., debt slavery was common, and the practice was rooted in customs rather than formal laws (Levinthal, 1918). Between the 1st and 16th century AD, a second phase of insolvency practices developed; debt slavery received formal legal sanction, but certain sections of the society (members of higher political standing) were granted immunity from such a punishment. With the dawn of enlightenment, rational-legal principles began to take centre stage, and by the mid-16th century, formal law offered some protection to the debtor in default but also empowered the state (more precisely, its embodiment, the crown) to impose the death penalty (Carlos, 2019; Bhattacharya & Ghosh, 2022)

Across these three phases, the purpose of the law (or the custom) was to enable the creditor to reclaim their debt. Further, another common feature unites these three phases - the lenders and borrowers were mostly singular entities[3] and natural persons. However, there were exceptions to this rule, i.e., some institutions did lend and borrow. After the 10th century, institutions like the church and the crown(s) often received or disbursed credit. The terms of such credit were, however, governed by bilateral agreements between the lender and the borrower rather than a codified national-level law. Starting from the turn of the 19th century, the modern era marks a significant departure from the earlier phases. Natural persons no longer occupy centre stage, neither as creditors[4] nor as borrowers[5]. With the invention of "companies", and with such companies receiving the lion's share of credit (earlier for trade and manufacturing, and later for services), they emerge as the key focus for insolvency and bankruptcy regimes (Bhattacharya & Ghosh, 2022).

Axiomatically, we know that corporations are different from natural persons. The former can be carved into pieces and liquidated. The latter, on the other hand, have inalienable rights. Therefore, modern-day insolvency and bankruptcy regimes have attempted to move beyond the express purpose of enabling creditors to reclaim their debt. Now, they aim to balance the rights of the debtors against the creditors. In India, the Presidency Towns Insolvency Act (1909) and the Provincial Insolvency Act (1920) attempted to do this before Part III of the IBC sought to replace them with new provisions. However, those earlier acts remain in force since the majority of Part III of the IBC is yet to be notified[6].

Apart from procedural aspects – such as the identification of the forum having jurisdiction over the subject, the presence of a moratorium, the time-bound nature of the processes, the need for an insolvency resolution professional etc., – the key difference between the British-era statutes and the IBC is the Fresh Start Process (FSP). The FSP is a low-cost quasi-bankruptcy process applicable for low-income, asset-light debtors holding minimal debt. It allows for a complete discharge of their debt provided they satisfy specific economic and procedural criteria. Thus, the FSP process mimics the gate-kept bankruptcy process whereby the debtor may get a complete discharge from their obligations (Bhattacharya & Ananth, 2021).

---

[3] one person would be lending to another rather than a consortium of persons lending to one or more people
[4] replaced by banks, and banking institutions
[5] replaced by corporates
[6] since section (243) of the IBC which repeals the Presidency Towns Insolvency Act (1909) and the Provincial Insolvency Act (1920) has not been notified, these laws remain in-force.

In the present form, an individual (debtor) applying for FSP under the IBC must satisfy four economic criteria, as specified in sections 80(2)(a) - 80(2)(c) and 80(2)(e) of the IBC. These include the income criterion (the debtor must have annual income not exceeding ₹ 60,000), the asset criterion (the aggregate value of the debtor's assets ought not to exceed ₹ 20,000), the debt criterion (the eligible debt owed by the individual must not exceed ₹ 35,000) and an extension of the asset criterion, whereunder for a debtor to be eligible, they must not own a "dwelling unit". Further, the IBC specifies that these criteria should be jointly applied, meaning that a debtor would qualify for the FSP if it satisfies all four (IBC, 2016).

The criteria, however, leave significant scope for interpretation in their definitions. For instance, it is unclear which income streams would be considered income under the income criterion. For an individual operating a proprietorship, all revenues from the business venture are essentially personal income and that aggregate number is very likely to exceed the ceiling, thus making most ineligible for the remedy. Furthermore, it is unclear whether direct benefits transfers by the government will be considered income. If they are, that would even further reduce the eligible debtor numbers. The asset criterion and its extension present several dilemmas also. How do we ascertain the value of household goods? Who should be considered the owner if the asset is a common asset? Regarding the ownership of a dwelling, how should structures that are not wholly residential but used for residential purposes be treated (e.g. a hut on agricultural land used as the residence and storage unit for grains)?

Thus, estimating the impact of the IBC, especially the FSP, using an as-is interpretation of Part III must be accompanied by a set of assumptions that seek to resolve interpretive concerns such as the ones identified in the previous paragraph. The following section discusses these assumptions and the data sources (and their transformations) in detail.

III. Data Sources and Methods

In India, no pan-national official data source simultaneously captures an individual's income, the assets owned by them and their debts. These data reside in fragmented silos. For income, the official data resides within the income tax department. However, with only 7.4 crore people filing income tax returns in 2022-23 and given the widespread informal economy in the country, the data is neither comprehensive nor adequately representative. For data on debt owed by the individual, the hurdles are similar. Credit Information Companies (CICs) capture the cumulative credit outstanding for individuals and businesses, but the data only represents formal credit, thus reducing representativeness and comprehensiveness. Most importantly, however, neither of the above data sources is public. Finally, capturing the asset ownership of an individual through any consolidated database is virtually impossible. So, official data sources are of little help, and reliance must be placed on nationally representative surveys for estimation efforts.

Currently, two such surveys exist – the All-India Debt and Investment Survey (AIDIS), conducted by the National Sample Survey Organisation (NSSO) and the Consumer Pyramids Household Surveys (CPHS), conducted by the Center for Monitoring Indian Economy (CMIE). Both surveys have their limitations.

AIDIS is a sample survey that captures quantitative information on assets and liabilities but not income. Further, most of the relevant data for our analysis is captured at a household level and not at the individual level, which ought to be the unit of analysis given the construct of the FSP. The CPHS, on the other hand, provides complementary details, like the quantum of income and ownership of debt, at an individual level. And also, across most asset segments such as household durables, jewellery, vehicles, etc., the CPHS data only indicates whether a particular asset types is owned or not, and not its value (if owned). Thus, from CPHS, we may only learn that a household has jewellery, but not how much it is worth. The qualification criteria for FSP, however, are based on values.

Thus, neither the CPHS dataset nor the AIDIS dataset can be used in isolation to estimate the number of borrowers the FSP will cover. However, together, both datasets complement each other. The CPHS dataset presents select insights at an individual level and captures income. In contrast, the AIDIS dataset captures granular details on asset ownership and debt owed, though at a household level. Thus, a combined analysis of both datasets is critical, necessitating us to adopt an approach to match households from one dataset to another.

Section 3.1: Matching the Datasets

Matching observations between datasets is a common yet intricate challenge, especially when dealing with sample surveys representing the same universe. This task becomes particularly complex since the AIDIS (for the year 2019) and CPHS (for the year 2019) datasets have a multitude of variables, both categorical and continuous. These variables must be taken into account simultaneously for any accurate matching. This objective can, therefore, be recast as a classification problem. To elucidate, let us consider there are three households, "a_1", "a_2", and "a_3" from the AIDIS dataset and "c_1", "c_2" and "c_3" from the CPHS dataset. Further, let us consider there are three variables common between the two datasets, "V_1", "V_2", and "V_3". The values of each variable for the different households are given below in Tables 1 (A) and 1(B).

Table-1(A): Snippet from AIDIS dataset

Household	V_1	V_2	V_3
a_1	Male	7	15000
a_2	Male	5	30000
a_3	Female	6	45000

Table-1(B): Snippet from CPHS dataset

Household	V_1	V_2	V_3
c_1	Male	7	15000
c_2	Female	12	150000
c_3	Female	6	48000

Datasets like the AIDIS and CPHS often contain variables like the gender of the head of the household, the number of members in the family and the income (at a given frequency). Thus, we can assume V_1, V_2, and V_3 represent these categories. With the presented information, it would appear that households a_1 and c_1 are identical since any and all given variables have identical values. Conversely, households a_2 and c_2 are very different. In the case of a_3 and c_3, however, concluding whether the households are identical (or different) is an arduous task, especially when we consider that data may have been collected at different points in time. Thus, statistical models must be used to systematically calculate similarities between two households using their properties (i.e., variables). One of the most popular methods for solving such classification problems is the *k*-nearest-neighbour (KNN) method (Cover & Hart, 1967).

The KNN method is often used for classification and regression tasks (Fix & Hodges Jr, 1951). Its flexibility and simplicity make it a valuable tool in data-matching exercises. The method operates on the premise that similar instances in the feature (i.e., variable) space tend to share similar labels (Song et al., 2017). In the context of our exercise, it means that households with similar characteristics, such as – the number of members, location, social group, expenditure, etc., are likely to be the same, i.e., they reflect identical characteristics. Thus, the KNN method essentially establishes similarities (Mehta et al., 2018), which can then be inferred to mean that household "a" from AIDIS is identical to household "c" from CPHS.

Before proceeding further, it is important to understand the KNN method's three key aspects. First, how is the distance between the neighbours calculated? Second, how is the value of "K" assigned? Third, how is the assignment decision made (decision rule)?

Table-2: Variables selected for identifying similar households

Variable Description	Variable Type
Region (Urban/Rural)	Categorical
District	Categorical
Social Group	Categorical
Religion	Categorical
Age Groups	Categorical
Gender Groups	Categorical
Household Size Groups	Categorical
Household Expenditure	Continuous
#Similar HHs in the Country	Continuous

Source: Authors' Calculations

On the choice of distance measure, we note first that we are working with two types of variables: categorical ones and continuous ones. A categorical variable can assume a finite number of categories without a natural ordering. For example, the states of India may be coded as numbers, with 1 representing Andhra Pradesh, 2 for Arunachal Pradesh, 28 representing West Bengal, and so on (assignment per alphabetical order). Here, the numbers 1 to 28 have a natural order, where 28 is greater than 27, which in turn is greater than 26, and so on. However, such ordering is meaningless. Just because West Bengal is 28 and Andhra Pradesh is 1, it doesn't mean West Bengal is greater than Andhra Pradesh. Similarly, in our case, the categorical variables discussed in Table-2 do not share a natural order, despite often being coded as numbers.

The second variable type is a quantitative measurement (on the integers or real numbers line). In this case, there is a natural order. Further, the difference between the values are also meaningful. For example, an expense of ₹ 10 is less than one of ₹ 100. Similarly, the difference between ₹ 10 and ₹ 100 is meaningful since we can now learn that one household consumed more goods valued and we can quantify that difference as ₹ 90 in value terms.

Several distance functions are available when dealing with all categorical or non-categorical variables (Abu Alfeilat et al., 2019; Van de Velden et al., 2019). However, options are limited for datasets with mixed-type variables, which is common in survey data.

The continuous variables are normalised first so that the values lie between 0 and 1. The normalisation is achieved by subtracting the minimum value of the variable in the dataset from the value to be normalised and dividing this difference by the difference between the maximum and minimum values of the variable in the dataset. Thereafter, we compute the scalar distance for the normalised variable between the two households (from the AIDIS and CPHS datasets). Thus, we obtain 2 distances, one for each variable. To resolve these 2 distances into a single measure that combines the distance for all (both) continuous variables, we square each scalar difference, then sum the squares and then take the square root (this is a Euclidean metric). This result is divided by 2 to obtain a continuous distance distribution (between 0 and 1).

For categorical variables, the process is more straightforward. For each of the categorical variables, either there will be a perfect match or not. If there is a perfect match, we calculate that distance as zero. If not, then we calculate that distance as 1. We then sum the 7 distances (for the seven categorical variables) to obtain a combined measure of the distance for all categorical variables. This result is divided by 7 to obtain a step-separated[7] categorical distance (between 0 and 1).

Finally, the two distance measures, one for continuous variables and the other for categorical variables, are resolved into a single distance measure using the modified Gower method (Gower, 1971), and this too produces a number between 0 and 1. This concludes the discussion on the first of the three aspects of the KNN method.

The second and third aspects are the value assigned to "K" and the assignment algorithm for the nearest match. We discuss these together as they relate closely to each other. For our estimation, we assign the value of 5 to 'K', meaning that the KNN method will consider the "5" nearest neighbours (based on the collapsed distance as measured through the modified Gower's distance) before assigning which is the closest match (based on the individual distances across all variables). For our analysis, we can consider that the operation is being carried out for household "a" from AIDIS across all households "c_1" to "c_n" from CMIE. Thus, in the first step, the KNN method will select 5 closest neighbours from the CPHS dataset using only one distance measure, the modified Gower distance. Thus, we obtain 5 possible assignments: household 'a' matched to 'c_1' (denoted as $c_1 \rightarrow a$), or $c_2 \rightarrow a$, $c_3 \rightarrow a$, $c_4 \rightarrow a$, and c5 → a. In a scenario where only one pair has the minimum distance between two households, such a pair is considered to be the final match. To exemplify, if the distance between $c_1 \rightarrow a$ is 0.1 and the distances between $c_2 \rightarrow a$, $c_3 \rightarrow a$, etc. are all greater than 0.1, household 'c_1' is assigned to household 'a'. However, if the minimum distance is shared by two (or more) pairs, i.e., the distance between, say, $c_1 \rightarrow a$ and $c_2 \rightarrow a$ are identical and the minimum, then there is a tie. In such a scenario, to resolve the tie, the model computes 9 measures of distance for each pair of households, i.e., for the pair ($c_1 \rightarrow a$), the model computes the distance using the 'region' variable, then the 'district' variable, and so on, across all variables listed in Table-2. So, instead of comparing just one distance measure, the model now compares nine distance measures to find which pair has the maximum number of minimum distances. It is still theoretically possible not to be able to resolve the tie; however, since we did not face the situation, a discussion of the same is avoided. Through this process, the KNN method chooses which of the five households from CMIE is the closest match to household 'a' of AIDIS.

---

[7] The distances are step separated, since it can only assume discrete values of 0/7 (i.e., all the categorical variables match), or 1/7 (i.e., only one categorical variable does not match), and so on.

The KNN method also has a few drawbacks (Guo et al., 2003). Firstly, its computational complexity increases with the size of the dataset (Maillo et al., 2015; Maillo et al., 2017; Deng et al., 2016). Secondly, in high-dimensional spaces where instances tend to be equidistant[8], a challenge arises, impacting the method's performance, known as the curse of dimensionality. Finally, the KNN method is sensitive to imbalanced datasets, potentially leading to biased predictions (Goyal, 2022). In this estimation exercise, the first two drawbacks, computational complexity and distances in higher dimensional spaces, are mitigated by reducing the total observations and dimensions. Observation reduction was done by selecting one state at a time from both datasets, and dimension reduction was done by selecting only 9 common variables across both AIDIS and CPHS datasets.

The third challenge that arises due to imbalanced datasets, resulting in higher and lower density regions, remains. For example, we expect to find more households earning between ₹ 10,000 and ₹ 1,00,000 than between ₹ 10,00,000 and ₹ 10,90,000, despite the interval being equal. Thus, when all variables are considered together, regions of overpopulation (and higher densities) and regions of underpopulation (and lower densities) emerge. This prevents us from achieving a 1:1 (unique) match. To mitigate this hurdle, we also use the "Hungarian Algorithm" to find matching households between the two datasets.

The Hungarian method, developed by Hungarian mathematicians Dénes Kőnig and Jenő Egerváry in the 1930s, has found applications in various fields. It solves the classification problem where the goal is to find the optimal assignment of a set of tasks to a set of agents, minimising the total cost (Hahn et al., 1998). In our context, the goal is to assign households from the CPHS dataset to households in the AIDIS dataset while minimising the total distance.

Operationally, the task is carried out by constructing a table, say 'X'. Each element in the table, 'X_{ac}', represents the distance between household '*a*' from AIDIS and household *'c'* from CPHS datasets. The distance measure used for the Hungarian method is identical to that of the KNN. The lower the distance between the two households, the more similar they are. The Hungarian method then iteratively selects pairs of unique households in a manner such that the sum of all distances (between two matched households) is minimised. We can consider an example to understand this better. Say there are two households, a_1 and a_2 from AIDIS and c_1, c_2, and c_3 from CPHS. Thus, there are six possible assignments: $c_1 \rightarrow a_1$, $c_2 \rightarrow a_1$, $c_3 \rightarrow a_1$, $c_1 \rightarrow a_2$, $c_2 \rightarrow a_2$, $c_3 \rightarrow a_2$. Firstly, the Hungarian method considers the assignment, $c_1 \rightarrow a_1$, as a given (say, with a distance of 0.2). At this stage, both c_1 and a_1 are considered assigned, and thus, the model only computes the distance for $c_2 \rightarrow a_2$ (say, a distance of 0.3) and c_3 a_2 (say, a distance of 0.4), i.e., the residual pairs. Thus, in the first iteration, the optimal match is found to be c_1 a_1 and $c_2 \rightarrow a_2$, with a total distance of 0.5. The model then considers the pair $c_2 \rightarrow a_1$ as fixed and computes the distance for the residual pairs, which, let us say, results in a minimum total distance of 0.4, with $c_2 \rightarrow a_1$ and $c_1 \rightarrow a_2$ representing the matches. Finally, in the third iteration of the model, $c_3 \rightarrow a_1$ will be considered fixed, and the distance of the residual pairs will be computed. Out of these three iterations, let us say the second iteration resulted in the lowest sum of distances. In such a scenario, the resultant pair from the second iteration is considered final.

Thus, combining the KNN and the Hungarian methods provides a comprehensive and effective approach to household matching. The former's

[8] It can be intuitively understood in the following example: Assume we compare two countries based on one parameter, say "GDP". Then, we are likely to find a difference. As we start adding dimensions, say population, growth rates, gender distribution, life expectancy, majority religion, etc., in some cases, the distances will start increasing (e.g., if we were comparing India and Bangladesh), while in others the distances will start reducing (e.g., if we were comparing Iran and Turkey, which have similar population, life expectancy, and so on). So, as the number of variables (dimensions) increase, the chances that two countries may appear similar increases, especially when we are adding the difference in the variables.

flexibility in handling mixed variable types and adaptability to complex distributions, combined with the latter's precision in achieving an optimal one-to-one mapping, creates a synergistic effect that addresses the discussed challenges in the matching process.

To generate unique one-to-one mapping, we must match from the smaller dataset to the bigger one – meaning that for the states where AIDIS has the smaller number of households, we will try to find for each AIDIS household a corresponding and unique household from the CMIE dataset that is its closest match. Thus, to combine both models, we start with KNN. Assuming that AIDIS has fewer households for all states compared to the CPHS, the KNN model shall result in some one-to-one matching (one household from the CPHS dataset will be assigned to one from AIDIS), some one-to-many matching (one household from CPHS will be assigned to many households of AIDIS), as well as some residual households (of CPHS who were not assigned to any households in AIDIS).

These unique (one-to-one) matches are considered final matches. For the one-to-many matches, we consider the closest match as the final match. To exemplify, say, household c1 of CPHS was matched with households a_1, a_2, and a_3 of AIDIS. The distance between each pair c_1-a_1, c_1-a_2, and c_1-a_3 are 0.2, 0.25 and 0.35, respectively. So, despite three matches, we only consider the c1-a1 pair since this has the lowest distance. We obtain a set of matched and unmatched households using these one-to-one matches and by resolving the one-to-many matches. These matched households are used for final analysis, whereas the unmatched households are then passed onto the Hungarian method for final matching.

Section 3.2: Data Transformations

Data cleaning is a crucial step in the pre-processing pipeline, especially when dealing with datasets that include both categorical and continuous variables. Following are the key strategies adopted for data cleaning before employing the K-Nearest Neighbors (KNN) and the Hungarian method.

- Handling Missing Values: KNN and the Hungarian methods are sensitive to missing data. Given the negligible occurrence of such missing data across the variables used for matching and estimating the impact of the FSP, imputation methods are avoided since they may introduce bias or distort the original distribution. Instead, such households were dropped.
- Standardising and Scaling: KNN relies on distance metrics, and the Hungarian method involves optimisation, both of which are influenced by the scale of variables. Thus, observations were standardised by subtracting the minimum value and dividing by the range (maximum observed value – minimum observed value of the variable).
- Recasting Categorical Variables: Categorical variables, wherever in the form of non-numeric values, were converted into a numerical format.
- Ensuring Compatibility with Methods: Finally, since the two methods have specific requirements regarding the input data format, the datasets were reorganised and variables were appropriately pre-processed to ensure compatibility.

Upon completion of the data transformation, the KNN and the Hungarian methods were used to obtain the final data structure based on which estimations were carried out.

Before discussing the final data structure, it is important to discuss one final aspect of the matching procedure: the quantum of data loss. It is evident that whether the matching happens from AIDIS to CPHS or from CPHS to AIDIS, the final results will not differ since the final result will indicate that households "a" and "c" (from AIDIS and CPHS, respectively) are identical. However, the number of households in each state may differ. For example, in Bihar, AIDIS has 7708 households and CPHS has 9236 households, and thus, 1528 households[9] from CPHS do not get any households from AIDIS assigned to them. The data pertaining to these (1528 in case of Bihar) residual households are thus not accounted for in the final dataset. Appendix-A presents the number of households that were residual households for each of the analysed states.

Section 3.3: Final Data Structure

The final dataset contains all the variables used for merging, along with additional variables from the AIDIS and CMIE datasets. Table-3 presents the description of the variables and their source data[10]:

Table-3: Variables present in the final data (used for estimations)

Sl	Variable Name	From AIDIS	From CMIE
1	Region (Urban/Rural)	Yes	Yes
2	District	Yes	Yes
3	Social Group	Yes	Yes
4	Religion	Yes	Yes
5	Age Groups	Yes	Yes
6	Gender Groups	Yes	Yes
7	Household (HH) Size	Yes	Yes
8	HH Expenditure	Yes	Yes
9	#Similar HHs in the State	Yes	Yes
10	Value of assets owned by the HH (across various types of assets)	Yes	-
11	Amount of Debt Outstanding	Yes	-
12	Occupational Sector of the Head of the HH	-	Yes
13	Household Income	-	Yes

Source: Authors' Calculations

Estimations were done using these variables for the households across Indian states and union territories. The analysis however excludes Andaman & Nicobar Islands, Arunachal Pradesh, Dadra & Nagar Haveli, Daman & Diu, Lakshadweep, Manipur, Mizoram, and Nagaland as the CMIE CPHS does not report data for those states in 2019.

In addition to sample-level estimations, we also use the weights provided by the two datasets to project the estimations onto the population level. For states where the base dataset is AIDIS, i.e., where all households of AIDIS are assigned a corresponding household from the CPHS dataset, we use the weights in the AIDIS dataset to compute state-population-level results. Similarly, for states where the CPHS dataset is used as a base dataset, CPHS weights are used. For most states, we rely on the AIDIS dataset as the base dataset due to its smaller state-specific sample size.

[9] 9236 (households in CPHS) - 7708 (households in AIDIS) =1528 Households from CPHS who were not assigned a corresponding household from the AIDIS dataset.

[10] The total number of variables used for the estimation is 156, but between them they contain the data pertaining to the themes discussed in the table. All 156 variables are not reproduced here to enhance ease of understanding.

In case of AIDIS, weights are assigned at the stratum or district level. To compute the total number of FSP-eligible households in the population, we identify qualifying households in the sample, multiply their eligibility by the assigned weight, and sum up these values for a population-level estimate (National Sample Survey Organisation, 2019). However, for Assam, Delhi, Meghalaya, Sikkim, and Tripura, we turn to the CMIE CPHS as the base dataset. When using CMIE CPHS as the base, we apply the dataset's provided weights, utilising state-level weights for households and a non-response factor. The weight of an observation is calculated by scaling the state-level weight with the non-response factor, yielding a measure for each household per month. These constructed weights are averaged to derive a final measure for each household in the year 2019, which is then employed for all population-level estimates (Consumer Pyramids Household Survey, 2019). The estimation results are discussed in the next section.

IV. Estimation Results

Households were matched using a tiered approach. The first layer of matching was done using KNN, and the second layer using the Hungarian model. The following table, Table-4, presents the total number of households (of AIDIS) matched in each stage and their mean distances.

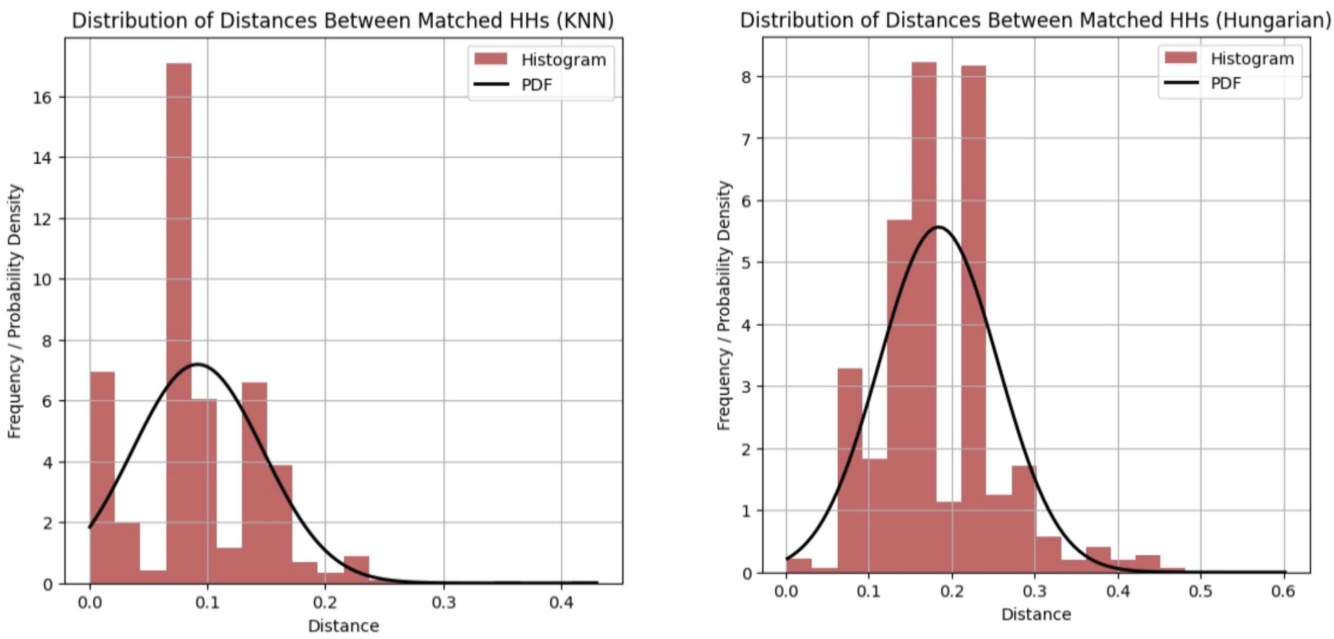
Table-4: Households matched through each model (and summary statistics of the distances)

	# HH (from AIDIS/CMIE) Matched	Mean Distance (Modified Gower)	Std. Dev.	Median Distance (Modified Gower)
KNN Method	37323	0.09	0.055	0.08
Hungarian Method	69093	0.18	0.071	0.16

Source: Authors' Calculations

Figures-1(A) and -1(B) present the distribution of distances of the matched households across the two methods.

Figure-1: Distribution of distances between matched households using KNN (A) and Hungarian Method (B)



Source: Authors' Calculations

As discussed in the earlier section, any pair of matched households will have two distances—one combined distance for categorical variables and one combined distance for continuous variables. Given that we summed the distance of all categorical variables and then divided it by 7, we obtained a stepwise distribution for categorical variables (between 0 and 1). Similarly, we obtain a continuous distribution (between 0 and 1) for continuous variables. Thus, Figures 1(A) and 1(B) suggest that most of the matched households were fairly close to one another.

Using the merged data, we estimate the eligibility of the households for FSP. Table-5 presents the summary statistics of the relevant variables (at the sample level).

Table-5: Summary statistics of the relevant variables for determining eligibility under the FSP

Variable	Count	Mean	Standard Deviation	1st quartile (Q)	2nd Q (Median)	3rd Q
Total Annual Income	1,06,416	2,41,405.7	2,09,040.5	1,21,081.5	1,81,735	2,92,447
Outstanding Debt	50,058	3,19,361.3	10,36,296	34,570	87,000	2,67,659
Value of Assets	1,06,416	23,67,650	85,67,444	2,72,175	8,95,000	23,76,150
Home Ownership[11]	1,06,416	0.83	NA	NA	NA	NA

Source: Authors' Calculations

While the summary statistics presented above are for the sample, the estimation results have been calculated for the population level by applying appropriate weights, as described in the previous section. Table-6 presents how many households qualify under each of the four criteria laid out for the FSP.

Table-6: Number of households qualifying for FSP under each of the eligibility criteria

	Qualifying Households (from Matched Dataset)[12]
FSP Criterion-1: Annual Income < ₹ 60,000	45,02,187
FSP Criterion-2: Outstanding debt amount < ₹ 35,000, but > ₹ 0	2,17,58,764
FSP Criterion-3: Value of Assets < ₹ 20,000	40,24,937
FSP Criterion-4: No home ownership	92,89,643
Combining all criteria	1,50,408

Source: Authors' Calculations

[11] Home ownership is a categorical value. The mean is represented since it presents the ratio of number of people who own a residential property (from the data it appears that 94% of the sample owns a residential property).

[12] The following results have been calculated only for households that have reported owing some debt.

Combining all four criteria, we find that only 1,50,408 out of the 26,56,71,317 households with outstanding debt qualify for FSP. This represents 0.057% of all households. The number of qualifying households across each state is represented in Appendix C. Table-7 presents the number and proportions of qualifying households at the state level.

Table-7: Share of households qualifying for FSP under the income criteria

State	Count	Qualifying HHs (Expenditure criterion)		Base weights used
		Total	%	
Andhra Pradesh	14198806	22525.75	0.1586	CMIE CPHS
Assam	12200471	1115	0.0091	AIDIS
Bihar	17748050	554.83	0.0031	CMIE CPHS
Chandigarh	252275.0938	0	0.0000	CMIE CPHS
Chhattisgarh	5672758.5	0	0.0000	CMIE CPHS
Delhi	4922844	0	0.0000	AIDIS
Goa	308249.5	0	0.0000	CMIE CPHS
Gujarat	12531386	12.5	0.0001	CMIE CPHS
Haryana	5414255.5	0	0.0000	CMIE CPHS
Himachal Pradesh	1716132.75	0	0.0000	CMIE CPHS
Jammu & Kashmir	2272021.25	0	0.0000	CMIE CPHS
Jharkhand	6516384.5	198.5	0.0030	CMIE CPHS
Karnataka	13810240	304.5	0.0022	CMIE CPHS
Kerala	8910524	2663.25	0.0299	CMIE CPHS
Madhya Pradesh	14949053	6806	0.0455	CMIE CPHS
Maharashtra	24223068	3692.5	0.0152	CMIE CPHS
Meghalaya	770592	546	0.0709	AIDIS
Odisha	10015405	63726.63	0.6363	CMIE CPHS
Puducherry	288658	0	0.0000	CMIE CPHS
Punjab	6019335	515.25	0.0086	CMIE CPHS
Rajasthan	13273183	0	0.0000	CMIE CPHS
Sikkim	555504	2175	0.3915	AIDIS
Tamil Nadu	19161852	27230.17	0.1421	CMIE CPHS
Telangana	9276134	0	0.0000	CMIE CPHS
Tripura	1261376	0	0.0000	AIDIS
Uttar Pradesh	35141980	10009.25	0.0285	CMIE CPHS
Uttarakhand	1947767	0	0.0000	CMIE CPHS
West Bengal	22313012	8332.58	0.0373	CMIE CPHS
Total	**265671317.1**	**150407.71**	**0.0566**	

Source: Authors' Calculations

The estimates reveal that Odisha (with 63,727 households), Tamil Nadu (27,230 households), and Andhra Pradesh (22,526 households) are the states with the highest number of households that qualify for FSP. Together, these states account for 75% of the total number of qualifying households per the income criterion. These states also constitute 77% of the total outstanding debt that qualifies for FSP. Further, there are twelve states without any qualifying households. The estimation results thus suggest that there are pockets of concentration where FSP may have a higher uptake, assuming the ratio of qualifying households vis-à-vis households that seek refuge remains constant across regions, states, and cultures.

We also explore an alternative estimation approach. Earlier, we used four criteria (given in Table-8). However, if we replace criterion-1, i.e., "the income of the household must be less than ₹ 60,000 annually", with "expenditure of the household must be less than ₹ 60,000 annually", we find that the number of households that qualify for FSP increases from ₹ 1,50,408 to 4,42,802. We construct this scenario (by replacing income with expenditure) since most measures of poverty focus on the expenditure of the individual or household rather than income. Table-8 provides the number of households that qualify for this revised criteria.

Table-8: Number of households qualifying for FSP under the revised criteria (expenditure-based)

	Qualifying Households (from Matched Dataset)[13]
FSP Criterion-1: Annual Income < ₹ 60,000	1,04,05,050
FSP Criterion-2: Outstanding debt amount < ₹ 35,000, but > ₹ 0	2,17,58,764
FSP Criterion-3: Value of Assets < ₹ 20,000	40,24,937
FSP Criterion-4: No home ownership	92,89,643
Combining all criteria	4,42,802

Source: Authors' Calculations

Combining the revised criteria (replacing income with expenditure), we find that only 4,42,802 households out of the 26,56,71,317 households with outstanding debt qualify for FSP, i.e., only 0.166% of households qualify for FSP. Table-9 presents the state-level qualifications.

Under the revised criteria, Odisha still has 1,03,537 qualifying households, which is the highest in the country. It is followed by West Bengal with 96,159 and Uttar Pradesh with 54,641 qualifying households. These three states together account for 57% of the total number of qualifying households and 56% of the total qualifying outstanding debt, considering the expenditure criterion (alongside asset, debt and home ownership criteria). In this scenario, the number of states with zero qualifying households comes down to six. The number of qualifying households across each state is represented in Appendix D. Table-9 presents the number and proportions of qualifying households at the state level.

[13] The following results have been calculated only for households that have reported owing some debt.

Table-9: Share of households qualifying for FSP under the expenditure criteria

State	Count	Qualifying HHs (Expenditure criterion)		Base weights used
		Total	%	
Andhra Pradesh	14198806	43708	0.3078	CMIE CPHS
Assam	12200471	44566	0.3653	AIDIS
Bihar	17748050	11563	0.0651	CMIE CPHS
Chandigarh	252275	418	0.1657	CMIE CPHS
Chhattisgarh	5672759	69	0.0012	CMIE CPHS
Delhi	4922844	0	0.0000	AIDIS
Goa	308250	137	0.0444	CMIE CPHS
Gujarat	12531386	1755	0.0140	CMIE CPHS
Haryana	5414256	2266	0.0418	CMIE CPHS
Himachal Pradesh	1716133	0	0.0000	CMIE CPHS
Jammu & Kashmir	2272021	0	0.0000	CMIE CPHS
Jharkhand	6516385	2659	0.0408	CMIE CPHS
Karnataka	13810240	8314	0.0602	CMIE CPHS
Kerala	8910524	1013	0.0114	CMIE CPHS
Madhya Pradesh	14949053	1816	0.0121	CMIE CPHS
Maharashtra	24223068	175167	0.0723	CMIE CPHS
Meghalaya	770592	0	0.0000	AIDIS
Odisha	10015405	103537	1.0338	CMIE CPHS
Puducherry	288658	0	0.0000	CMIE CPHS
Punjab	6019335	5946	0.0988	CMIE CPHS
Rajasthan	13273183	426	0.0032	CMIE CPHS
Sikkim	555504	1262	0.2272	AIDIS
Tamil Nadu	19161852	34779	0.1815	CMIE CPHS
Telangana	9276134	7051	0.0760	CMIE CPHS
Tripura	1261376	0	0.0000	AIDIS
Uttar Pradesh	35141980	54641	0.1555	CMIE CPHS
Uttarakhand	1947767	3201	0.1643	CMIE CPHS
West Bengal	22313012	96159	0.4310	CMIE CPHS
Total	**265671317**	**442802**	**0.1667**	

Source: Authors' Calculations

V. Discussion

The estimates reveal that only 1.5 lakh out of the 26 crore households with debt qualify under all four criteria laid down by the IBC. The number of qualifying households increases to 4.4 lakhs if we replace the income criterion with a similar criterion for expenditure. Thus, as discussed earlier, only 0.05% of the households qualify for the Fresh Start Process in India. If the expenditure criterion were to be considered, then that would allow 0.16% of the total households to qualify for FSP.

This three-fold increase in the proportion of qualifying households (when considering the expenditure criterion also implies that the debt that must be written off increases from ₹ 264 crores to ₹ 705 crores. The ₹ 705 crores may appear to be a substantial amount in isolation; however, it amounts to a mere 0.86% of the credit outstanding of micro-finance institutions, specifically entities licensed as Non-Banking Financial Companies (NBFC-MFIs) (MFIN, 2023). If we consider the banking sector (RBI, 2023), it amounts to only 0.05% of total unsecured personal loans.

Under the existing criteria (contained in the IBC), Odisha has the highest number of qualifying households, 63,727. This forms a minuscule fraction of the state's total population of households, only 0.63%. For the states with the second and third highest number of qualifying households, Tamil Nadu (27,230) and Andhra Pradesh (22,526), the qualifying households only represent 0.14% and 0.15% of the total populations, respectively. The state with the highest proportion of qualifying households is Sikkim (2,175 households), amounting to 0.39% of the total population. Thus, in a scenario where all qualifying households seek refuge under the FSP (per the current criteria), there are no states where even 1% of the households will be covered under the IBC. Appendices C and E, present the number of qualifying households and the quantum of qualifying debt (that has to be written off in case all qualifying households seek refuge under the FSP) at a state level, respectively.

In the case of the income criteria, the maximum qualifying debt belongs to Odisha, followed by Andhra Pradesh and Tamil Nadu. A closer study of the number of qualifying households and the outstanding debt suggests that the average outstanding debt per household varies significantly across states (See Appendix B for further details). For example, Andhra Pradesh and Tamil Nadu have similar numbers of households qualifying for FSP (22,526 and 27,230 households, respectively). However, the amount of qualifying debt differs significantly. For Andhra Pradesh, the qualifying debt is ₹ 63.7 crores and for Tamil Nadu (despite having more qualifying households), the qualifying debt is ₹ 39.9 crores. Thus, the average qualifying debt per household for Andhra Pradesh (₹ 28.2 thousand) emerges to be almost twice that of Tamil Nadu (₹ 14.6 thousand). We posit that two factors contribute to this disparity. First, some states have a higher degree of credit penetration in the low-income segments. Second, some states have a higher degree of overindebtedness. Thus, it is not necessary that as the number of qualifying households increases, the qualifying debt must increase in similar proportions.

The scenario remains similar, even when considering the constructed expenditure criteria (alongside the asset, debt and home ownership criteria). Using the expenditure criteria, Odisha still has the highest number of qualifying households. According to this criteria, the share of households that qualify for FSP from the state increases to 1.03% of the total population. The constructed criteria lead to West Bengal having the second-highest number of eligible households, constituting 0.43% of its total population. Uttar Pradesh follows as the state with the third-highest number of eligible households, comprising 0.15% of its total population qualifying for the FSP. Appendices D and F present the number of qualifying households and qualifying debts for each state, respectively.

Under the constructed expenditure criteria, we again observe that states like Andhra Pradesh and Tamil

Nadu have a similar number of qualifying households, but a sizeable difference in the qualifying debt. West Bengal and Odisha also follow a similar trend. States with lesser populations are also not immune to the differentiated average household debt, as is evidenced by Himachal Pradesh and Uttarakhand.

While the estimates provide an insight into the total number of households that may qualify under different scenarios and the debt that correspondingly must be written off, there are a few limitations of the study that we must acknowledge. The results are at a household level and not at an individual level. The IBC defines FSP as a process whereunder an individual may seek refuge. However, one key factor inhibits estimations at the individual level. We can leverage the matched dataset to obtain individual-level income and outstanding debt, but we cannot obtain asset ownership (including home ownership) details. This hurdle arises since there is no singular approach through which assets may be apportioned between the household members. Let us consider a household (of four members) that owns a few utensils, a gas stove, a refrigerator, and a bicycle.

The head of the household is a 55-year-old male who works as a casual labourer. His wife, aged 50, works seasonally during harvest, and two adult children are in college. Assuming they don't own a house, how can we decide who owns the household assets? One approach is to consider who are the beneficiaries. In this example, all the members are. Alternatively, we may inquire who purchased the items initially. This approach may work when cases are adjudicated one at a time, but it is hardly implementable during a sample survey. Another is apportioning the assets according to the current or historical income patterns. In this example, it would mean that we take the income ratio of the head of the household and his wife and then apportion the assets in that ratio. It may work if we can trace this ratio for a long enough period, but it does not account for disproportionate gifts. We can consider many alternate approaches, but none truly capture the nuances. Thus, the limitation around apportioning assets is unlikely to be mitigated ex-ante, i.e., before the section is notified and there is enough jurisprudence to guide estimations.

It is also important to recognise that we have only estimated the number of households that will qualify under FSP and not the number of households likely to seek refuge under it. It is well documented that households tend to make sacrifices, ranging from skipping festivals to skipping meals and pulling children out of school before they turn delinquent. The stigma and shame associated with being delinquent will likely magnify when they try to seek a formal discharge from their debts. Shaping this belief system is the underlying culture that the individual subscribes to. Thus, even if the number of eligible individuals increases, it does not mean everyone will start seeking refuge under the FSP.

Further, the Bankruptcy process, defined under the IBC, presents an interesting alternative to the FSP. In both FSP and bankruptcy, the final outcome is that the debtor is discharged from their repayment obligation, though the process of achieving this differs. In the case of the bankruptcy process, the debtor’s assets are attached to an estate administered by a bankruptcy trustee. The trustee is responsible for selling such assets to recover the dues from the borrower and repay her creditors. Thus, in case of bankruptcy, an alienation of assets is posited to occur. However, such alienation of assets is not absolute, as specific assets are excluded from being attached to the bankruptcy estate. Section 79(14) of the IBC lists these excluded assets; however, it does not assign any value to most.

The first two sub-sections read as under:

> …unencumbered tools, books, vehicles and other equipment as are necessary to the debtor or bankrupt for his personal use or for the purpose of his employment, business or vocation and unencumbered furniture, household equipment and provisions as are necessary for satisfying the basic domestic needs of the bankrupt and his immediate family.

Thus, all assets essential for the debtor's vocation will likely be protected, irrespective of their value. The third exclusion on personal ornaments allows the competent authority to set a value beyond which assets will not be excluded. The sub-section reads 'any unencumbered personal ornaments of such value, as may be prescribed, of the debtor or his immediate family which cannot be parted with, in accordance with religious usage'. Under the Insolvency and Bankruptcy (Application to Adjudicating Authority for Bankruptcy Process for Personal Guarantors to Corporate Debtors) Rules, 2019, personal jewellery worth up to ₹ 1,00,000 is excluded. Similarly, unencumbered single dwelling units having a value of up to ₹ 20,00,000 in urban areas are excluded. In rural areas, single dwelling units having a value of up to ₹ 10,00,000 are excluded. Thus, the quantum of asset protection under the bankruptcy process is significantly greater than that of the FSP. This suggests that households that do not meet the current thresholds present in the FSP may file for the insolvency resolution process (IRP) and later bankruptcy and reap similar benefits as prescribed under the FSP. Thus, the incentive structures poise such households to not act in good faith during the IRP so that they can reap similar benefits (as the FSP) during the bankruptcy process.

This possibility that some may reap the benefits of FSP, despite not qualifying for it, begs the question: are the current FSP thresholds appropriate? To answer, we must decipher the motive of the parliament when enacting the law. Though there is no stated motive for the FSP, a closer reading of the code may provide some insights. Since the fundamental objective of the code is to balance the rights of creditors and debtors, higher thresholds for FSP may erode significant creditor value (which in turn can impact the credit market, but such a discussion is beyond the scope of this paper). Conversely, the low thresholds may indicate an intent to cover the poorest of the poor. We can discuss this in two contexts: minimum wages and minimum per capita consumption expenditure. In India, Nagaland has the lowest minimum wage. At ₹ 5280 per month, it corresponds to ₹ 62,760 annually (Dezan Shira & Associates, 2023). In Delhi, the minimum wage is ₹ 17,494 (monthly) or approximately ₹ 2,10,000 annually (The Mint, 2023). Setting the FSP thresholds lower than the minimum wage suggests an intent to protect the most vulnerable. The question, however, remains. Does the code adequately protect all that needs protection?

One approach that the parliament may consider is replacing the income criterion with an expenditure-based criterion. If there is only one earning member in a household of four, the member must consume items worth ₹ 74,463 annually to ensure that the household stays above India's poverty line (Bhattacharya & Ananth, 2021). In our alternate estimate, we assumed household expenditure thresholds to be ₹ 60,000 and found that only 17 out of 7708 households in the sample qualify for FSP.


29.79
22.10
18.07
25.01
30.12
07.28

VI. Conclusion

The inclusion of FSP suggests that the framers of the IBC envisioned it to embody the evolving moral standard of insolvency regimes. The code makes a visible effort to distinguish and protect natural persons. Despite this intent, the fact that majority of the Part III of the Code is still not notified underscores that resolving the tussle between moral hazard and debtor protection is an arduous task. The methodology suggested in this paper allows policymakers to estimate the impact of the FSP, aiding in the process of resolving the tussle.

Appendix A: State-wise number of residual households

State	AIDIS sample size	CMIE CPHS sample size	Residual	Residual dataset
Andhra Pradesh	4710	8080	3370	CMIE CPHS
Assam	3577	1755	1822	AIDIS
Bihar	7708	9382	1674	CMIE CPHS
Chandigarh	190	456	266	CMIE CPHS
Chhattisgarh	2281	4799	2518	CMIE CPHS
Delhi	1650	1375	275	AIDIS
Goa	235	1064	829	CMIE CPHS
Gujarat	5095	9066	3971	CMIE CPHS
Haryana	2181	5538	3357	CMIE CPHS
Himachal Pradesh	1054	1280	226	CMIE CPHS
Jammu & Kashmir	1603	2588	985	CMIE CPHS
Jharkhand	2830	4710	1880	CMIE CPHS
Karnataka	5750	9717	3967	CMIE CPHS
Kerala	3610	4786	1176	CMIE CPHS
Madhya Pradesh	6164	9200	3036	CMIE CPHS
Maharashtra	10181	19834	9653	CMIE CPHS
Meghalaya	1368	1040	328	AIDIS
Odisha	4080	6761	2681	CMIE CPHS
Puducherry	359	1140	781	CMIE CPHS
Punjab	2691	6760	4069	CMIE CPHS
Rajasthan	5978	10886	4908	CMIE CPHS
Sikkim	858	816	42	AIDIS
Tamil Nadu	7075	10938	3863	CMIE CPHS
Telangana	2999	5830	2831	CMIE CPHS
Tripura	2304	1192	1112	AIDIS
Uttar Pradesh	13769	22868	9099	CMIE CPHS
Uttarakhand	1136	2042	906	CMIE CPHS
West Bengal	8559	10502	1943	CMIE CPHS
Total	**109995**	**174405**		

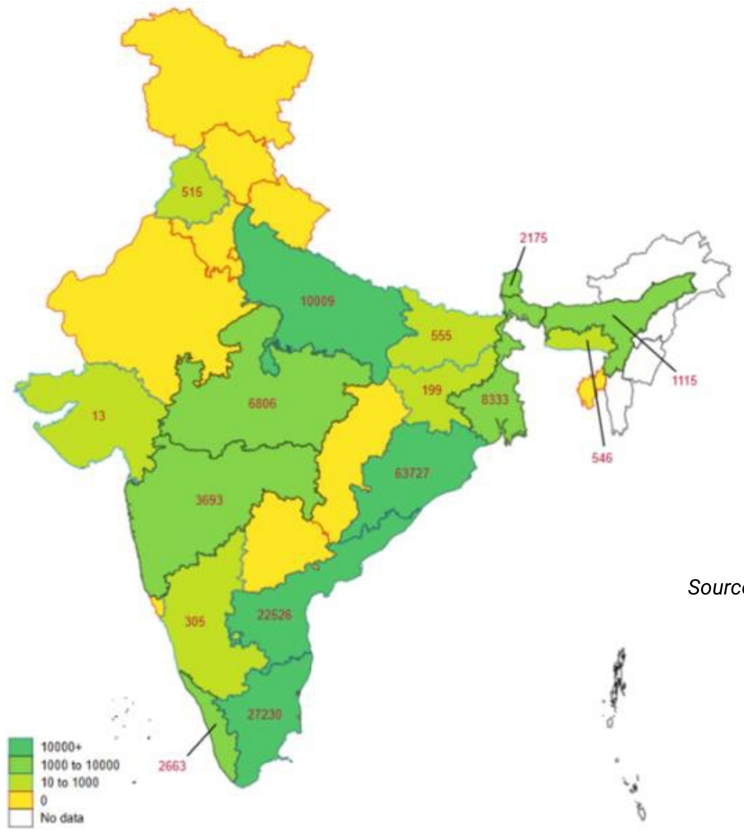
Appendix B: Detailed table (non-rounded) on the share of qualifying households under FSP (income and expenditure criteria)

Note: The figures presented in the table above are for the population-level. It lays out the number of households that qualify for FSP for every criterion, namely, home ownership, total income, asset value, and outstanding debt. We also do a similar calculation for total expenditure. We then calculate the final number of households that would qualify if all the criteria were to be applied. Although the calculations under outstanding debt and final income and expenditure criteria only account for households that owe some debt, for the other calculations we present the figures for households that currently do not owe any debt as there is a chance that they may become indebted in the future

State	Count	Home Ownership		Total Expenditure		Total Income		Asset Value		Debt	Qualifying HHs (Income criteria)			Qualifying HHs (Expenditure criteria)		
		Debt=0	Debt>0	Debt=0	Debt>0	Debt=0	Debt>0	Debt=0	Debt>0		Total	%	Outstanding Debt	Total	%	Outstanding Debt
Andhra Pradesh	14198806	635071.1	2055244	300570.2	933535.9	123232.1	418175.3	303645	10385.40	1121078	22525.75	0.15865	637073344	43708.01	0.307829	637009920
Assam	12200471	218533	211512	324608	378199	754423	708745	80608	67845	1630800	1115	0.00914	27643080	44566	0.365281	716585920
Bihar	17748050	62322.27	178552.3	191464.4	628172.3	130923.6	303135.4	22422.55	51908.9	2250797	554.83	0.00313	8152780	11562.83	0.06515	109707776
Chandigarh	252275.0938	2957.35	8092.56	0	418	0	0	1179.1	447.33	4302.18	0	0.00000	0	418	0.165692	6270000
Chhattisgarh	5672758.5	38101.92	53087.14	188878.2	249048.1	54171.54	67487.82	11377.5	32097.5	475626.5	0	0.00000	0	69	0.001216	759000
Delhi	4922844	279297	218586	0	0	0	0	1523.55	91644	121473	0	0.00000	0	0	0	0
Goa	308249.5	2481.16	3191.6	171.5	137	0	978.5	921.33	960	3642	0	0.00000	0	137	0.044445	2740000
Gujarat	12531386	106929.1	160018.2	39374.67	101095.8	20334.68	54626.35	53752.35	36691.82	516420.1	12.5	0.00010	125000	1754.5	0.014001	20250950
Haryana	5414255.5	52494.37	130411.2	5066.22	17507.8	0	489	28060.26	34045.22	242019.9	0	0.00000	0	2265.5	0.041843	21726300
Himachal Pradesh	1716132.75	9061.75	36834.5	1286.83	28279.85	0	941	6870.5	5089.23	62388.11	0	0.00000	0	0	0	0
Jammu & Kashmir	2272021.25	1644.33	6278.71	1212.5	32003.25	2713.5	41685.15	475	934.75	96731.64	0	0.00000	0	0	0	0
Jharkhand	6516384.5	46793.54	72585.01	158298.4	252192.7	48890.92	49006.32	31440.52	18491.62	723043.1	198.5	0.00305	459700	2658.5	0.040797	59063076
Karnataka	13810240	130869.6	620216.6	142331.1	615740.6	25612.3	107278	51237.65	203080.7	873674.8	304.5	0.00220	3093750	8314.06	0.060202	139030528
Kerala	8910524	193367.7	581543.7	47264.92	111829.3	36633.58	117894	82075.43	191504.9	528871.1	2663.25	0.02989	49662676	1013.25	0.011371	29803676
Madhya Pradesh	14949053	96381.99	279502	213969.3	1095408	12866.93	95783.74	39889.98	97341.1	1298148	6806	0.04553	162286512	1816.14	0.012149	41296864
Maharashtra	24223068	284476.4	658418.5	256065.5	878256.1	93447.33	331163.8	145866.5	222628.5	1251901	3692.5	0.01524	57615000	17516.92	0.072315	443746880
Meghalaya	770592	8505	43800	546	3048	1092	2472	6324	17244	108090	546	0.07085	7985250	0	0	0
Odisha	10015405	217898.5	500311.8	637429	1372801	265615.1	740722.8	113689.6	257125.5	1900215	63726.63	0.63629	992387584	103536.9	1.033776	1688397952
Puducherry	288658	7653.21	2709.67	324	2672.88	0	195.5	0	7060.29	20554.13	0	0.00000	0	0	0	0
Punjab	6019335	67820.64	162778.6	29740.5	9158.84	3225.75	4726.46	25173.71	83828.5	415468.9	515.25	0.00856	8914340	5946	0.098782	72510000
Rajasthan	13273183	159006.7	213954.1	126264.2	201695.7	34057.81	89070.67	42575.74	92092.61	954387.3	0	0.00000	0	426	0.003209	10692000
Sikkim	555504	26312	32326	11707	10311	58136	101262	20513	15518	50508	2175	0.39154	47166752	1262	0.227181	24341750
Tamil Nadu	19161852	675535.1	1272215	179500.7	549265.9	87498.69	272428.7	272762.7	503439.9	1243834	27230.17	0.14211	398806944	34779.09	0.181502	530335648
Telangana	9276134	186490	662977.9	119193.9	716122.8	16969.08	189609.9	69181.71	242215.1	746335.8	0	0.00000	0	7051.25	0.076015	190397104
Tripura	1261376	15349	19518	12121	27443	5460	3798	9224	3899	154021	0	0.00000	0	0	0	0
Uttar Pradesh	35141980	189551.9	539266.4	279559.4	1411141	72191.38	427222.5	114561.1	313916.2	2632484	10009.25	0.02848	155658848	54641.19	0.155487	538766080
Uttarakhand	1947767	13561.16	63116.09	2849.17	17136.5	2118.67	3870.5	3716.33	14550.58	110704.2	0	0.00000	0	3201	0.164342	67219416
West Bengal	22313012	483270.4	478225.4	691942	762429.4	314548.3	369419	335395.3	380796.8	2221245	8332.58	0.03734	87452976	96159.3	0.430956	1704838784
Total	**265671317.1**	**4211739**	**9289643**	**3961739**	**10405049.72**	**2164162**	**4502187**	**2025294**	**4024937**	**21758764**	**150407.7**	**0.05661**	**2644484536**	**442802.4**	**0.166673**	**7055489624**

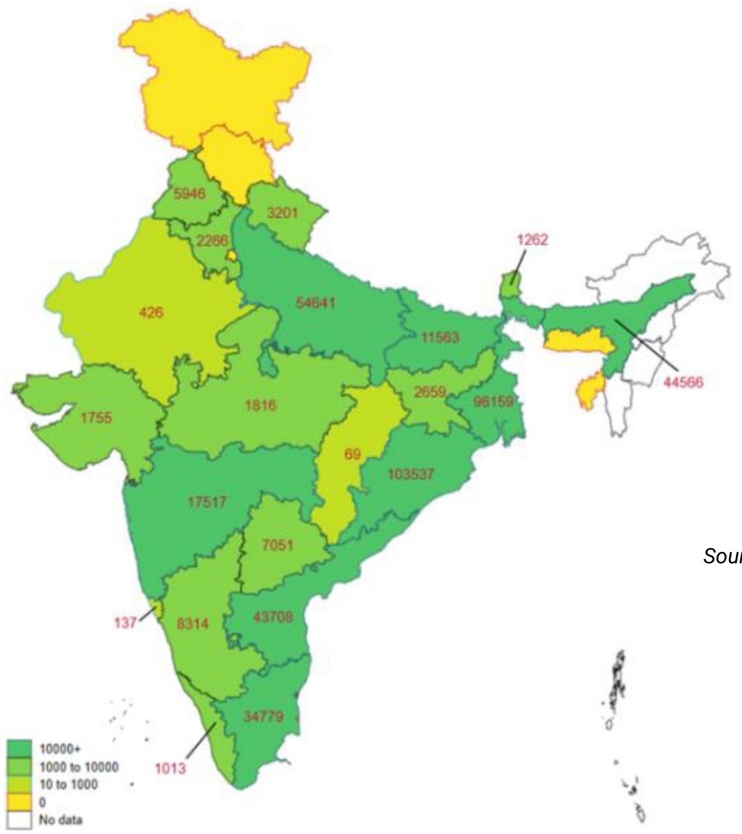
Source: Authors' Calculations

Appendix C: State-wise number of qualifying HHs (Income criterion)



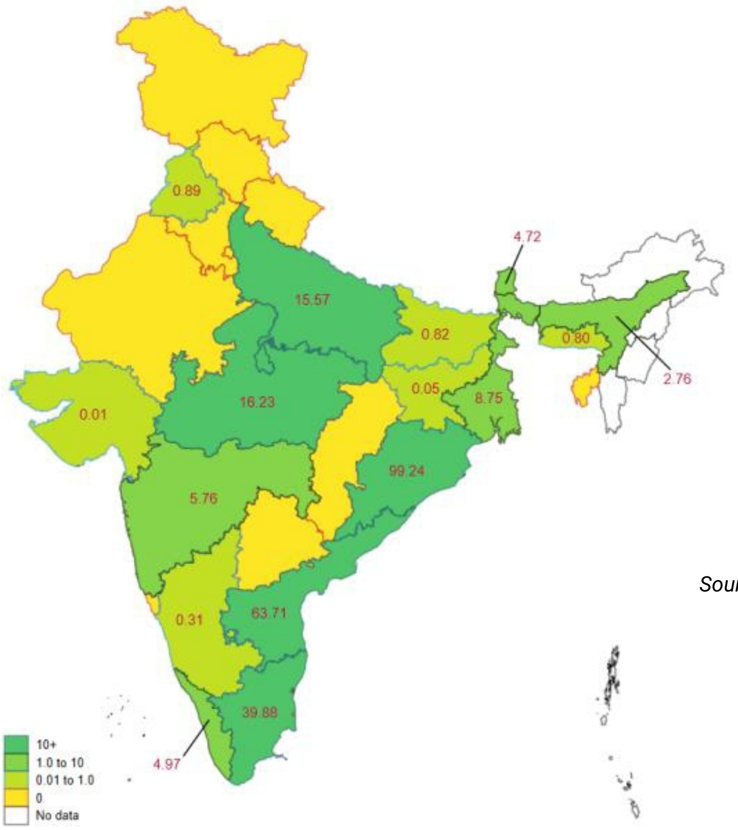
Source: Authors' Calculations

Appendix D: State-wise number of qualifying HHs (Expenditure criterion)



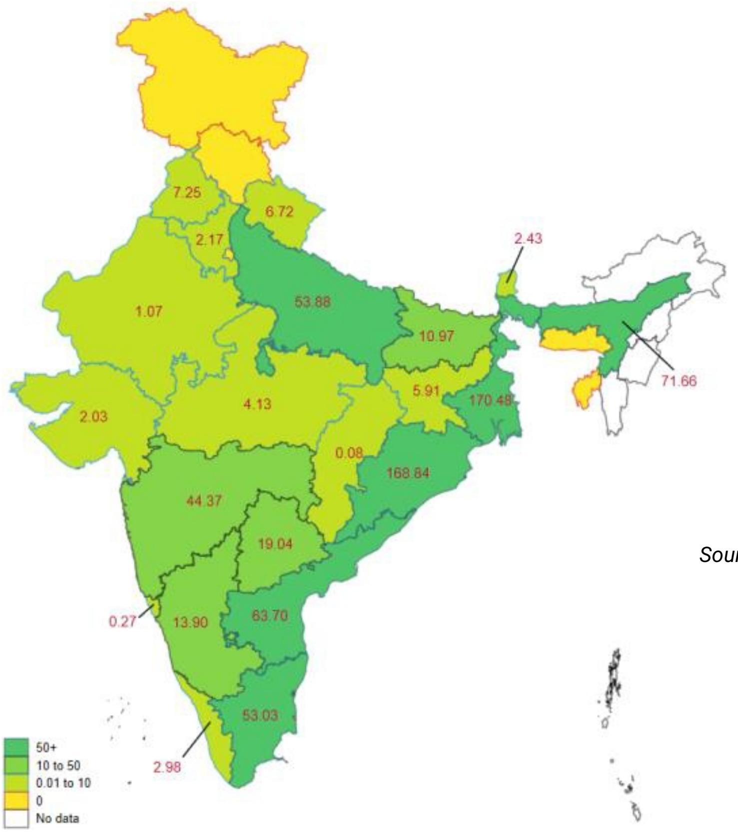
Source: Authors' Calculations

Appendix E: State-wise amount of qualifying debt in INR Crores (Income criteria)



Source: Authors' Calculations

Appendix F: State-wise amount of qualifying debt (Expenditure criteria)



Source: Authors' Calculations

References

Abu Alfeilat, H. A., Hassanat, A. B., Lasassmeh, O., Tarawneh, A. S., Alhasanat, M. B., Eyal Salman, H. S., & Prasath, V. S. (2019). Effects of distance measure choice on k-nearest neighbor classifier performance: a review. Big data, 7(4), 221-248.

Bhattacharya, D., & Ghosh, I. (2022). Must Shylock Receive his Pound of Flesh? An Ethical Inquiry into the Insolvency and Bankruptcy Regimes for Natural Persons . https://ibbi.gov.in/uploads/whatsnew/b5fba368fbd5c5817333f95fbb0d48bb.pdf.

Bhattacharya, D., & Ananth , B. (2021). The Curious Case of Fresh Start Process: An inquiry into historical antecedents, international analogues, and domestic substitutes. https://ibbi.gov.in/uploads/whatsnew/1d8b31fc65f7ac6f09a973be8f12f868.pdf

Carlos A. et al. (2019), "Bankruptcy, Discharge, and the Emergence of Debtor Rights in Eighteenth-Century England." Enterprise & Society, Vol. 20(2) pp. 457-506.

Chomboon, K., Chujai, P., Teerarassamee, P., Kerdprasop, K., & Kerdprasop, N. (2015, March). An empirical study of distance metrics for k-nearest neighbor algorithm. In Proceedings of the 3rd international conference on industrial application engineering (Vol. 2).

Consumer Pyramids Dx. (2019). Consumer Pyramids Household Survey: Weights. Retrieved from https://consumerpyramidsdx.cmie.com

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. IEEE transactions on information theory, 13(1), 21-27.

Deng, Z., Zhu, X., Cheng, D., Zong, M., & Zhang, S. (2016). Efficient kNN classification algorithm for big data. Neurocomputing, 195, 143-148.

Dezan Shira & Associates (2023), "A Guide to Minimum Wage in India in 2023", India Briefing, 26 October.

D'Orazio, M. (2021). Distances with mixed type variables some modified Gower's coefficients. arXiv preprint arXiv:2101.02481.

Fix, E., & Hodges, J. L. (1989). Discriminatory analysis. Nonparametric discrimination: Consistency properties. International Statistical Review/Revue Internationale de Statistique, 57(3), 238-247.

Government of India. (2016, May 28). Insolvency and Bankruptcy Code. https://ibbi.gov.in//webadmin/pdf/legalframwork/2017/Jul/IBC%202016.pdf.

Gower, J. C. (1971). A general coefficient of similarity and some of its properties. Biometrics, 857-871. Citation count: 6086 (Source: Google Scholar)

Goyal, S. (2022). Handling class-imbalance with KNN (neighbourhood) under-sampling for software defect prediction. Artificial Intelligence Review, 55(3), 2023-2064.

Guo, G., Wang, H., Bell, D., Bi, Y., & Greer, K. (2003). KNN model-based approach in classification. In On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE: OTM Confederated International Conferences, CoopIS, DOA, and ODBASE 2003, Catania, Sicily, Italy, November 3-7, 2003. Proceedings (pp. 986-996). Springer Berlin Heidelberg.

Hahn, P., Grant, T., & Hall, N. (1998). A branch-and-bound algorithm for the quadratic assignment problem based on the Hungarian method. European Journal of Operational Research, 108(3), 629-640.

Hu, L. Y., Huang, M. W., Ke, S. W., & Tsai, C. F. (2016). The distance function effect on k-nearest neighbor classification for medical datasets. SpringerPlus, 5(1), 1-9.

Levinthal L. (1918), “The Early History of Bankruptcy Law” University of Pennsylvania Law Review and American Law Register, Vol. 66(5), pp. 223-250.

Livemint (2023), “Kejriwal government announces hike in minimum monthly wages for workers in Delhi”, The Mint, 19 October.

Maillo, J., Ramírez, S., Triguero, I., & Herrera, F. (2017). kNN-IS: An Iterative Spark-based design of the k-Nearest Neighbors classifier for big data. Knowledge-Based Systems, 117, 3-15.

Maillo, J., Triguero, I., & Herrera, F. (2015, August). A mapreduce-based k-nearest neighbor approach for big data classification. In 2015 IEEE Trustcom/BigDataSE/ISPA (Vol. 2, pp. 167-172). IEEE.

Mehta, S., Shen, X., Gou, J., & Niu, D. (2018). A new nearest centroid neighbor classifier based on k local means using harmonic mean distance. Information, 9(9), 234.

MFIN. (2023). MFIN Micrometer Data as on 30 September 2023. Retrieved from https://mfinindia.org/Resources/micrometer

National Sample Survey Office. (2019). Note on Sample Design and Estimation Procedure of NSS 77th Round.

Reserve Bank of India. (2023). Reserve Bank of India Bulletin. Retrieved from https://rbidocs.rbi.org.in/rdocs/Bulletin/PDFs/RBIBULLETINDECEMBER20238F42E694CD9E4937824D9FCCF408F7A5.PDF

Song, Y., Liang, J., Lu, J., & Zhao, X. (2017). An efficient instance selection algorithm for k nearest neighbor regression. Neurocomputing, 251, 26-34.

Van de Velden, M., Iodice D'Enza, A., & Markos, A. (2019). Distance-based clustering of mixed data. Wiley Interdisciplinary Reviews: Computational Statistics, 11(3), e1456.

INSOLVENCY LAW ACADEMY
5, Mathura Road, Jangpura-A, New Delhi-110014
message@insolvencylawacademy.com

DVARA RESEARCH FOUNDATION
Dvara Holdings, 10th Floor
Phase 1 IIT-Madras Research Park,
Kanagam Village, Taramani, Chennai 600113

www.insolvencylawacademy.com

www.dvararesearch.com

insolvency-law-academy insolacademy facebook.com/insolacademy insolvency_law_academy