Housing deficit and census microdata: methodological framework, challenges, and perspectives

Déficit habitacional e microdados censitários: fluxo metodológico, desafios e perspectivas

Priscila Kauana Barelli FORCEL [I]
Elza Luli MIYASAKA [II]
Tiago Augusto da CUNHA [III]

Abstract

This article discusses the methodological framework for calculating the housing deficit in Brazil using microdata from the Population Census, based on the methodology developed by the João Pinheiro Foundation (FJP). It presents the data processing procedures, the variables used, and the criteria for identifying the components of the deficit. The study highlights the gains in transparency and the detailed territorial analysis enabled by the use of microdata, while also addressing methodological and operational challenges, such as the adaptation of variables across different databases and the opacity of certain routines applied by FJP. Finally, the article outlines the main challenges and prospects regarding the application of the methodology to the 2022 Census data and advocates for systematic documentation and open science as key strategies to improve the measurement of the housing deficit in Brazil.

Keywords: social housing; population census; open science; quantitative methodology.

Resumo

Este artigo discute o fluxo metodológico de cálculo do déficit habitacional a partir dos microdados do Censo Demográfico, com base na metodologia da Fundação João Pinheiro (FJP). São apresentados os procedimentos de tratamento dos dados, as variáveis utilizadas e os critérios para identificar os componentes do déficit. A análise destaca os ganhos de transparência e detalhamento territorial proporcionados pelo uso dos microdados, ao mesmo tempo que problematiza desafios metodológicos e operacionais, como a adaptação das variáveis entre diferentes bases e a opacidade de rotinas aplicadas pela FJP. Por fim, o texto aponta desafios e perspectivas para a aplicação da metodologia com os dados do Censo 2022 e defende a documentação sistematizada e a ciência aberta como caminhos para o aprimoramento da mensuração do déficit habitacional.

Palavras-chave: *habitação social; censo demográfico; ciência aberta; metodologia quantitativa.*

Introduction

The relationship between industrial capital and the production of the city has shaped the multiple morphological dimensions of the urban, particularly the unfolding of social classes and the location of housing. Beginning in the 1960s, the accelerated process of urbanization-driven by the pursuit of better living conditions and integration into the labor market-generated a growing demand for housing, transportation, and infrastructure. This movement unfolded in a context of severe exploitation of the labor force and concentration of wealth (Maricato, 2002).

Urban space came to be configured as a territory of disputes, in which the distribution of social groups does not occur randomly (Villaça, 2012). The less valued areas of intra-urban space were allocated to the most vulnerable populations, whose income pushed them toward the peripheries and devalued regions of the cities. This displacement forced the poor to occupy the urban margins, while at the same time masking social inequalities and naturalizing processes artificially produced, such as neoliberal investments in the spatial distribution of jobs, differentiated patterns of urban mobility, and the logic of real estate valorization (Bonduki, 1998; Maricato, 1999, 2002).

The disparity between the minimum wage and the high costs of housing made it impossible for most workers to access the formal housing market (Maricato, 1999). The unequal distribution of classes in urban space expresses a structure of political, economic, and ideological domination (Villaça, 2011), which appropriates the advantages of produced space, especially with regard to time and accessibility.

In 2022, the housing deficit in Brazil surpassed 6 million households, while more than 26 million presented some type of inadequacy (FJP, 2024a). Both phenomena show higher percentages among women and *pardos* (a census category referring to mixed-race individuals), with the Southeast region standing out (FJP, 2024a). The housing deficit is a historical, political, and structural problem, influenced by the economic processes of ultraneoliberalism, which directly affect land issues and the provision of social housing (Arantes, Vainer & Maricato, 2002; Maricato, 2015; Rolnik, 2015; Prieto & Laczynski, 2020).

Understanding the housing deficit increasingly requires an approach that considers the complexity of population and mobility dynamics that shape urban space. Internal and international migrations, for example, directly influence the unequal occupation of territory and precarious forms of housing insertion, especially in metropolitan areas. Studies indicate that territorial mobility processes condition access to both labor markets and housing, with housing precariousness serving as a concrete expression of the social and spatial inequalities that structure Brazilian cities (Pasternak et al., 2022; Bógus & Magalhães, 2020).

In this context, this article aims to analyze the application of Census microdata (i.e., anonymized individual-level records made available for research purposes) to calculate the housing deficit according to the methodology of Fundação João Pinheiro (FJP). The study presents the historical evolution of the methodology and the methodological flow adopted for its application in the Statistical Package for the Social Sciences (SPSS), based on the 2010 Census microdata.[1] It also discusses the possibilities of replicating FJP's methodology with the release of 2022 Census microdata. The choice of SPSS is justified by its capacity to efficiently handle large databases and by its user-friendly interface, which does not require advanced programming skills.

Furthermore, the article discusses the advances and challenges in applying the methodology at the weighting area level using 2022 Census microdata, especially given the absence of variables that directly correspond to the criteria defining the housing deficit. Thus, understanding the applicability of this methodology requires a detailed analysis of its evolution and the criteria it comprises.

Housing deficit and the FJP methodology

The housing deficit is an essential indicator for planning public housing policies, as it reflects unmet housing needs, particularly among the most vulnerable groups. This concept encompasses not only the shortage of dwellings but also housing inadequacy, tenure insecurity, and barriers to accessing decent and affordable housing (UN, 2015).

According to the United Nations (UN, 2015), adequate housing must ensure privacy, structural stability, security of tenure, and access to basic infrastructure such as water supply, sanitation, waste collection, and electricity. In addition, it should be built with durable and healthy materials, be located in areas with access to essential services, and provide sufficient space to prevent overcrowding.

The measurement of the housing deficit in Brazil has been the subject of debate since the 1990s. For Russo (2017), the indicator seeks to represent housing deprivation based on criteria such as precarious housing, excessive density, and lack of urban infrastructure. Its analysis makes it possible to identify the most vulnerable territories and population groups and to guide the design of public policies and the allocation of investments aimed at reducing the deficit and improving housing conditions.

More than a statistical measure, the housing deficit is a concept shaped by political and methodological disputes. Researchers, policymakers, and economic sectors play a fundamental role in shaping housing policies, producing official diagnoses, and allocating public resources (Russo, 2017). In Brazil, the main reference for this measurement is the methodology developed by Fundação João Pinheiro, which is widely used by governmental and academic institutions. However, this methodology presents limitations, both in the interpretation of data and in the definition of criteria used to characterize the deficit.

The FJP methodology was developed in partnership with the United Nations Development Programme (PNUD) and was first released in 1995, based on the 1990 National Household Sample Survey (PNAD) and the 1991 Population Census. Since then, it has undergone conceptual and operational modifications to adapt to the variables available in official surveys (FJP, 2013).

According to FJP (ibid.), the housing deficit is composed of two main axes: (I) the deficit in the strict sense, which represents the need for new housing units; and (II) housing inadequacy, which refers to the need for improvements in existing dwellings (Genevois & Costa, 2001; FJP, 2013). To measure these demands, FJP distinguishes between the quantitative deficit – related to the shortage of housing units – and the qualitative deficit, linked to structural precariousness and the lack of infrastructure in existing dwellings. This article focuses on the quantitative deficit, as it can be estimated from census microdata disaggregated by weighting areas.

The methodology has undergone several reformulations since 1991 in order to keep pace with changes in urban dynamics, data availability, and housing planning needs. Initially, the concept was restricted to the need for replacement or expansion of housing, but

beginning in 2000 it started to distinguish between replacement and expansion. Accordingly, its components were also refined.

The component of precarious housing has been part of the methodology since 1991 and includes two subgroups: rustic dwellings, characterized by materials such as reused wood, uncovered wattle and daub, straw, or other fragile materials; and improvised dwellings, located in spaces not originally intended for housing, such as commercial buildings, underpasses, boats, or shacks (FJP, 2010). This component seeks to quantify the share of the population living in precarious conditions in order to support the formulation of public policies aimed at improving housing conditions and combating poverty, and it refers to the materiality of the buildings. In 2016, the *Cadastro Único* (CadÚnico, Brazil's Unified Registry for Social Programs) began to be used to estimate improvised dwellings; and in 2019, complementary administrative data were integrated to improve classification (FJP, 2021).

Involuntary family cohabitation was initially treated in absolute terms – that is, without distinguishing between voluntary and involuntary situations (FJP, 1995; FJP, 2019). Starting in 2000, the concept was segmented, and in 2007 it was reformulated to consider only cases with a manifest intention to form a new household (FJP, 2001; FJP, 2009). Since the Census does not have a specific variable for this information, FJP developed an algorithm to estimate this intention (FJP, 2010; Miranda--Ribeiro, Viana & Azevedo, 2015). In 2016, it became possible to identify secondary family units within households, and in 2019 a new refinement was introduced to reduce the overestimation of this component (FJP, 2020; FJP, 2021).

The component of excessive rent burden was initially considered housing inadequacy in 1991 and was incorporated into the deficit in 2000, when financial overload with rent was recognized as a factor of housing exclusion (FJP, 2001). Since 2010, households with an income of up to three minimum wages that spend more than 30% of their monthly income on rent have been classified as part of the deficit (FJP, 2009). In 2019, discussions began on the need to differentiate structural from conjunctural rent burden, considering fluctuations in the housing market (FJP, 2020; FJP, 2021).

Excessive crowding, in turn, was not part of the deficit until 2000, being considered only a form of housing inadequacy. In 2007, it began to be incorporated into the deficit, but only for rented households; for owner-occupied households, it remained classified as inadequacy (FJP, 2009; Miranda-Ribeiro, Viana & Salis, 2013). The criterion was maintained in 2010 and 2016, with adjustments to the residents-to-bedroom ratio. In 2019, the impact of excessive crowding on urban housing dynamics was reassessed, and in 2022 it was removed from the methodology due to its overlap with the cohabitation criterion (FJP, 2021; FJP, 2024a). According to FJP (2020), the justification was to avoid double counting and improve the accuracy of estimates. However, several authors had already pointed out the limits of this criterion, as it evaluated crowding solely by bedroom density, without considering socioeconomic and cultural factors or the overall floor area of the dwellings (Feitosa, 2022; Givisiez & Oliveira, 2022; Arriagada Luco, 2022).

The methodology adopts a hierarchy of criteria: a household classified as improvised, for instance, is automatically excluded from the analysis of other deficit components, such as cohabitation, rent burden, or crowding (FJP, 2010). Applying the methodology to Census microdata makes it possible to calculate the Housing Deficit at the municipal level (ibid.; Miranda-Ribeiro, Viana & Azevedo, 2015). The municipality is the immediate territory of everyday life; it is where deficiencies are

expressed and where public policies are implemented. Local analysis is, therefore, the first step toward understanding and reorganizing the city (Cunha et al., 2024).

Despite its widespread adoption, the FJP methodology presents important limitations. Feitosa (2022) and Givisiez & Oliveira (2022) point out that its traditional application disregards population dynamics, such as the formation of new families and migratory flows, treating the deficit as a static phenomenon. In addition, there is evidence of overestimation of involuntary cohabitation and difficulties in territorializing the deficit at intra-urban scales, given the aggregation by weighting areas (Bonduki, 2022; Arriagada Luco, 2022).

Another critical issue lies in the separation between the quantitative and qualitative deficit. FJP assumes that land tenure inadequacy can be resolved through regularization and improvements, without the need for new housing units. However, studies indicate that this approach underestimates the real housing demand, since many informal settlements face legal and structural obstacles that make such regularization unfeasible (Nascimento & Braga, 2009; Lacerda, Pimenta & Ferreira, 2022). Added to this is the high cost of rent and the precarization of the housing market, which intensify the deficit over time.

Historically, the housing deficit in Brazil has been addressed in a predominantly quantitative manner, emphasizing the shortage of dwellings. However, this approach overlooks fundamental aspects such as location, housing quality, and access to urban infrastructure. Authors such as Bonduki (2022); and Nascimento & Braga (2009) argue for a broader concept of the deficit, one that incorporates the actual living conditions of the population. The existence of millions of vacant dwellings in the country, in contrast with the reality of families living in involuntary cohabitation or in precarious

housing, demonstrates that the problem is not limited to the production of new units but also involves the unequal distribution of housing and its economic accessibility.

Therefore, the housing deficit in Brazil reflects a complex set of factors that go beyond the mere absence of dwellings. It involves socioeconomic inequalities, insecure tenure, and limited access to infrastructure. While the Fundação João Pinheiro methodology has been consolidated as a national reference, it requires refinements that better incorporate demographic and territorial dynamics in order to provide a more accurate reading of Brazil's housing reality.

Population census for housing deficit calculation

The methodology for calculating the housing deficit, developed and consolidated by Fundação João Pinheiro (FJP), is primarily based on data from the Continuous National Household Sample Survey (PNAD Contínua).[2] This database allows for regular monitoring of the deficit; however, its sample-based nature limits territorial disaggregation and restricts estimates to the state level and, in some cases, to metropolitan regions. Consequently, it does not allow for detailed analysis at the municipal or sectoral level, which compromises more precise diagnoses of local housing conditions.

Given this limitation, the Population Census emerges as a viable methodological alternative, offering full coverage of the population and enabling analyses at the municipal level, especially by weighting areas. Unlike the annually conducted PNAD Contínua, the Census is decennial and provides a more comprehensive snapshot of housing conditions. However, its periodicity represents a

significant challenge, as data become outdated in intercensal periods, which justifies the complementary use of PNAD Contínua for periodic updates (FJP, 2010; Miranda-Ribeiro, Viana & Azevedo, 2015).

Applying the FJP methodology to the 2010 Population Census allowed significant gains in spatial precision and enabled the calculation of the housing deficit at the municipal scale, particularly through weighting areas. This granularity was essential for understanding the territorial distribution of the phenomenon and identifying areas of higher housing vulnerability. However, the use of Census data

presents methodological challenges, particularly due to the absence of direct information on involuntary cohabitation – a central component in measuring the housing deficit.

To overcome this limitation, FJP developed an algorithm that indirectly estimates secondary cohabiting families with the intention of forming their own household, using a reduction factor. This factor is calculated based on the average percentages of cohabiting families in deficit observed in the 2009 and 2011 PNAD surveys and is applied according to municipal size (Miranda-Ribeiro, Viana & Azevedo, 2015), as presented in Chart 1.

Chart 1 – Groups of municipalities for the formulation of the algorithm to identify secondary cohabiting families with the intention of forming an exclusive household

Criteria applied to secondary cohabiting families observed in the 2010 Population Census to determine the total number of cohabiting families intending to form an exclusive household, according to municipality groups		
Group	**Group description**	**Criterion for determining the total number of cohabiting families in deficit**
Group 1	Municipalities belonging to the Metropolitan Regions (RM) of Belém, Fortaleza, Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo, Curitiba and Porto Alegre	Average of the percentages observed in the 2009 and 2011 PNAD surveys in the respective MR
Group 2	State capitals not included in Group 1	Average of the percentages observed in the 2009 and 2011 PNAD surveys in the MR of the same geographic region (when an MR exists in the region) Average of the percentages observed in the 2009 and 2011 PNAD surveys in an MR from another geographic region (when no MR exists in the region)
Group 3	Municipalities with a population above 200,000 not included in Groups 1 and 2	Adopts the criterion of the state capital
Grupo 4	Other municipalities	Average of the percentages observed in the 2009 and 2011 PNAD surveys for the respective state

Source: FJP (2013) and Miranda-Ribeiro, Viana & Azevedo (2015).

The use of the reduction factor allows for the alignment of Census data with the patterns observed in the PNAD surveys and helps to smooth out variations in the time series of the housing deficit. However, this strategy also has limitations: it assumes stability in the relationship between the two data sources over time, which does not always hold given economic, social, and political changes. In addition, its uniform application may obscure regional specificities, particularly in Group 4, which encompasses municipalities ranging from fewer than 1,000 to 199,999 inhabitants, reflecting considerable internal heterogeneity.

Comparisons between the Continuous PNAD and the Population Census show that both sources have advantages and limitations. PNAD provides frequent updates but lacks spatial detail, whereas the Census allows for more precise diagnoses, although at decennial intervals. Therefore, the combination of both sources, together with improvements in estimation methods, is essential for the accurate measurement of the housing deficit in Brazil.

In 2010, the methodological alignment between PNAD and the Census was particularly favorable, allowing for the direct application of the FJP methodology at the intra-urban level, based on variables available in both surveys.

Chart 2 presents a comparison of the variables used in the two data sources for calculating the components of the housing deficit.

By manipulating these variables, it becomes possible to calculate the housing deficit not only at the municipal level but also by weighting areas—geographical units formed by exclusive groupings of census tracts. These areas allow for the application of more refined statistical procedures, such as calibrating estimates based on total population data. According to IBGE (2012), sample data can be recovered down to the district, subdistrict, or weighting area level, provided there are at least 400 occupied private dwellings.

Weighting areas correspond to the Census sample questionnaire, which collects more detailed information on the population's socioeconomic characteristics, such as employment, income, education, and migration. They can be accessed through the 2010 Census microdata (variable V0011 – Weighting Area) or directly on the IBGE website via https://censo2010.ibge.gov.br/apps/areaponderacao/.

Regarding the 2022 Population Census, although it has advanced in the collection of housing data, it also presents methodological challenges. One of the main issues is the impossibility of directly estimating the excessive

Chart 2 – Variables for calculating the Housing Deficit:
comparison between PNAD and the 2010 Population Census

Deficit component		PNAD	2010 Population Census
Precarious Housing	Improvised Dwellings	V0201 – Dwelling Category	V4001 – Category of Unit Visited Classification: 5 – Occupied improvised private dwelling
	Rustic Dwellings	V0202 – Type of Dwelling V0203 – Predominant material used in the construction of external walls of the building	V0202 – Predominant material of external walls Classification: 5 – Uncoated wattle and daub 6 – Reused wood 7 – Straw 8 – Other material 9 – No wall
Involuntary family cohabitation	Room	V0202 – Type of Dwelling	V4002 – Type of Unit Visited Classification: 14 – Housing in: rooming house, tenement, or substandard dwelling
		V0403 – Family Number	V0201 – Occupancy Status Classification: 1 – Owned by a household member – fully paid 2 – Owned by a household member – still being paid 3 – Rented 5 – Granted by other means 6 – Other condition
	Cohabiting Families	V0410 – Intention to move and build another dwelling	V5020 - Family Number Classification: Secondary to ninth cohabiting family V5130 - Logical order of the family. Classification: 1 V0502 - Kinship or cohabitation relationship with the household head Classification: 4 – Child of head and spouse 5 – Child of head only 6 – Stepchild 7 – Son-/Daughter-in-law 8 – Parent or stepparent 9 – Parent-in-law 10 – Grandchild 11 – Great-grandchild 12 – Sibling 13 – Grandparent 14 – Other relative 15 – Household member 16 – Cohabitant 17 – Boarder See Chart 1 for the application of the reduction factor by municipal size
Excessive Rent Burden		V4105 – Dwelling Situation	V1006 – Dwelling Situation Classification: 1 – Urban
		V0203 – Predominant material used in the construction of external walls of the building	V0201 – Dwelling Condition Classification: 3 – Rented
		V0105 – Total residents	–
		V4621 – Household *per capita* monthly income	V6529 – Household income
		V0208 – Monthly rent paid in the reference month	V2011 – Rent amount
		V4722 – Household monthly income for all housing units	–
Excessive Occupancy in Rented Households		V0207 – Dwelling Condition	V0201 – Dwelling Condition Classification 3 – Rented
		V0105 – Total residents	V6204 – Residents per bedroom. When greater than three (3)
		V0206 – Number of rooms used as bedrooms	–

Source: Fundação João Pinheiro (2013); Miranda-Ribeiro, Viana & Azevedo (2015); 2010 Population Census; adapted by the authors in 2025.

rent burden component (FJP, 2024b), which highlights not only the need to integrate different data sources but also the urgency of reviewing and expanding census questions to ensure the measurement of fundamental components of the housing deficit. Chart 3 summarizes the main changes between the 2010 and 2022 Censuses regarding the variables used in calculating the housing deficit.

Chart 3 summarizes the main methodological changes observed between the 2010 and 2022 Population Censuses and considers the four components of the housing deficit defined by Fundação João Pinheiro: precarious housing, family cohabitation, excessive urban rent burden, and overcrowding in rented households. These changes reflect a significant methodological review by Fundação João Pinheiro, but also highlight growing challenges in maintaining the continuity of historical housing deficit series, particularly regarding comparability between censuses.

For precarious housing, methodological continuity is observed, with direct estimates available for both improvised and rustic dwellings. Regarding family cohabitation, there was a methodological revision: while in 2010 the estimate was obtained indirectly through the identification of cohabiting families intending to establish a separate household, in 2022 FJP proposes a reconstruction of family units based on the use of rooms as bedrooms and the degree of overcrowding in these spaces.

The greatest methodological challenge lies in measuring the excessive urban rent burden, which was unfeasible in the 2022 Census due to the absence of a variable directly reporting rent expenditure. This gap prevents the application of traditional criteria for household income committed to housing and requires the development of alternative strategies, such as the use of auxiliary data from PNAD Contínua or the construction of imputation models.

Chart 3 – Survey items for calculating the housing deficit, 2010 and 2022

Deficit component		Population Census 2010	Population Census 2022
Precarious Housing	Improvised Dwellings	Direct estimate	Direct estimate
	Rustic Dwellings	Direct estimate	Direct estimate
Cohabiting Families	Cohabiting Families (2010) Exclusive Cohabiting Units (2022)	Cohabiting families obtained indirectly. Only families with the intention to establish a separate household (own indirect estimate)	Estimate through reconstruction of family units + overcrowding in rooms serving as bedrooms > 2
	Room	Excludes rooms provided by employer	Excludes rooms provided by employer
Excessive urban rent burden		Direct estimate	Absence of variable/ question on rent expenditure makes estimation impossible
Excessive occupancy in rented households		Direct estimate	Not included

Source: FJP (2023), adapted by the authors in 2025.

Given the unavailability of an explicit variable on rent payments in the 2022 Census, FJP proposed an alternative approach to estimate the excessive rent burden by constructing a proxy based on the ratio between monthly rent and household income reported in the microdata. This strategy requires methodological caution, as the income effectively available for rent payment is not directly reported in the census questionnaire. Moreover, the quality of the estimate strongly depends on the completeness and consistency of respondents' information. Nevertheless, the proxy allows capturing, even if partially, the financial effort of renting families. The analysis can be further refined by cross-referencing variables such as family composition, age range of household heads, and geographic location, enabling a more nuanced assessment of housing burden in urban territories.

Additionally, the component of overcrowding in rented households, included in the application of the methodology to the 2010 Census, was removed in 2022. This exclusion represents a significant setback in the approach to the deficit, as it eliminates a relevant qualitative dimension regarding housing conditions in urban areas. Its removal stems from the incompatibility between the variables available in the 2022 Census and the criteria defined by FJP. Estimating this specific type of overcrowding requires two combined pieces of information: (1) dwelling occupancy type (owned, rented, granted, etc.); and (2) the number of rooms for exclusive use, which allows the calculation of household density. However, the 2022 Census does not safely link these two dimensions, making the application of the criterion unfeasible (FJP, 2024b).

Unlike PNAD Contínua, where these variables are available in a compatible format, the 2022 Census does not allow replication of this component, which justified its removal from the methodology. This limitation highlights the challenges faced when transitioning between different data sources and reinforces the need for caution when comparing historical housing deficit series.

Methodological flow in SPSS

The methodology of Fundação João Pinheiro for calculating the housing deficit was applied to the 2010 Population Census microdata using SPSS software. The process began with the organization and cleaning of the microdata by fragmenting the original files into smaller analytical units. This step required separating information related to households and individuals while preserving the common identifier, variable V0300 – control, which allows for the subsequent merging of the databases.

Next, the routine for calculating the housing deficit components was executed based on FJP criteria. Derived variables were created to identify: precarious housing; Involuntary family cohabitation; excessive rent burden; and excessive occupancy. The SPSS syntax developed automates case selection, applies filters, classifies households according to the four components, and generates a synthetic variable indicating the presence or absence of a deficit.

The methodological flow was structured to ensure replicability and consistency, leveraging SPSS capabilities for processing large datasets and executing statistical routines in an accessible and transparent manner (Forcel, 2023). The following sections detail the criteria used to identify each dimension of the housing deficit .

Precarious Housing:

a) Improvised Dwellings

The variable V4001 – Type of Dwelling was used. Dwellings classified as "5," corresponding to "occupied improvised private dwelling", were considered improvised. A variable "IMPROVISADO" was created, assigned a value of 1 for this group and 0 for all others. According to IBGE (2010), this type of dwelling refers to locations where residents cohabit due to family ties, domestic dependency, or social norms, but which are situated in structures not originally intended for housing, such as shops, factories, tents, shacks, trailers, or unfinished constructions. Although inadequate, these spaces were occupied as residences at the Census reference date.

b) Rustic Dwellings

A variable "RUSTICO" was created to indicate the presence of at least one of the conditions recorded in V0202 – Predominant material of external walls, in classifications 5 to 9. These categories include dwellings with uncoated adobe walls, reused wood, straw, materials such as zinc or plastic, or constructions without any walls, representing precariousness and inadequacy for residential purposes. If any of these conditions were present, the dwelling was assigned a value of 1 in the "RUSTICO" variable.

From these subcategories, the variable "DOMPREC" was created, identifying dwellings classified as 1 in either "IMPROVISADO" or "RUSTICO". This consolidation allowed the identification of dwellings falling under the precarious housing dimension.

Involuntary family cohabitation:

The cohabitation variable was created in the household database using V5020 – Family Number, considering values from 2 to 9, i.e., second- to ninth-order cohabiting families. This variable derives from identifying families within households based on their relationship to the household head, as well as information on maternity and marital status (IBGE, 2010).

To compose the cohabitation component, V5130 – Logical order in the family, indicating each individual's sequence within the family, and V0502 – Relationship or cohabitation with the household head, categories 4 to 17 as detailed in Chart 2, were also considered. The FJP-recommended reduction factor, according to municipal size (see Chart 1), was applied to adjust cases classified as cohabitation.

The variable "COABITACAO" was assigned a value of 1 for individuals meeting these criteria. In SPSS, this required aggregating individual data at the household level to verify the presence of multiple families in the same dwelling.

For identification of cohabitation by rooms, it would be necessary to select dwellings with classification 1, 2, 3, 5, or 6 in V0201 – Occupancy condition, and a value of 14 in V4002 – Type of unit, corresponding to rooming houses, tenements, or similar dwellings. However, this condition was not analyzed for 2010, as specified in Chart 3.

Excessive Rent Burden:

The composition of this dimension required combining criteria. Households classified as urban in V1006 – Dwelling situation (value 1) and rented in V0201 – Occupancy condition (value 3) were considered. When monthly household income (V6529) was greater than zero, the ratio between rent (V2011 – Rent value, in BRL) and household income was calculated. Households with a ratio above 0.3 were classified as having a deficit due to excessive rent. The variable "ONUS_EXCESSIVO" was assigned to cases meeting all these conditions.

Overcrowding in Rented Dwellings:

To identify overcrowded households, dwellings with V6204 – Resident density per bedroom greater than 3 were selected, provided they were classified as rented in V0201 – Occupancy condition (value 3). The variable "ADENSAMENTO" was assigned to cases meeting these criteria.

Consolidation of Housing Deficit Components:

To identify households with a total housing deficit, the variable "DEFICIT_PRIORITARIO" was created to store the priority deficit component, according to the hierarchy defined by FJP. Initially, this variable is set to 0, indicating the absence of a deficit. Prioritization rules were applied in the following order: households classified as precarious were assigned a value of 1; if not precarious but exhibiting family cohabitation with the reduction factor applied, they received a value of 2; households with excessive rent, not covered by previous conditions, were classified as 3; finally, rented households with overcrowding, not falling under other criteria, were assigned a value of 4.

From this variable, "DEFICIT_TOTAL" was created, with a value of 1 for households exhibiting at least one deficit component (i.e., DEFICIT_PRIORITARIO > 0) and 0 for all others. In addition, indicator variables were created for each component to allow disaggregated analyses: "DOMPREC", "COABITACAO", "ONUS_EXCESSIVO" and "ADENSAMENTO", which take a value of 1 when the household was primarily classified under that component and 0 otherwise.

Chart 4 summarizes the variables extracted from the 2010 Demographic Census microdata used to operationalize the criteria defined by the Fundação João Pinheiro. Each component of housing deficit – precarious housing, family overcrowding, excessive rent burden, and overcrowding – was linked to specific variables from the Census questionnaire, to which technical recording criteria were applied. The variables were renamed and structured in SPSS to indicate, in binary form, the presence or absence of conditions characterizing the deficit. The consolidation of these components into the "DEFICIT_TOTAL" variable allows identification of households experiencing housing deficit according to FJP parameters, ensures methodological standardization, and enables comparative analyses over time and across different territorial scales.

Chart 4 – Variables used for calculating the housing deficit
based on the 2010 Census microdata

Component		Original variable (IBGE)	Applied criterion	Variable name (SPSS)
Precarious Housing	Improvised Dwellings	V4001 – Type of Dwelling	5	IMPROVISADOS
	Rustic Dwellings	V0202 – Predominant material of external walls	de 5 a 9	RÚSTICOS
	Precarious Housing	IMPROVISADO RÚSTICO	1	DOMPREC
Family Cohabitation	Co-resident Families	V5020 – Family Number	de 2 a 9	COABITAÇÃO
		V5130 – Logical order in the family	1	
		V0502 – Kinship or co-residence relationship with the head of household	de 4 a 17	
	Rooms	V0201 – Dwelling, occupancy status	1, 2, 3, 5 ou 6	COAB_COMODOS
		V4002 – Type of unit in the condition	14	
Excessive Rent Burden		V1006 – Dwelling status	1	ONUS_EXCESSIVO
		V0201 – Dwelling occupancy status	1	
		V6529 – Monthly Household Income in July 2010	> 0	
		V2011 – Rent amount (in BRL)	ratio between V2011 / V6529	
Overcrowding in Rented Dwellings		V6204 – Residents per bedroom density	>3	ADENSAMENTO
		V0201 – Dwelling occupancy status	3	

Source: authors' elaboration, 2025.

Final considerations

This text presented the methodological workflow for calculating the housing deficit based on the Demographic Census microdata, following the methodology proposed by the Fundação João Pinheiro. Direct application of census variables allows greater transparency in constructing the indicators and enables more detailed and territorially nuanced analyses of the housing phenomenon.

The text also highlighted challenges in replicating the methodology, particularly regarding variable adaptation across different databases, opacity in certain official procedures, and expected changes in the 2022 Census. Therefore, two main areas for improvement are identified: (a) documentation and standardization of routines used to calculate the deficit; and (b) provision of processed databases that allow replication and validation by the scientific community.

This discussion reinforces the importance of expanding dialogue between data-producing institutions and researchers to consolidate a culture of open science and collaborative knowledge production. Greater accessibility to calculation routines can not only improve methodology but also foster debates about the very definition of "housing deficit."

The systematization presented supports future research focused on the territorial disaggregation of the deficit and encourages critical reflection on the limits and potential of the current concept. With the release of 2022 Census data, it becomes even more urgent to develop strategies that combine technical rigor and sensitivity to socio-spatial transformations, ensuring that the indicators used in housing policies more accurately and fairly reflect housing inequalities in the country.

The Demographic Census, conducted by IBGE, is the main source of data on housing in Brazil. However, its decennial periodicity poses significant challenges for continuous interpretation of the housing deficit, since major urban and social changes between censuses are not captured in real time. Researchers such as Bonduki (2022) and Feitosa (2022) recommend intercensal surveys to mitigate this gap.

Another limitation of the Census is the lack of detailed information on property tenure and housing conditions. Many irregular households are classified as qualitative deficit when they should be considered quantitative, because the Census does not investigate the possibility of land regularization – information that directly affects social housing policy formulation.

Applying the FJP methodology to census microdata involves challenges, including identifying compatible variables and adapting criteria to the available level of disaggregation. The 2022 Census, in particular, introduced significant changes, including the removal or reformulation of central variables such as excessive rent burden and overcrowding in rented dwellings, which complicates reproducibility and historical comparability of the series.

The use of the PNAD Contínua reduction factor highlights a methodological tension in maintaining consistency with the Census, demonstrating the need for a deeper review of the methodology in light of new urban dynamics and worsening housing inequalities.

Therefore, it is urgent to advance flexible methodologies and continuous data updating strategies, incorporating remote sensing, administrative databases, and machine learning techniques. These approaches enhance the ability to capture nuances of urban reality and support more effective public policies aimed at reducing housing deficit and guaranteeing the right to adequate housing.

[I] https://orcid.org/0000-0002-1321-4716
Universidade Federal de São Carlos, Centro de Ciências Exatas e de Tecnologia,
Programa de Pós-Graduação em Engenharia Urbana. São Carlos, SP/Brasil.
priscilaforcel@ufscar.br

[II] https://orcid.org/0000-0003-4480-9672
Universidade Federal de São Carlos, Centro de Ciências Exatas e de Tecnologia, Departamento de
Engenharia Civil. São Carlos, SP/Brasil.
elza.miyasaka@ufscar.br

[III] https://orcid.org/0000-0002-4634-8497
Ministério das Cidades, Secretaria Nacional de Habitação, Departamento de Produção Social da
Moradia. Brasília, DF/Brasil
tiagoaugustodacunha@gmail.com

Acknowledgments

Notes

(1) Due to the unavailability of the 2022 microdata, it is still not possible to calculate the deficit by weighting
area; these are scheduled to be released in August 2024.

(2) The PNAD Contínua replaced the traditional PNAD starting in 2012, with official data released from 2015
onwards.

References

ARANTES, O.; VAINER, C.; MARICATO, E. (org.) (2002). *A cidade do pensamento único: desmanchando
consensos*. Petrópolis, Vozes.

ARANTES, P. F.; FIX, M. (2009). Minha Casa, Minha Vida, o pacote habitacional de Lula: Como o governo Lula
pretende resolver o problema da habitação. *Correio da Cidadania.* Disponível em: https://unmp.org.br/
wp-content/uploads/2009/08/PacotePedroMariana.pdf. Acesso em: 8 ago 2023.

ARRIAGADA LUCO, C. (2022). "Carências habitacionais na América Latina: formas de cálculo e desafios
emergentes". In: SANTOS, E. C. (org.). *Ensaios e discussões sobre o déficit habitacional no Brasil.*
Belo Horizonte, FJP/MDR. Disponível em: https://drive.google.com/file/d/1stJWtC0dMFtK_
Zz0G0zFBjDgw2WxjL_i/view. Acesso em: out 2023.

BÓGUS, L.; MAGALHÃES, L. F. A. (2020). "Imigração internacional na Macrometrópole Paulista (MMP): perfil da inserção laboral e especificidades regionais". In: BÓGUS, L.; PASTERNAK, S.; MAGALHÃES, L. F. A. (org.). *Metropolização, governança e direito à cidade: dinâmicas, escalas e estratégias*. São Paulo, Educ/PIPEq, pp. 337-374.

BONDUKI, N. (1998). *Origens da habitação social no Brasil: arquitetura moderna, lei do inquilinato e difusão da casa própria*. São Paulo, Estação Liberdade/Fapesp.

_____ (2022). "Metodologias de avaliação do déficit habitacional: avaliação crítica e perspectivas". In: SANTOS, E. C. (org.). *Ensaios e discussões sobre o déficit habitacional no Brasil.* Belo Horizonte, FJP/MDR. Disponível em: https://drive.google.com/file/d/1stJWtC0dMFtK_Zz0G0zFBjDgw2WxjL_i/view. Acesso em: out 2023.

CUNHA, T. A.; DA SILVA, J. C.; ALMEIDA, A. M.; FORCEL, P. K. B.; MORENO, F. A.; GARCIA, R. P.; MIYASAKA, E. L.; CUERVA, G. A. (2024). Método de estimativa do déficit habitacional por aprendizado de máquina para uma cidade de porte médio: Ribeirão Preto-SP, Brasil. *PLURIS*. Disponível em: https://www.civil.uminho.pt/planning/Pluris2024/Atas/Papers/Paper817.pdf.

FEITOSA, F. F. (2022). "Aspectos territoriais na estimativa das necessidades habitacionais: Construindo alternativas metodológicas". In: SANTOS, E. C. (org.). *Ensaios e discussões sobre o déficit habitacional no Brasil*. Belo Horizonte, FJP/MDR. Disponível em: https://drive.google.com/file/d/1stJWtC0dMFtK_Zz0G0zFBjDgw2WxjL_i/view. Acesso em: out 2023.

FJP – Fundação João Pinheiro (1995). *Centro de Estudos Políticos e Sociais. Déficit habitacional no Brasil.* Belo Horizonte.

_____ (2001). *Centro de Estatística e Informações. Déficit habitacional no Brasil 200*0. Belo Horizonte.

_____ (2005). *Centro de Estatística e Informações. Déficit habitacional no Brasil, municípios selecionados e microrregiões geográficas*. 2. ed. Belo Horizonte.

_____ (2006). *Centro de Estatística e Informações. Déficit habitacional no Brasil 2005*. Belo Horizonte.

_____ (2008). *Centro de Estatística e Informações. Déficit habitacional no Brasil 200*6. Belo Horizonte.

_____ (2009). *Centro de Estatística e Informações. Déficit habitacional no Brasil 2007*. Belo Horizonte.

_____ (2010). *Centro de Estatística e Informações. Déficit habitacional no Brasil 2008*. Belo Horizonte.

_____ (2011). *Centro de Estatística e Informações. Déficit habitacional no Brasil 2009*. Belo Horizonte.

_____ (2013). *Centro de Estatística e Informações. Déficit habitacional municipal no Brasil 2010*. Belo Horizonte.

_____ (2020). *Metodologia do déficit habitacional e da inadequação de domicílios no Brasil – 2016-2019*. Belo Horizonte.

_____ (2021). *Déficit Habitacional no Brasil – 2016-2019*. Belo Horizonte.

_____ (2023). *Nota Técnica: Os desafios do cálculo do déficit habitacional frente ao censo demográfico de 2022*. Belo Horizonte.

_____ (2024a). *Déficit Habitacional no Brasil 2022*. Belo Horizonte.

_____ (2024b). *Nota Técnica: As voltas que o ônus dá: alternativas para o cálculo do ônus excessivo com o aluguel urbano municipal*. Belo Horizonte.

FORCEL, P. K. B. (2023). *Sintaxe do déficit habitacional – Censo 2010* [Repositório GitHub]. GitHub. Disponível em: https://github.com/PriscilaForcel/Deficit-Habitacional/blob/main/SINTAXE%20DO%20D%C3%89FICIT%20_%202010/Sintaxe_DEFICIT.txt

GENEVOIS, M. L. B. P.; COSTA, O. V. (2001). Carência Habitacional e Déficit de Moradias: questões metodológicas. *São Paulo em Perspectiva*, v. 15, n. 1, pp. 73-84, jan. DOI: https://doi.org/10.1590/S0102-88392001000100009.

GIVISIEZ, G. H.; OLIVEIRA, E. L. de. (2013). Privacidade intradomiciliar: um estudo sobre as necessidades de ampliações em residências. *Revista Brasileira de Estudos de População,* v. 30, n. 1, pp. 199-223.

_____ (2022). "Privacidade interna das moradias e os arranjos familiares". In: SANTOS, E. C. (org.). *Ensaios e discussões sobre o déficit habitacional no Brasil*. Belo Horizonte, FJP/MDR. Disponível em: https://drive. google.com/file/d/1stJWtC0dMFtK_Zz0G0zFBjDgw2WxjL_i/view. Acesso em: out 2023.

IBGE (RJ) (2010). *Censo demográfico 2010: documentação*. Rio de Janeiro. Disponível em: https://ftp.ibge.gov. br/Censos/Censo_Demografico_2010/Resultados_Gerais_da_Amostra/Microdados/Documentacao.zip. Acesso em: 6 jul 2022.

_____ (RJ) (2012). *Censo demográfico 2010: resultados gerais da amostra.* Rio de Janeiro. Disponível em: ftp:// ftp.ibge.gov.br/Censos/Censo_Demografico_2010/Resultados_Gerais_da_Amostra/Microdados/. Acesso em: 6 jul 2022.

LACERDA, G. C.; PIMENTA, I. S.; FERREIRA, F. P. M. (2022). "O ônus excessivo com aluguel urbano e o déficit habitacional: abrindo a caixa-preta". In: SANTOS, E. C. (org.). *Ensaios e discussões sobre o déficit habitacional no Brasil*. Belo Horizonte, FJP/MDR. Disponível em: https://drive.google.com/file/ d/1stJWtC0dMFtK_Zz0G0zFBjDgw2WxjL_i/view. Acesso em: out 2023.

MARICATO, E. (1999). *Habitação e cidade*. São Paulo, Atual Editora.

_____ (2002). "As ideias fora do lugar e o lugar fora das ideias: planejamento urbano no Brasil". In: ARANTES, O.; VAINER, C.; MARICATO, E. *A cidade do pensamento único: desmanchando consensos*. Petrópolis, Vozes.

_____ (2015). *Para entender a crise urbana*. São Paulo, Expressão Popular.

MIRANDA-RIBEIRO, A. de; VIANA, R. de M.; SALIS, R. M. (2013). Déficit Habitacional no Brasil em 2007 e 2008: notas metodológicas e principais resultados. *Revista Geografias*, v. 9, n. 1, pp. 97-115. DOI: https://doi. org/10.35699/2237-549X.13356.

MIRANDA-RIBEIRO, A. D.; VIANA, R. M.; AZEVEDO, S. D. (2015). Déficit habitacional municipal em Minas Gerais. *Caderno de Geografia,* v. 25, n. 43, pp. 144-162. Disponível em: https://www.redalyc.org/articulo. oa?id=333238464010. Acesso em: maio 2023

NASCIMENTO, D. M.; BRAGA, R. C. Q. (2009). Déficit habitacional: um problema a ser resolvido ou uma lição a ser aprendida? *Risco Revista de Pesquisa em Arquitetura e Urbanismo* (online), n. 9, pp. 98-109. DOI: 10.11606/issn.1984-4506.v0i9p98-109.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS (2015). *Déficit habitacional en América Latina y el Caribe: una herramienta para el diagnóstico y el desarrollo de políticas efectivas en vivienda y hábitat.* Nova York, ONU. Disponível em: https://unhabitat.org/books/deficit-habitacional-en-america-latina-y-elcaribe/. Acesso em: jan 2024.

PASTERNAK, S. (2012). Um olhar sobre a habitação em São Paulo. *Cadernos Metrópole.* São Paulo, n. 9, pp. 81-117. Disponível em: https://revistas.pucsp.br/index.php/metropole/article/view/9206. Acesso em: dez 2024.

PASTERNAK, S.; D'OTTAVIANO, C.; BÓGUS, L. M. M.; MAGALHÃES, L. F. A.; PEDROSA, R. (2022). "Dinâmica demográfica, desigualdades socioespaciais e precariedade habitacional". In: BÓGUS, L. M. M. et al. (org.). *Reforma urbana e direito à cidade: São Paulo.* São Paulo, Observatório das Metrópoles/PUC-SP, pp. 45-62.

PRIETO, G. F. T.; LACZYNSKI, P. (2020). São Paulo à venda: ultraneoliberalismo urbano, privatização e acumulação de Capital (2017-2020). *GeoUSP Espaço e Tempo* (Online), v. 24, n. 2, pp. 1-20. DOI: https:// doi.org/10.11606/issn.2179-0892.

ROLNIK, R. (2015). *Guerra dos lugares*. São Paulo, Boitempo.

RUSSO, E. de O. (2017). "O que o indicador indica? O déficit habitacional no Brasil e as disputas em torno de sua elaboração, em meados dos anos 1990". In: CARDOSO, A. L.; JAENISCH, S. M.; ARAGÃO, T. A. *Vinte e dois anos de política habitacional no Brasil: da euforia à crise.* Rio de Janeiro, Letra Capital/Observatório das Metrópoles, cap. 1, pp. 51-83.

VILLAÇA, F. (2011). São Paulo: segregação urbana e desigualdade. *Estudos Avançados,* v. 25, n. 71.

_____ (2012). *Reflexões sobre as cidades brasileiras*. São Paulo, Studio Nobel.

Authorship contribution

Priscila Kauana Barelli Forcel: projet adminstration; formal analysis; funding acquisition; conceptualization; data curation; investigation; methodology; resources; writing–original draft; software; visualization.
Elza Luli Miyasaka: funding acquisition; conceptualization; methodology; writing–review & editing; supervision; validation.
Tiago Augusto da Cunha: formal analyasis; conceptualization; data curation; methodology; writing–review & editing; software; supervision; validation.

Declaration of conflict of interest

The authors declare that there is no conflict of interest.

Data Availability Statement

All the supporting data for the results of this study has been made available in the Institutional Repository of the Federal University of São Carlos and can be accessed at https://hdl.handle.net/20.500.14289/22766. It is also available on GitHub at the following link: https://github.com/PriscilaForcel/Deficit-Habitacional_JACAREI.

Editors: Lucia Bógus and Luiz César de Queiroz Ribeiro
Organizers of this issue: Suzana Pasternak and Luís Felipe Aires Magalhães

Translation: this article was translated from Portuguese to English by the authors themselves.