



UNIVERSITÀ
DI PAVIA

Dipartimento di Scienze Economiche e Aziendali

Corso di Laurea Magistrale in Economics, Development and Innovations

EQUITY IN ACCESS TO
HEALTHCARE IN THE ITALIAN
NATIONAL HEALTH SERVICE:
REGIONAL DISPARITIES AND
SOCIOECONOMIC
INEQUALITIES IN FORGONE
CARE

Relatore:
Chiar.mo Prof. Di Novi Cinzia

Tesi di Laurea
di Biagio Dipalma
Matr. n. 562307

Anno Accademico 2025–2026

Abstract

Equitable access to healthcare, understood as access determined by health need rather than socioeconomic position, is a core objective of the Italian National Health Service. Existing evidence on income-related horizontal inequity in Italy has largely relied on single cross-sections or earlier periods, while recent work documents pro-rich inequity in specialist and diagnostic care in 2023. This thesis extends that evidence over time by estimating a need-adjusted Erreygers concentration index for four healthcare services — blood tests, emergency department visits, specialist visits and specialist diagnostic examinations — using six consecutive annual waves (2018–2023) of the Italian National Institute of Statistics’ *Aspetti della Vita Quotidiana* survey. Estimates are produced at both the national and regional level, with sampling weights applied consistently throughout the analysis.

The results show that need standardisation materially changes the assessment of equity. Blood-test utilisation displays only a weak socioeconomic gradient in the raw data, yet the need-adjusted index is consistently positive, revealing pro-rich inequity that would be understated by the unadjusted distribution alone. Specialist visits and diagnostic examinations display the largest positive need-adjusted indices across the period, while blood tests show smaller but persistent pro-rich inequity; emergency department estimates remain small and negative, with statistical significance varying across waves. For the three programmable services, the 2020–2021 need-adjusted point estimates are generally higher than the corresponding 2019 estimates, although no formal tests of cross-wave differences are conducted. For specialist visits, decomposition of the 2019–2020 point-estimate difference shows that the age contribution becomes less negative, mainly in association with a lower estimated marginal effect of being aged 65 or over, while the prevalence of older respondents changes little.

At the regional level, the individual point estimates do not display a clear North–South ordering in the waves examined; however, the analysis does not formally test average differences between macro-areas and therefore does not establish the absence of an underlying territorial gradient. Finally, the weighted descriptive evidence on forgone care shows that cost is more frequently reported than waiting time in 2018–2019, whereas waiting time is more frequently reported in 2022–2023. Cost also shows a clearer socioeconomic gradient than waiting time. Overall, the findings indicate that pro-rich horizontal inequity in programmable healthcare utilisation is a persistent feature of the Italian SSN context over 2018–2023, while the magnitude and composition of the estimated inequity vary across waves.

Contents

Introduction	1
1 The Italian National Health Service and the Concept of Equity in Access to Healthcare	5
1.1 The Origins and Founding Principles of the Italian National Health Service	5
1.2 Regionalisation and Decentralisation of the National Health Service .	7
1.3 Financing of the National Health Service and the Allocation of Resources across Regions	8
1.4 The Essential Levels of Care and the Nuovo Sistema di Garanzia . . .	10
1.5 Regional Disparities: Financial Recovery Plans and Interregional Patient Mobility	12
1.6 Horizontal and Vertical Equity: A Conceptual Framework	13
1.7 Literature Review on Equity of Access to Healthcare in Italy and the Positioning of the Present Thesis	15
2 Data and Methods	23
2.1 Data	23
2.1.1 The ISTAT Survey on Aspects of Daily Life	23
2.1.2 Waves and Access Conditions	24
2.1.3 Construction of the Estimation Sample	25
2.2 Measuring Inequality: the Concentration Index	26
2.2.1 Definition and Interpretation	26
2.2.2 From the Standard Index to the Erreygers Index	27
2.2.3 Utilisation Outcomes	28
2.2.4 The Socioeconomic Status Index	29
2.2.5 Need and Non-Need Variables	30
2.2.6 Forgone Care	30
2.3 Harmonisation Across Waves	31

2.3.1	Chronic Morbidity	31
2.3.2	The 2022 Change in the Coding of Forgone Care	33
2.3.3	COVID-19 Coding and the Residual Reason Category	34
2.4	Measuring Inequity: Need Standardisation and the ENA Index	35
2.4.1	Horizontal Equity	35
2.4.2	Indirect Standardisation	36
2.4.3	The Need-Adjusted Index	36
2.5	Decomposition of the Index	37
2.5.1	Purpose	37
2.5.2	Reading a Contribution	38
2.5.3	Covariates Decomposed and the Residual	38
2.5.4	Decomposing Differences Across Waves	38
2.6	Estimation and Inference	39
2.6.1	Design of the Analysis	39
2.6.2	The Socioeconomic Ranking	39
2.6.3	Standard Errors	40
2.6.4	Reliability of the Regional Estimates	40
2.6.5	Weighting	41
2.6.6	Software	42
2.7	Forgone Care: Descriptive Analysis	42
2.7.1	Denominators and Breakdowns	42
3	Results	43
3.1	Sample and Descriptive Statistics	43
3.1.1	Utilisation of Healthcare Services	43
3.1.2	Need and Socioeconomic Variables	44
3.1.3	Utilisation by Socioeconomic Status Quintile	44
3.1.4	Utilisation by Age Class	45
3.1.5	Utilisation by Macro-Area	46
3.1.6	Sensitivity to the 2023 Sample Definition	47
3.2	National Indices: EI and ENA	48
3.2.1	Levels and Ordering Across Services	48
3.2.2	The Gap Between $E(y)$ and ENA	49
3.2.3	Trajectory, 2018–2023	49
3.2.4	The Emergency Department	52
3.3	Decomposition of the ENA Index	52
3.3.1	What Generates the Level of ENA	55
3.3.2	Accounting for the 2020 Point-Estimate Difference	58

3.3.3	The Emergency Department: a Different Composition	60
3.4	Regional Estimates	61
3.4.1	Precision at the Regional Level	61
3.4.2	Regional Rankings and the North–South Question	61
3.4.3	Stability and Volatility Across Waves	63
3.4.4	Consistency with the National Pattern	64
3.4.5	Blood Tests and the Emergency Department	64
3.5	Forgone Care	65
3.5.1	Rates of Forgone Care	65
3.5.2	Forgone Care by Quintile, Age and Macro-Area	70
3.5.3	COVID-19 Coding and the Residual Reason Category	72
3.5.4	The Changing Relative Importance of Cost and Waiting Time	73
3.5.5	Cost and Waiting Time by Macro-Area	74
3.5.6	The Socioeconomic Gradient in Reported Barriers	77
3.5.7	Summary	77
4	Discussion	78
4.1	Summary of Main Findings	78
4.2	The Raw Index Conceals Inequity: the Case of Blood Tests	79
4.3	The Higher 2020 Inequity Point Estimates and Their Subsequent Pat- tern	80
4.3.1	Timing and Magnitude of the Point-Estimate Difference	80
4.3.2	Changes in the Age and Socioeconomic Contributions	81
4.3.3	Post-2019 Patterns and Reported Barriers	81
4.4	The Emergency Department: A Small and Unstable Pattern	82
4.5	Regional Estimates and the North–South Question	84
4.6	Comparison with the Existing Literature	85
4.6.1	Levels Relative to Di Novi (2025)	85
4.6.2	Positioning Relative to Di Novi, Piacenza, Robone and Turati (2019)	86
4.6.3	What This Thesis Adds	86
4.7	Limitations	87
4.7.1	Precision of the Estimates	87
4.7.2	Construction of the Data	88
4.8	Policy Implications	89
	Conclusion	91

Bibliography	95
List of Legislation	98
Appendix	100

Introduction

Equitable access to healthcare is one of the founding commitments of the Italian National Health Service (*Servizio Sanitario Nazionale*, SSN). When the SSN replaced the pre-existing system of occupational sickness funds in 1978, it did so explicitly to sever the link between an individual's entitlement to care and their occupational or social position, establishing instead a universal right to healthcare financed through general taxation. More than four decades later, whether this commitment is met in practice remains an open empirical question. The evidence reviewed in Chapter 1 indicates that the SSN continues to operate within a framework of comparatively constrained resources, unevenly distributed across Regions with markedly different fiscal capacities, and that the essential levels of care the State is constitutionally required to guarantee uniformly are met unevenly, with sub-threshold performance concentrated, though not exclusively, in the Regions of the Mezzogiorno. Whether these institutional and financial disparities translate into unequal access to specific healthcare services, once the greater healthcare need of some population groups is properly accounted for, is the question this thesis sets out to answer.

Motivation

The empirical literature on horizontal equity of access to healthcare in Italy, reviewed in full in Section 1.7, is small but internally consistent. Masseria and Giannoni (2010) document pro-rich inequity in general practitioner, specialist and emergency care using data from 2000; Glorioso and Subramanian (2014) offer a more refined account based on 2005 data, showing that inequity in specialist and diagnostic care is concentrated among individuals in comparatively good health; and Di Novi et al. (2019) examine, over the period 1994–2007, how fiscal decentralisation affected the distribution of health outcomes across and within Regions. Most recently, Di Novi (2025) applies the need-adjusted Erreygers concentration index (ENA) to the same four healthcare services considered in this thesis — blood tests, emergency department visits, specialist visits and specialist diagnostic examinations — using

the 2023 wave of the ISTAT *Aspetti della Vita Quotidiana* (AVQ) survey. That study reports significant pro-rich inequity in specialist visits, diagnostic examinations and blood tests, while emergency department access is approximately equitable at the national level.

Three of these four contributions conclude their empirical analysis by 2007; the fourth reports estimates for a single annual wave. None documents annual need-adjusted inequity estimates across the years immediately preceding the COVID-19 pandemic, the pandemic itself, and the subsequent period through 2023. The existing literature is therefore not positioned to show how the point estimates and their decomposition vary across this recent window. This is the gap addressed by the present thesis: rather than treating 2023 as an isolated cross-section, it places that wave within a six-year sequence of repeated cross-sections and examines the descriptive temporal pattern of horizontal inequity in the Italian SSN context.

Aims and Contribution

This thesis extends the need-adjusted Erreygers methodology of Di Novi (2025) along the temporal dimension. Where that contribution estimates the ENA index on a single annual wave, the analysis presented here estimates the same index for the same four healthcare services, and with the same construction of socioeconomic status, on six consecutive repeated cross-sections of the AVQ survey covering 2018 to 2023, at both the national and regional level. This design places the 2023 cross-section within a sequence of annual estimates and allows the magnitude and composition of the ENA point estimates to be compared descriptively across waves. Because the survey consists of independent repeated cross-sections, these comparisons do not identify individual-level change, and the thesis does not formally test all year-to-year differences. The same multi-wave design is also used to compare the reported composition of forgone-care barriers across internally comparable periods and to examine the heterogeneity of regional ENA point estimates over time.

Three aspects of the empirical results are emphasised. First, need standardisation materially changes the assessment of equity for blood tests. Their raw utilisation gradient across socioeconomic groups is weak, whereas the need-adjusted index is consistently positive, showing that an apparently flat raw distribution can coexist with pro-rich inequity once differential health need is taken into account. Specialist visits and diagnostic examinations form the group with the largest positive ENA estimates, but the small differences between those two services are not interpreted as a substantively meaningful ranking.

Second, the decomposition helps clarify the 2019–2020 difference in point estimates for specialist visits. The age contribution becomes less negative, from approximately -0.034 to -0.022 . Examining its components shows that this arithmetic difference is associated mainly with a lower estimated marginal effect of being aged 65 or over, while the weighted prevalence of older respondents changes little and the age concentration index becomes only moderately less negative. Since no formal test of the cross-wave difference in these components is conducted, the result is interpreted as a decomposition of point estimates rather than as evidence of a statistically established structural change in the role of age.

Third, the weighted descriptive evidence on forgone care shows a change in the relative ordering of the two main reported barriers across the more internally comparable endpoint periods: cost is more frequently reported in 2018–2019, whereas waiting time is more frequently reported in 2022–2023. Cost also displays a clearer socioeconomic gradient than waiting time. At the regional level, the individual ENA point estimates do not display a clear North–South ordering in the waves reported; however, no formal comparison of average inequity across macro-areas is conducted, so the analysis does not establish that an underlying territorial gradient is absent.

Structure of the Thesis

The thesis is organised in four chapters, in addition to this introduction.

Chapter 1 sets out the institutional and financial context within which access to the SSN’s services must be understood — the Service’s founding principles, its progressive regionalisation, the architecture of its financing, the persistent regional gradient in essential levels of care, and the mechanisms of financial recovery plans and interregional patient mobility through which regional disparities continue to manifest themselves — before introducing the concept of horizontal equity that constitutes the analytical core of the thesis and reviewing, in full, the existing literature on equity of access to healthcare in Italy against which the present contribution is positioned.

Chapter 2 presents the data and the empirical strategy. It describes the AVQ survey and the construction of the six annual estimation samples; defines the four utilisation outcomes, the socioeconomic status index, and the need and non-need variables entering the analysis; sets out how these variables are harmonised across six waves of a questionnaire that was revised during the period, including a substantive change in 2022 to the coding of forgone care; introduces the Erreygers concentration index and its need-adjusted counterpart, the ENA index, together with

the criterion of horizontal equity it is designed to measure; explains how the index is decomposed into the contributions of socioeconomic status, need and non-need characteristics; and details the estimation design, including the bootstrap procedure used for inference and the treatment of sampling weights.

Chapter 3 reports the results. It documents utilisation of the four healthcare services and its relationship to socioeconomic position, age and geography; presents the annual national point estimates of the standard and need-adjusted concentration indices between 2018 and 2023; decomposes the ENA index into its constituent contributions and examines how those components differ across waves; reports regional estimates for specialist visits and diagnostic examinations, together with an assessment of their statistical reliability; and presents six waves of weighted descriptive evidence on forgone care, including the reasons reported for it and how those reasons vary by socioeconomic quintile, age class and macro-area.

Chapter 4 interprets these results, situates them against the literature reviewed in Chapter 1, and sets out the limitations of the analysis and its implications for policy. It examines why need standardisation matters for services such as blood tests, whose raw utilisation gradient is weak; discusses the component-level accounting of the 2019–2020 point-estimate difference for specialist visits; considers the emergency department as a service whose ENA remains small and negative while its statistical significance varies across waves; and discusses the regional estimates without treating the absence of a clear North–South ordering in the reported point estimates as proof that a territorial gradient does not exist.

Chapter 1

The Italian National Health Service and the Concept of Equity in Access to Healthcare

1.1 The Origins and Founding Principles of the Italian National Health Service

Before 1978, health protection in Italy was structured according to a Bismarckian social health insurance model, with coverage organised around occupational sickness funds rather than universal entitlement (Toth, 2016). By the mid-1970s, more than 300 separate sickness funds (*casse mutue*) operated in the country, with assignment determined exclusively by occupation and both contribution rates and benefits varying considerably from one fund to another (Toth, 2016). Even at their broadest reach, these funds covered no more than 93 per cent of the population, leaving over three million Italians without any healthcare coverage at all, in a system whose main limitations resided in “organisational fragmentation and disparity in treatment” (Toth, 2016).

This combination of fragmentation and unequal treatment across occupational categories meant that, in the decades preceding 1978, access to healthcare in Italy depended not on health need but on one’s position within a stratified system of occupational entitlement — a configuration structurally at odds with the principle of horizontal equity that constitutes the analytical core of this thesis and that is formally introduced in Section 1.6. It was against this backdrop, and amid a broader wave of social and political mobilisation that swept through Italy during the 1970s, that Parliament approved a radical reform of the entire healthcare financing and

provision system (Giorgi, 2023).

The Italian National Health Service (*Servizio Sanitario Nazionale*, SSN) was established by Law No. 833 of 23 December 1978, which replaced the pre-existing system of occupational sickness funds with a Beveridge-type model financed primarily through general taxation (Toth, 2016). Article 1 of the law explicitly referred to Article 32 of the Italian Constitution, defining the SSN as the set of functions, structures, services and activities aimed at promoting, maintaining and restoring the physical and mental health of the entire population, with no distinction based on individual or social conditions, and in a manner ensuring equality of citizens with respect to the service (Giorgi, 2023). Five guiding principles can be identified around which the reform was structured (Toth, 2016): universality of coverage, extended to the entire resident population regardless of occupational status; comprehensiveness of the benefit package; equity of financing, whereby contributions were to be proportional to individual means rather than tied to occupational category; unity of administration, bringing the previously fragmented public and private actors under a single public entity; and, as a synthesis of the preceding four, equal treatment for all users, irrespective of age, sex, income, or region of residence. Taken together, these principles marked a decisive break with the logic of the pre-1978 system: health-care access was no longer to be mediated by occupational status, but was instead reconceived as a universal social right (Giorgi, 2023; Toth, 2016).

To implement this vision, the reform assigned powers and responsibilities to three institutional levels — the State, the Regions, and the municipalities — with the State retaining overall planning and financing authority, the Regions responsible for programming the service locally, and municipalities managing the newly created Local Health Units (Bifulco & Neri, 2022). In practice, however, this three-tier structure soon proved a source of persistent jurisdictional conflict: responsibility for planning and financing was not clearly divided between the national and regional levels, Regions increasingly considered the resources transferred by the central government insufficient, and the management of Local Health Units suffered from excessive politicisation (Ferré et al., 2014). It was in response to these mounting tensions that the Italian government introduced, in the early 1990s, a further reform of the system, discussed in the following section.

1.2 Regionalisation and Decentralisation of the National Health Service

The tensions that had accumulated under the original 1978 architecture prompted the Italian government to undertake a substantial reform of the National Health Service in the early 1990s (Ferré et al., 2014). This reform, enacted through Legislative Decrees No. 502 of 30 December 1992 and No. 517 of 7 December 1993, transformed the Local Health Units into autonomous entities with managerial and financial independence, and introduced a quasi-market logic within the otherwise publicly financed system: a formal separation between the purchaser function, exercised by the Local Health Authorities, and the provider function, exercised by public and private accredited facilities alike, reimbursed on the basis of prospective Diagnosis-Related Group tariffs rather than retrospective costs (Brenna, 2011; Ferré et al., 2014). This shift was intended to increase provider accountability, contain expenditure growth, and improve service quality through competition among a plurality of providers (Brenna, 2011).

The practical implementation of this reform was far from uniform across the Italian territory. Lombardy, in particular, made full use of the discretion it afforded, enacting Regional Law No. 31 of 1997 to establish what has come to be known as the “Lombardy model”: strict purchaser-provider separation, competition between public and private accredited providers, and full patient freedom of choice (Brenna, 2011). Other Regions pursued markedly different paths, from Emilia-Romagna’s more integrated model to a substantially larger role for private accredited beds in Regions such as Lazio (Brenna & Spandonaro, 2015). Rather than producing a single, homogeneous model of managed competition, the 1992–93 reform thus created a regulatory space within which Regions constructed substantially different health systems — a divergence further reinforced by the constitutional reform of 2001.

In 2001, Constitutional Law No. 3 amended Title V of the Constitution, expanding regional legislative autonomy in health matters while leaving the definition of the essential levels of care (LEA) — to be guaranteed uniformly across the country — under national competence (Ferré et al., 2014). The reform also reaffirmed the founding principles of universalism, comprehensiveness and public funding, while recommending that Regions replace open competition with a model of “contractual planning” between purchasers and providers (Brenna, 2011). Its scope and ambiguity, however, left commentators concerned that a further strengthening of regionalisation, superimposed on an already uneven pattern of regional development,

risked fragmenting the National Health Service into a set of substantially distinct regional systems (Ferré et al., 2014).

This same period saw the principle of fiscal federalism progressively enshrined in law, from the replacement of the National Health Fund with a National Solidarity Fund in 2000, through the Title V reform itself, to the framework law of 2009 that specified the criteria for redistributing local taxes across levels of government (Brenna, 2011; Ferré et al., 2014); in practice, however, regional healthcare financing has continued to combine regional taxes, revenue-sharing arrangements, central transfers and equalisation mechanisms, rather than resting on own-source resources alone. Its implementation, moreover, has remained incomplete: regional financing continues to rely, to a considerable extent, on centrally determined transfers and equalisation mechanisms rather than on genuinely autonomous regional taxation — an unresolved tension that is the starting point for the analysis of healthcare financing developed in the following section.

1.3 Financing of the National Health Service and the Allocation of Resources across Regions

Financing of the Italian National Health Service relies on regional taxes, including the regional tax on productive activities (IRAP) and the regional surcharge on personal income tax, introduced in 1998 to replace direct transfers from the national budget, together with VAT revenue-sharing and equalisation mechanisms for Regions unable to raise sufficient revenue on their own (Di Novi et al., 2019; Ferré et al., 2014). This reform was explicitly conceived as an instrument of fiscal decentralisation, intended to increase local accountability and thereby improve the efficiency of healthcare spending (Di Novi et al., 2019). In practice, however, a Region’s capacity to raise revenue through these instruments depends directly on its underlying economic base, so that Regions with weaker fiscal capacity have often been required to apply the highest permissible tax rates simply to approach fiscal balance (Ferré et al., 2014).

This financing architecture has operated for over two decades within a context of expenditure containment rather than expansion. Between 2008 and 2019, current public health expenditure as a share of GDP in Italy remained virtually unchanged, whereas it rose substantially in France, Germany and the United Kingdom over the same period (Bifulco & Neri, 2022). This divergence indicates that Italy’s comparatively low position relative to other major European economies is not a recent phenomenon but the outcome of more than a decade of sustained relative under-

financing. Since 2001, the Ministry of Health has conducted an annual evaluation of whether each Region effectively supplies the essential levels of care (LEA) constitutionally guaranteed to its residents; this monitoring has consistently shown that Regions of southern Italy tend to provide lower standards of healthcare largely independently of the funding formally provided by the State (Signorelli et al., 2020), indicating that fiscal disparities alone do not fully explain the persistence of unequal service provision — a pattern examined in detail in the following section.

In 2023, Italy’s current health expenditure amounted to 8.4 per cent of GDP, 1.6 percentage points below the EU average, with per capita spending around 19 per cent below the EU average once adjusted for purchasing power (OECD & European Observatory on Health Systems and Policies, 2025). This shortfall is attributable primarily to a comparatively limited public component, only partially offset by a private component — composed, to the extent of roughly 90 per cent, of direct out-of-pocket payments — that exceeds the European average (OECD & European Observatory on Health Systems and Policies, 2025). The COVID-19 pandemic introduced a marked, if temporary, discontinuity in this trajectory: real health expenditure grew by more than 8 per cent between 2019 and 2021, before moderating to only around 3 per cent above its 2019 level by 2023 (OECD & European Observatory on Health Systems and Policies, 2025). The accessibility consequences of this constrained spending remain tangible in the most recent data: in 2023, long waiting lists were identified as the principal barrier to care in Italy, with over 7 per cent of the population reporting that they had forgone needed medical care for this reason, rising to more than 2.5 times the general population rate among adults at risk of poverty in 2024 (OECD & European Observatory on Health Systems and Policies, 2025).

Taken together, the evidence presented in this section indicates that the Italian National Health Service operates within a framework of comparatively constrained aggregate resources, unevenly distributed across Regions with markedly different fiscal capacities. Whether this translates into a genuinely uniform standard of care is a question that admits of a direct empirical answer, provided by the Ministry of Health’s own monitoring of the essential levels of care through the *Nuovo Sistema di Garanzia*, to which the following section is devoted.

1.4 The Essential Levels of Care and the Nuovo Sistema di Garanzia

The essential levels of care (*Livelli Essenziali di Assistenza*, LEA) constitute the set of healthcare services that the National Health Service is required to guarantee uniformly to all residents, irrespective of their Region of residence (Signorelli et al., 2020). Ministerial monitoring of this principle, in place since 2001, was substantially reformed with the Ministerial Decree of 12 March 2019, which established the *Nuovo Sistema di Garanzia* (NSG) as a more analytically sophisticated successor to the earlier “griglia LEA” scoring system (Lispi et al., 2025).

The NSG evaluates regional performance across three macro-areas of care — collective prevention, district-level care, and hospital care — through a subset of CORE indicators, each converted onto a common 0–100 scale with a sufficiency threshold of 60 (Lispi et al., 2025). Crucially, a Region’s guarantee of LEA provision is considered achieved only when this threshold is met separately in each of the three macro-areas, with no compensation permitted across them — a design intended to prevent a strong performance in one domain from concealing severe shortfalls in another (Lispi et al., 2025). Compliance with the CORE indicators is, moreover, directly linked to Regions’ access to a supplementary share of financing under Article 2, paragraph 68, of Law No. 191/2009, making LEA monitoring an active mechanism connecting service delivery to financial resources, not merely a descriptive exercise (Lispi et al., 2025).

As Figure 1.1 shows, thirteen of the twenty-one regional health systems met the sufficiency threshold in all three macro-areas in 2023, while the remaining sub-threshold Regions are heavily concentrated in the Mezzogiorno and among smaller Alpine territories (Lispi et al., 2025). This pattern is not, however, a simple North–South dichotomy: Campania, Puglia and Sardinia achieve full compliance despite belonging to the Mezzogiorno, while Liguria and the Province of Bolzano fall short despite being northern territories (Lispi et al., 2025). This regional gradient is corroborated by independent evidence: the National Outcomes Programme coordinated by AGENAS systematically identifies central and northern Regions as superior performers relative to the South (Toth, 2016), a pattern mirrored in citizen perceptions, with 83 per cent of southern residents considering their regional health service inadequate compared with approximately 30 per cent in the North (Toth, 2016).

The persistence of this gradient is closely intertwined with the decentralised and uncoordinated accreditation of private providers discussed in Section 1.2: several northern and central Regions accredited well-resourced private facilities from an

Legend

- Above threshold in all 3 macro-areas
- Below threshold in 2 macro-areas
- Below threshold in 1 macro-area

11

early stage, while a comparable process took place far more slowly across most southern Regions (Brenna & Spandonaro, 2015). This asymmetry in the availability of care is a central driver of the interregional patient mobility examined in the following section, together with the related mechanism of financial recovery plans (*Piani di Rientro*).

1.5 Regional Disparities: Financial Recovery Plans and Interregional Patient Mobility

Regional disparities in fiscal capacity and service provision found their most acute expression in the structural healthcare deficits accumulated by a subset of Regions over the course of the 2000s. Beginning in 2007, the central government required these Regions to adopt formal financial recovery plans (*Piani di Rientro*), conditioning state coverage of accumulated deficits on multi-year programmes of expenditure containment (Ferré et al., 2014). In aggregate terms, this regime achieved a marked reduction in overspending — from €4.48 billion in 2006 to €1.04 billion by 2012 — though this consolidation coincided with, and should be read against, the broader decade-long pattern of relative under-financing already documented in Section 1.3 (Bifulco & Neri, 2022; Ferré et al., 2014).

A further manifestation of persistent regional disparities concerns interregional patient mobility. Italian citizens retain the right to seek treatment at any public or private accredited hospital in the country, with the cost of care settled between the Region of residence and the Region of treatment (Signorelli et al., 2020). This mobility is considerable in scale — approximately 700,000 patients migrate between Regions each year for hospital care — and follows a marked geographical pattern, with patients from southern Regions such as Calabria, Campania, Puglia and Sicily disproportionately seeking treatment in Lombardy, Emilia-Romagna and Tuscany (Brenna & Spandonaro, 2015; Signorelli et al., 2020). This asymmetry has intensified rather than attenuated over time, and is corroborated by independent evidence on regional healthcare performance discussed in Section 1.4, suggesting that patients migrate towards Regions that are independently assessed as offering better care (Signorelli et al., 2020; Toth, 2016).

This pattern is rooted, at least in part, in the decentralised and uncoordinated accreditation of private providers: since 1999, hospital accreditation has followed a four-stage procedure carried out autonomously at regional level, without central supervision ensuring an equitable distribution of accredited capacity across the country (Brenna & Spandonaro, 2015). The resulting dynamics carry a twofold implication

for the National Health Service. From an efficiency standpoint, interregional mobility does not resolve, but merely displaces, the underlying imbalance between healthcare supply and demand, since Regions that lose patients continue to bear the fixed costs of their own infrastructure while also financing the treatment their residents receive elsewhere (Brenna & Spandonaro, 2015). From an equity standpoint, exercising the formal right to free choice of treatment location requires resources — for travel and accommodation — that are not equally available to all patients, meaning that the practical benefit of interregional mobility accrues disproportionately to the better-off (Brenna & Spandonaro, 2015).

Taken together, the evidence presented in this chapter — spanning the founding principles of the National Health Service, its progressive regionalisation, the architecture of its financing, the persistent gradient in essential levels of care, and the mechanisms through which regional disparities continue to manifest themselves — establishes the institutional and empirical backdrop against which the concept of horizontal equity in access to healthcare, introduced in the following section, must be understood.

1.6 Horizontal and Vertical Equity: A Conceptual Framework

Having documented, in the preceding sections, the institutional architecture, financing arrangements, and persistent regional disparities that characterise the Italian National Health Service, this section turns to the conceptual framework required to assess whether access to its services is distributed equitably. Equitable access to healthcare constitutes a core objective of the Italian healthcare system (Masseria & Giannoni, 2010), yet the notion of equity itself requires careful specification before it can be operationalised empirically.

The relevant standard in this literature is horizontal equity, understood as the principle that access to healthcare should depend exclusively on an individual's health need, with socioeconomic factors unrelated to that need exerting no independent influence on utilisation (Masseria & Giannoni, 2010). Income-related equity in healthcare use is typically assessed through the standardisation-by-regression methodology, whereby the observed relationship between utilisation and need is used to construct an estimate of the utilisation that each individual would be expected to receive on the basis of need alone (Masseria & Giannoni, 2010).

This approach gives rise to the horizontal inequity index (HI), formally defined as the difference between income-related inequality in actual healthcare use, denoted

C_M , and income-related inequality in need-expected use, denoted C_N (Masseria & Giannoni, 2010). The sign of each component carries a specific interpretation: C_M is positive when the better-off make greater use of care than the worse-off, and negative in the opposite case; C_N is negative when the distribution of need itself favours the worse-off, and positive when it favours the better-off (Masseria & Giannoni, 2010). No inequity in access to care is deemed to be present whenever C_M and C_N are equal, or whenever their difference fails to reach statistical significance; conversely, a positive value of HI indicates inequity favouring the better-off, while a negative value indicates inequity favouring the worse-off (Masseria & Giannoni, 2010).

The practical implementation of this conceptual apparatus depends critically on the choice of a summary measure capable of expressing income-related inequality in a manner suitable for utilisation variables that are frequently binary in nature. The measurement of income-related inequality in healthcare utilisation is conventionally carried out through the concentration index, a summary statistic that captures the extent to which a given outcome variable is disproportionately concentrated among richer or poorer individuals. When applied to variables measuring health, healthcare utilisation, or health-related behaviour, however, the standard concentration index presents a specific limitation: its bounds depend on the mean of the variable under consideration, so that the range the index can attain is itself a function of the average level of utilisation. This complicates comparisons across populations, regions, or time periods characterised by different average levels of that variable (Di Novi et al., 2019).

To address this limitation, the literature has adopted a corrected version of the index, proposed by Erreygers (2009) and generally referred to as the Erreygers index. The Erreygers correction yields fixed bounds for bounded variables, thereby overcoming the mean-dependent bounds of the standard concentration index (Erreygers, 2009), which makes it possible to compare populations — such as different Regions — with substantially different average levels of utilisation (Di Novi et al., 2019). A further methodological advantage of the Erreygers index concerns its behaviour when applied to a health-related outcome and to its logical complement. Whereas the standard concentration index may yield conflicting conclusions when applied separately to a measure of health and to the corresponding measure of ill-health, the Erreygers index satisfies what is known as the “mirror property”: inequalities detected in a measure of health necessarily mirror those detected in the corresponding measure of ill-health, ensuring internal consistency between the two perspectives (Di Novi et al., 2019). These two properties — bounds that remain fixed regardless of the mean of the outcome, and the mirror property — make the Erreygers

index a particularly suitable instrument for the analysis of income-related inequality in binary healthcare utilisation variables of the kind considered in this thesis, and constitute the methodological basis for the need-adjusted horizontal inequity index applied in the empirical chapters that follow.

The conceptual apparatus set out above — horizontal equity, need standardisation, and the Erreygers-corrected concentration index — has been applied to the Italian National Health Service by a small but internally consistent body of empirical work, reviewed in full in the following section. Two features of that literature are worth anticipating here, because they connect directly to the institutional evidence presented in Sections 1.4 and 1.5. The first is a services-based hierarchy of inequity: hospital inpatient care is broadly equitable, its utilisation depending predominantly on need rather than on socioeconomic status, whereas specialist and emergency care display the most pronounced departures from equity (Ferré et al., 2014). The second is that these departures are consistently traced back to inter-regional variation alongside income and education, in a setting where physicians, hospital beds and outpatient facilities are disproportionately concentrated in the northern and central Regions, and where emergency care is, in some Regions, used as a surrogate for specialist care (Ferré et al., 2014). The following section reviews this literature in detail, situating the present thesis’s extension of the analysis to six consecutive repeated cross-sections spanning 2018–2023 within the broader trajectory of Italian research on equity of access to healthcare.

1.7 Literature Review on Equity of Access to Healthcare in Italy and the Positioning of the Present Thesis

The empirical literature applying the conceptual and methodological framework outlined in the preceding section to the Italian National Health Service originates with the work of Masseria and Giannoni (2010), whose study represents the first systematic estimation of income-related horizontal inequity across multiple healthcare services in Italy. Combining the Multiscopo survey of 2000 with the European Community Household Panel survey for Italy, the authors estimated income-related inequity indices for the probability of visiting a general practitioner, receiving specialist care, undergoing hospitalisation, and accessing emergency care, employing a range of need indicators including chronic conditions, self-reported disability, age and gender (Masseria & Giannoni, 2010).

The results revealed a markedly uneven pattern across services. Significant pro-rich income-related inequity was found for general practitioner visits, specialist care and emergency care, whereas no significant inequity emerged for inpatient care (Masseria & Giannoni, 2010). This inequity was attributed primarily to income and to inter-regional variation, with education and insurance coverage playing an additional role in the case of specialist care (Masseria & Giannoni, 2010). Specialist visits, in particular, were found to be associated with a comparatively large need-standardised income gradient, with need itself estimated to reduce inequality in utilisation by as much as 23 per cent, only partially offsetting the pro-rich pattern that remained once need had been accounted for (Masseria & Giannoni, 2010). At the same time, considerable regional heterogeneity emerged independently of income, with residents of southern Regions found to be systematically less likely to visit a specialist than residents of the rest of the country (Masseria & Giannoni, 2010).

A further, distinctive contribution of the study lies in its disease-based approach: horizontal inequity indices were computed separately for individuals reporting one of four chronic conditions — hypertension, arthritis, tumour, and heart disease — in addition to the estimates obtained for the adult population as a whole (Masseria & Giannoni, 2010). This disaggregated analysis confirmed statistically significant pro-rich inequity in the probability of specialist care for three of the four conditions examined, while revealing pro-poor inequity in access to general practitioner care across all four conditions considered (Masseria & Giannoni, 2010). On the basis of these findings, the decomposition suggested that reducing regional variation could substantially reduce the observed pro-rich inequity documented for general practitioner, specialist and emergency care (Masseria & Giannoni, 2010).

A subsequent and considerably more comprehensive contribution is provided by Glorioso and Subramanian (2014), who set out explicitly to clarify the mixed conclusions reached by earlier studies on horizontal equity in the Italian National Health Service. Drawing on the 2005 wave of the Italian National Institute of Statistics' Health Conditions and Health Care Utilisation Survey — a nationally representative sample of 103,651 adults — the authors employed three-level multilevel logistic regression models, with individuals nested within families and families nested within Regions, to estimate how the probability of utilising five distinct healthcare services varies with socioeconomic status among individuals reporting equal health need (Glorioso & Subramanian, 2014). The five services considered were general practitioner visits, specialist visits, basic medical tests, outpatient diagnostic services, and hospitalisation, thereby extending the set of indicators typically examined in

the previous literature to include preventive and diagnostic services (Glorioso & Subramanian, 2014).

The results reveal a differentiated pattern of inequity across services. The use of primary care was found to be inequitable in favour of the less well-off, with the predicted probability of a general practitioner visit falling from 15.1 per cent among individuals in the lowest socioeconomic position considered to 11.9 per cent among those in the highest, corresponding to an inequity ratio of 0.78 (Glorioso & Subramanian, 2014). Specialist care, by contrast, exhibited substantial inequity in favour of the well-off, with an inequity ratio of 1.40, indicating a predicted probability of a specialist visit some 40 per cent higher among the most advantaged individuals than among the least advantaged (Glorioso & Subramanian, 2014). A comparable, though more moderate, pro-rich pattern emerged for basic medical tests (inequity ratio of 1.10) and for outpatient diagnostic services (inequity ratio of 1.24), whereas hospitalisation was found to be essentially equitable, with an inequity ratio close to unity and not statistically distinguishable from an absence of inequity (Glorioso & Subramanian, 2014).

The principal methodological innovation of the study lies in its stratified analysis of inequity by level of self-rated health status. This analysis reveals that the degree of pro-rich inequity in specialist care, basic medical tests and diagnostic services tends to diminish, and in some cases to approach statistical non-significance, as health status worsens and need becomes more apparent, whereas the highest levels of inequity are concentrated precisely among individuals in good health (Glorioso & Subramanian, 2014). The authors interpret this pattern as evidence that the benefits of prevention and early diagnosis are unevenly distributed in favour of the well-off, since individuals of lower socioeconomic status appear to postpone the use of preventive and diagnostic services until the onset of more evident need (Glorioso & Subramanian, 2014).

In accounting for these results, the authors point to the flat-rate copayments required for specialist visits and diagnostic services, the comparatively lower quality of care and longer waiting lists affecting certain sectors of the National Health Service, particularly in the southern Regions, and the resulting propensity of higher-income individuals to turn to more accessible private provision (Glorioso & Subramanian, 2014). The pro-poor inequity identified in general practitioner care, meanwhile, is interpreted as a possible substitution effect: since the predicted-probability curve for general practitioner visits closely mirrors that for specialist visits, with one rising as the other falls, individuals of higher socioeconomic status may simply be substituting relatively costly specialist consultations for less expensive general practitioner

visits, so that access to outpatient care of some kind remains broadly equitable even as the composition of that care diverges by socioeconomic status (Glorioso & Subramanian, 2014).

A third strand of the Italian literature relevant to this thesis shifts the focus from the measurement of income-related inequity in utilisation to the institutional determinants of regional health disparities, and specifically to the consequences of fiscal decentralisation. Di Novi et al. (2019) study the 1998 reform that granted Italian Regions greater fiscal autonomy — substituting central government transfers with two new sources of autonomous regional revenue, discussed in Section 1.3 — using a quasi-experimental design applied to individual-level Multiscopo data for the period 1994–2007 (Di Novi et al., 2019). Identification does not rest on the timing of the reform alone: the authors exploit cross-regional variation in per capita GDP — and hence in the size of the tax base underpinning the two new sources of own revenue — as a continuous measure of a Region’s exposure to the reform, interacting it with a post-1998 indicator in a specification that includes Region and year fixed effects and standard errors clustered at the regional level. The common-trend assumption underlying this difference-in-differences design is assessed by augmenting the model with leads of the treatment, which are found not to be significantly associated with the outcome (Di Novi et al., 2019). It is also worth noting that the outcome of interest is not self-assessed health as such, but a median-based index of inequality in self-assessed health, computed at the Region-year level.

The study’s central finding is that fiscal decentralisation did not affect disparities in self-assessed health between Regions, but did contribute to a significant containment of disparities within Regions, in a period during which within-Region health inequalities were otherwise on the rise (Di Novi et al., 2019). According to the authors’ estimates, the impact of decentralisation on the within-Region inequality index amounted, on average, to approximately three times its within-Region standard deviation (Di Novi et al., 2019). Critically, this effect was found to be considerably stronger in richer Regions than in poorer ones, a result the authors attribute to the fact that greater fiscal autonomy only translates into a genuine increase in local accountability where the underlying tax base is sufficiently large to allow politicians real discretion over the composition and targeting of local health spending; in poorer Regions, by contrast, healthcare funding continues to depend heavily on transfers from the central government, limiting the practical scope for improved local governance (Di Novi et al., 2019).

Consistent with this interpretation, the authors’ analysis by educational subgroup shows that, following the reform, individuals with low and medium educa-

tional attainment residing in richer Regions were both more likely to report improved health and more likely to make use of preventive and home-care services, while reducing their reliance on inpatient and emergency care — a pattern the authors read as evidence that more accountable local politicians in richer Regions were able to supply more appropriate and better-targeted services to their populations (Di Novi et al., 2019). No comparable improvement was observed among the corresponding sub-groups in poorer Regions.

The most direct antecedent of the present thesis, and the contribution from which it takes its point of departure, is Di Novi (2025), a chapter devoted explicitly to inequality of access to healthcare services within a wider report on the sustainability of the Italian National Health Service. The analytical apparatus employed in the following chapters is adopted from this contribution: the need-adjusted Erreygers (ENA) concentration index, applied to the same four services considered here — blood tests, emergency department visits, specialist visits, and specialist diagnostic examinations — estimated on individual-level microdata from the ISTAT survey *Aspetti della Vita Quotidiana* (AVQ, Aspects of Daily Life), with socioeconomic status measured through an asset index obtained by extracting the first common factor from a matrix of polychoric correlations among household wealth proxies. Chapter 2 sets out each of these components in full; what follows here is a summary of the evidence they produce.

Drawing on the 2023 wave of the AVQ survey, the chapter documents a national pattern consistent with the one already reported by Masseria and Giannoni (2010) and Glorioso and Subramanian (2014): significant pro-rich inequity in specialist visits (ENA = 0.104), diagnostic examinations (ENA = 0.114) and, to a smaller extent, blood tests (ENA = 0.038), against a national ENA statistically indistinguishable from zero for emergency department access. At the regional level, the reported inequity in specialist care peaks in Calabria (ENA = 0.20), while diagnostic examinations exceed an ENA of 0.12 in Lombardy, Emilia-Romagna and Sardinia; pro-poor emergency department use, where statistically significant, is confined to Piedmont and Molise and remains close to zero in magnitude. A dedicated section of the same chapter examines waiting times and costs as factors of inequality in access, reporting cost as the more salient barrier in the Regions of the Mezzogiorno and waiting times as the more salient one across much of the Centre and North.

The present thesis takes this framework as given and extends it along the temporal dimension. Where Di Novi (2025) estimates the ENA index on a single annual wave, the following chapters estimate the same index, for the same four services and with the same construction of socioeconomic status, on six consecutive waves

of the AVQ survey covering 2018 to 2023, at both the national and regional level. The 2023 cross-section can thereby be placed within a sequence of annual point estimates rather than treated as an isolated observation. The contribution of the multi-wave design lies in documenting the temporal pattern of inequity and its decomposition. The 2023 relative levels of specialist-visit and diagnostic-examination ENA are already documented by Di Novi (2025) and are used here as a benchmark rather than treated as a novel finding. Two further extensions follow from the same design: weighted descriptive evidence on forgone care permits comparison of the relative importance of reported barriers across internally comparable periods, while the year-by-year regional estimates describe heterogeneity in regional ENA point estimates without formally testing territorial gradients.

As Table 1.1 makes clear, three of the four contributions reviewed are temporally concentrated in a window between 1994 and 2007 and rely in each case on a single cross-section or a pre-2008 panel; the fourth and most recent, Di Novi (2025), shares with the present thesis both the set of services examined and the estimator adopted, and provides the framework that the following chapters carry from a single annual wave (2023) to a multi-year series of repeated cross-sections. With the partial exception of Glorioso and Subramanian (2014), the earlier contributions consider a narrower set of services than those examined in the following chapters of this thesis.

Considered jointly, the four contributions reviewed above provide a rich picture of horizontal inequity in access to healthcare in Italy, spanning a period stretching from the mid-1990s to the immediate present. Masseria and Giannoni (2010) document the situation as it stood in 2000; Glorioso and Subramanian (2014) offer a more comprehensive and methodologically refined account based on 2005 data; Di Novi et al. (2019) examine the institutional determinants of regional health disparities over the period 1994–2007; and Di Novi (2025) reports pro-rich inequity in specialist and diagnostic care in 2023. The earlier studies conclude their empirical analysis by 2007, while the most recent contribution reports a single annual wave. What is therefore still to be documented is the annual pattern of need-adjusted inequity point estimates across the years immediately preceding the COVID-19 pandemic, the pandemic period, and the subsequent years through 2023.

This chapter has shown that the institutional and financial conditions underlying access to healthcare in Italy have continued to evolve since the period covered by the earlier literature, while regional differences in fiscal capacity, service provision and the role of private accredited providers remain salient in the evidence reviewed above. Di Novi (2025) provides the most direct recent benchmark by documenting pro-rich inequity in specialist and diagnostic care in 2023. The question taken up

Table 1.1: Comparison of the four principal Italian studies on horizontal equity of access to healthcare, and the positioning of the present thesis.

Study	Data & period	Method	Services examined	Key finding
Masseria & Giannoni (2010)	ISTAT Multiscopo (2000) + ECHP	Concentration-index-based horizontal inequity, need-standardised; disease-based subsample analysis	GP visits, specialist visits, hospitalisation, emergency care	Pro-rich inequity in GP, specialist and emergency care; hospitalisation equitable; driven mainly by income and regional variation
Glorioso & Subramanian (2014)	ISTAT Health Conditions Survey (2005), $n = 103,651$	Three-level multilevel logistic regression; inequity ratio; analysis stratified by self-rated health status	GP visits, specialist visits, basic medical tests, diagnostic tests, hospitalisation	Pro-poor inequity in GP care (0.78); pro-rich inequity in specialist care (1.40), basic tests (1.10) and diagnostics (1.24); hospitalisation equitable (0.92); inequity declines as health need increases
Di Novi et al. (2019)	ISTAT Multiscopo panel (1994–2007)	Difference-in-differences with continuous exposure (regional per capita GDP) around the 1998 fiscal decentralisation reform; Region and year fixed effects	Inequality in self-assessed health (median-based index); home, emergency, inpatient care; LHA appointments	No effect on between-Region disparities; significant reduction in within-Region disparities ($\approx 3\times$ within-Region SD), stronger in richer Regions
Di Novi (2025)	ISTAT AVQ, single wave (2023)	Need-adjusted Erreygers index (ENA), probit-based, national and regional estimates	Blood tests, emergency department visits, specialist visits, specialist diagnostic examinations	Significant pro-rich inequity in specialist visits (0.104), diagnostics (0.114) and blood tests (0.038); emergency care equitable nationally; specialist inequity peaks in Calabria (0.20)
<i>Present thesis</i>	ISTAT <i>Aspetti della Vita Quotidiana</i> , repeated cross-sections (2018–2023)	Need-adjusted Erreygers concentration index (ENA), following Di Novi (2025), estimated at national and regional level for each year	Blood tests, emergency department visits, specialist visits, specialist diagnostic examinations	<i>Empirical contribution of the present thesis (Chapters 3–4)</i>

in the following chapters is how that cross-sectional finding fits within the sequence of annual point estimates observed over 2018–2023, and how the composition of the need-adjusted index differs across those waves.

Adopting the methodological apparatus introduced in Section 1.6 — the standardisation of utilisation for health need and the Erreygers-corrected concentration index, employed in the fiscal-decentralisation literature reviewed above (Di Novi et al., 2019) and applied to these four services by Di Novi (2025) — this thesis constructs a need-adjusted horizontal inequity index for four specific healthcare services: blood tests, emergency department visits, specialist visits, and specialist diagnostic examinations. Unlike the studies reviewed in this section, which rely either on a single cross-section or on a pre-2008 panel, the present analysis uses six consecutive repeated cross-sections (2018–2023) of the Italian National Institute of Statistics’ *Aspetti della Vita Quotidiana* survey, allowing the resulting indices to be computed and compared descriptively at both the national and regional level for each year. The repeated cross-sectional analysis is complemented by weighted descriptive statistics on reported reasons for forgoing care, interpreted only across periods for which the questionnaire coding is sufficiently comparable. The data and methodology underlying this analysis are presented in full in the following chapter.

Chapter 2

Data and Methods

This chapter sets out the data and the empirical strategy underlying the analysis presented in Chapter 3. It proceeds in the order in which the components of that strategy are needed.

Section 2.1 describes the survey, the waves covered and the construction of the estimation sample. Section 2.2 introduces the concentration index, sets out the reasons for adopting the corrected version proposed by Erreygers (2009), and defines the variables the index requires: the utilisation outcomes, the socioeconomic ranking, and the covariates entering the utilisation model. Section 2.3 explains how those variables are made comparable across six waves of a questionnaire that was revised during the period.

The remaining sections concern estimation. Section 2.4 introduces the criterion of horizontal equity and the need-standardised index through which departures from it are measured; Section 2.5 sets out how the index is attributed to its components; and Section 2.6 describes the design of the analysis, the derivation of standard errors, and the treatment of sampling weights. Section 2.7 describes the complementary descriptive evidence on forgone care.

2.1 Data

2.1.1 The ISTAT Survey on Aspects of Daily Life

The empirical analysis in this thesis draws on individual-level microdata from the ISTAT multipurpose household survey *Aspetti della Vita Quotidiana* (AVQ, Aspects of Daily Life), one component of the integrated system of multipurpose household surveys coordinated by the Italian National Institute of Statistics. The survey has been conducted annually since 1993 and forms part of the National Statistical Pro-

gramme, which defines the set of statistical collections recognised as being of public interest.

Each wave covers approximately 25,000 households and more than 50,000 individuals of all ages, and collects information on education, employment, family and social relationships, housing and territorial context, leisure, social and political participation, and health and lifestyles. Every sampled unit is assigned a statistical weight designed to reproduce the resident population of the reference year, so that estimates are representative both nationally and at the regional level. The survey is administered as a repeated cross-section: each wave interviews an independently drawn sample, and respondents are not followed across years. This shapes the empirical strategy adopted throughout: the analysis describes how the distribution of access changes across successive representative samples of the Italian adult population, not how the circumstances of given individuals evolve over time.

Two features of the AVQ make it particularly suited to the study of equity in healthcare utilisation and access in the Italian context. First, it records utilisation of specific healthcare services over a well-defined reference period, alongside detailed information on health status, allowing utilisation to be related to need at the individual level. Second, it asks respondents directly whether they had to forgo a needed service and, if so, for what reason, which brings within reach a dimension of access that utilisation records alone cannot capture. The analysis is therefore situated within the institutional context of the Italian National Health Service (SSN), without assuming that every recorded episode of care was publicly delivered or publicly financed.

2.1.2 Waves and Access Conditions

The analysis covers six consecutive waves, from 2018 to 2023. The window is chosen so as to include two pre-pandemic years, the two years most affected by the COVID-19 emergency and the associated suspension of elective care, and two subsequent years, allowing the trajectory of access to be observed before, during and after the shock.

All six waves are drawn from ISTAT microdata files for the *Aspetti della Vita Quotidiana* survey. In each year, the record layout contains the variable `REGMF`, identifying the respondent's Region of residence, and the survey documentation states that the sample design supports estimation at the regional level. The same twenty-region classification based on `REGMF` is therefore used consistently across all six waves in the regional analysis.

2.1.3 Construction of the Estimation Sample

The analysis is restricted to the adult population, defined as respondents aged 18 and over, since the questions on healthcare utilisation and on forgone care are addressed to adults and since the socioeconomic and employment variables used below are not meaningful for minors. This restriction removes between 6,069 and 7,281 respondents per wave.

From the resulting adult sample, the main analysis uses a common complete-case sample within each wave. Observations are excluded where any of the variables required for the four utilisation analyses or for their common set of covariates is not recorded. These comprise the socio-demographic characteristics used as controls — marital status, educational attainment and employment status — the territorial classification, self-assessed health, and the variables recording use of the services covered by the health section of the questionnaire. The last group includes two services not analysed in this thesis, contacts with the out-of-hours medical service and receipt of home care, which are recorded alongside the four utilisation outcomes. This common-sample rule keeps the population underlying the main service estimates identical within a wave, but it also means that a missing value on one service can exclude an otherwise usable observation from the analysis of another service.

The 2023 wave requires additional scrutiny. After the general demographic and socio-demographic restrictions, 33,960 observations remain, of which 2,777 (8.18 per cent) have no recorded value for the emergency-department item (PROSOC). Inspection of the underlying records shows that this is not an isolated missing value on that item: among these 2,777 observations, 98.8 per cent are also missing on out-of-hours medical care (GUMED), 99.9 per cent on home care (ASSDO) and 99.5 per cent on hospitalisation (RICOV). At the same time, all 2,777 observations retain a valid self-assessed-health response, and approximately 37 per cent also contain a valid response to each of the later questions on specialist visits, blood tests and diagnostic examinations. The pattern is therefore consistent with concentrated non-recording in a block of healthcare-use questions rather than with general survey dropout.

The documentation available for this thesis does not identify a routing rule that would allow these 2,777 cases to be classified with confidence as structural missing values. The missingness rate is very similar by sex (8.02 per cent among men and 8.32 per cent among women; $p = 0.314$) and does not differ significantly across age categories ($p = 0.281$), whereas it varies significantly across Regions ($p < 0.001$), ranging from approximately 4.8 to 13.8 per cent. For this reason, some territorial selection in the 2023 emergency-department sample cannot be ruled out.

The main estimates nevertheless retain the common-sample definition in order to

preserve comparability with the other waves and across services. To assess whether the 2023 anomaly propagates materially to the results, all four 2023 indices are also re-estimated on service-specific samples in which the common demographic and covariate eligibility restrictions are retained, but observations are not additionally excluded because an outcome relating to another healthcare service is missing. The results of this sensitivity analysis are summarised in Section 3.1.6 and reported in full in the Appendix, Table 4.1.

Table 2.1: Main common estimation sample size by wave

Wave	Raw file	Adults	Common sample
2018	44,672	37,391	34,201
2019	45,483	38,378	34,798
2020	42,810	36,328	32,724
2021	45,597	39,140	35,760
2022	42,022	35,916	31,963
2023	41,750	35,681	31,098

Notes: Unweighted numbers of individual observations. “Raw file” includes all ages; “Adults” includes respondents aged 18 and over. “Common sample” applies the complete-case restrictions described in Section 2.1.3 and is shared by the four main utilisation analyses within each wave.

Source: Author’s calculations using ISTAT AVQ microdata, 2018–2023.

Table 2.1 reports the resulting main common samples. Unless explicitly stated otherwise, the concentration indices, their decomposition and the descriptive statistics in Chapter 3 use these samples. The service-specific 2023 estimates are used only as a sensitivity analysis and do not replace the main series.

2.2 Measuring Inequality: the Concentration Index

2.2.1 Definition and Interpretation

Socioeconomic inequality in healthcare utilisation is measured here by the concentration index, the standard rank-dependent measure in health economics, introduced by Wagstaff et al. (1991) and developed for the measurement of inequity in healthcare delivery by Wagstaff and van Doorslaer (2000). The index derives from the concentration curve, which plots the cumulative share of a healthcare variable against the cumulative share of the population ranked by socioeconomic status. The index is twice the area between this curve and the line of equality, and can be written as

$$C(y) = \frac{2}{\mu} \text{cov}(y_i, R_i), \quad (2.1)$$

where y_i is the utilisation indicator for individual i , μ its sample mean, and R_i the individual's fractional rank in the distribution of socioeconomic status.

The index takes the value zero when utilisation is distributed independently of socioeconomic position. Its sign indicates the direction of the association: a negative value means that utilisation is concentrated among the socioeconomically disadvantaged, a positive value that it is concentrated among the advantaged. Its magnitude describes the strength of that association. In what follows, a positive value is referred to as pro-rich and a negative value as pro-poor.

2.2.2 From the Standard Index to the Erreygers Index

Equation (2.1) was developed for variables measured on a ratio scale with no upper limit. All four utilisation measures used here are binary, and binary variables are bounded and, following the classification of Erreygers and Van Ourti (2011), treated as cardinal. This raises three difficulties for the standard index.

The first concerns the range the index can attain. Wagstaff (2005) shows that for a binary variable with mean μ the minimum and maximum values of $C(y)$ are $\mu - 1$ and $1 - \mu$, so that the attainable range narrows as the mean rises and collapses to zero as $\mu \rightarrow 1$. Erreygers (2009) generalises the result to any bounded variable. Since utilisation rates in this application range from around 7 per cent for emergency department access to around 60 per cent for blood tests, and vary across years and Regions as well, comparisons along any of these dimensions would be affected.

The second is that the standard index does not satisfy the *mirror* property: the inequality measured in a binary indicator and the inequality measured in its complement are not equal in magnitude. Since the choice between measuring utilisation and non-utilisation is arbitrary, an index whose magnitude depends on that choice is unsatisfactory.

The third is that $C(y)$ is invariant only to proportional transformations of the outcome, and therefore presupposes a ratio scale that a binary indicator does not possess.

Erreygers (2009) addresses all three by requiring a rank-dependent index to satisfy four properties — transfer, level independence, cardinal invariance and mirror — and showing that exactly one index satisfies them while having maximum bounds of -1 and $+1$. For a variable bounded between zero and one it reduces to

$$E(y) = 4\mu C(y). \quad (2.2)$$

The alternative normalisation proposed by Wagstaff (2005), which divides the

index by its mean-dependent bound, forces the range to $[-1, +1]$ but fails level independence, and with it monotonicity: Erreygers (2009) gives an example in which an unambiguously pro-rich change in the distribution leaves the Wagstaff index unchanged at its maximum, or even reduces it. Erreygers and Van Ourti (2011) show more generally that every index in this class satisfying the mirror property other than the Erreygers index is inverse-relative for some distributions, increasing in magnitude when relative differences fall while absolute differences are held constant. Erreygers (2009) illustrates the practical consequences in an application to child stunting across 47 countries, where the correlation between prevalence and measured inequality is negative under the standard and Wagstaff indices and positive under the corrected index.

One qualification should be stated explicitly. The rescaling fixes the *maximum* bounds of the index at -1 and $+1$, but the range actually attainable for a given mean remains $\pm 4\mu(1 - \mu)$ for a binary variable (Erreygers, 2009, eq. 28): it reaches unity only when $\mu = 0.5$ and tends to zero at either extreme. What equation (2.2) delivers is therefore not a mean-independent attainable range, but an index that is simultaneously relative and absolute, invariant to the direction in which the outcome is coded, and unique in satisfying the four properties above. The same index is adopted, in the same institutional setting, by Di Novi (2025).

2.2.3 Utilisation Outcomes

Four binary measures of healthcare utilisation are considered. For blood tests, specialist diagnostic examinations and specialist visits, the measure records whether the respondent used the service at least once in the twelve months preceding the interview. Emergency department access instead refers to attendance during the preceding three months. The AVQ items identify utilisation of the service but do not distinguish the channel through which care was obtained: a recorded episode may refer to a public provider, a private provider accredited within the SSN, or care purchased entirely privately. The outcomes should therefore not be interpreted as measures of publicly delivered SSN care alone. Rather, they provide a picture of healthcare utilisation in the Italian institutional context, in which the SSN constitutes the organising framework and public, accredited private and fully private provision coexist.

- **Blood tests** record whether the respondent underwent haematochemical analyses, either as part of monitoring their own health or on a physician's prescription.

- **Specialist diagnostic examinations** record whether the respondent underwent diagnostic investigations such as imaging or instrumental tests.
- **Specialist visits** record whether the respondent consulted a medical specialist.
- **Emergency department access** records whether the respondent attended an accident and emergency department.

The four differ in a way that matters for interpretation. The first three are programmable services: they are typically scheduled in advance and can be obtained through public, accredited private or fully private channels, with financing arrangements that may include co-payments or direct private payment. Emergency department access is not programmable and is not subject to the same booking arrangements.

2.2.4 The Socioeconomic Status Index

The concentration index requires individuals to be ranked along a socioeconomic dimension. The AVQ does not collect household income or wealth in a form suitable for this purpose, so the ranking variable is constructed as an asset-based index, following the approach set out by Vyas and Kumaranayake (2006) for settings in which direct income measures are unavailable.

The index is built from seven items describing housing characteristics and durable asset ownership: home ownership, presence of a terrace, presence of a garden, employment of domestic help, air conditioning, number of cars owned, and home internet access. Because these items are binary or discrete rather than continuous, a Pearson correlation matrix would systematically understate the association between them. The index is therefore obtained by extracting the first common factor from a matrix of *polychoric* correlations (Olsson, 1979), and normalising the resulting score to the $[0, 1]$ interval.

The single retained factor accounts for 89.97 per cent of the common variance in 2018, supporting a one-dimensional reading of the underlying construct. Loadings are highest for the number of cars owned (0.66), home ownership (0.54) and home internet access (0.53), and lowest for employment of domestic help (0.19) and presence of a terrace (0.23).

The role of this index is twofold. It defines the fractional rank on which the concentration index of the outcome, and of every covariate entering the decomposition, is computed (equation (2.1)); and, like the need and non-need variables

described below, it also enters the probit model in Section 2.5 as an explanatory variable, so that its own marginal effect and concentration index combine, through equation (2.3), into the SES contribution reported throughout Chapter 3.

2.2.5 Need and Non-Need Variables

Following the standard distinction in this literature, the covariates of the utilisation model are partitioned into variables that legitimately justify differences in utilisation and variables that do not.

Four **need** variables are used. Age is entered as an indicator for respondents aged 65 and over, and sex as an indicator for female respondents; both are demographic characteristics conventionally treated as legitimate determinants of the need for care. Self-assessed health is derived from the question asking respondents to rate their health in general on a five-point scale, and takes the value one for those reporting good or very good health. Chronic morbidity is the indicator constructed in Section 2.3.1. Self-assessed health is widely used as a summary measure of health status and is a strong predictor of subsequent healthcare utilisation; it is coded here in the direction of good health, so that its expected association with utilisation is negative.

The **non-need** variables comprise marital status, educational attainment and employment status. These are socioeconomic characteristics: differences in utilisation associated with them are not justified by differences in need, and are therefore part of what the analysis seeks to measure rather than to standardise away. The utilisation models also include a full set of regional indicators, capturing territorial heterogeneity in the organisation, capacity and accessibility of healthcare services within the Italian health system.

2.2.6 Forgone Care

Alongside utilisation, the AVQ records whether respondents had to forgo care they needed. Two of the four services examined here are covered: specialist visits and specialist diagnostic examinations. No equivalent question is asked in respect of blood tests or emergency department access.

For each of the two services, respondents who needed the service during the twelve months preceding the interview are asked whether they were unable to obtain it. Those who report having forgone the service are then asked to indicate why, from a set of listed options and with more than one response permitted. The options are grouped here into four categories: cost, waiting times, inconvenience of

location or opening hours, and other reasons. In the released 2020 microdata there is no dedicated item identifying COVID-19 as a reason for forgoing care; from 2021 onwards the questionnaire contains a separate COVID-19 response item for both specialist visits and diagnostic examinations. Official ISTAT aggregate reporting nevertheless identifies COVID-related reasons as an important component of forgone care already in 2020, so the absence of a dedicated 2020 microdata item should not be read as evidence that COVID-related renunciation was absent in that year (Istituto Nazionale di Statistica (Istat), 2021).

These variables measure something the utilisation outcomes cannot. The concentration index describes how delivered care is distributed across the socioeconomic ranking; it is by construction silent on need that never translated into a utilisation record. The forgone care variables identify that group directly, and the reasons reported by its members indicate which barriers prevented access. They are used descriptively, as set out in Section 2.7.

2.3 Harmonisation Across Waves

The variables described in the preceding section must be available in identical form in each of the six waves for the resulting series to be internally comparable. Since ISTAT periodically revises the questionnaire and its coding conventions, two of them require explicit construction rather than direct use, and a third source of variation concerns item non-response. Because each affects the interpretation of the results reported in Chapter 3, the treatment adopted is set out here in full.

2.3.1 Chronic Morbidity

The chronic morbidity indicator introduced in Section 2.2.5 is not taken directly from the survey. The questionnaire contains two distinct measures of chronic illness. The first is a single global question asking whether the respondent suffers from chronic illnesses or long-standing health problems, where “long-standing” is defined as lasting, or expected to last, at least six months. The second is a battery in which respondents are asked, condition by condition, whether they suffer from each of a list of named chronic diseases.

The summary variable derived from the global question is not released in the 2021 and 2022 microdata files, so an indicator based on it cannot span the observation window. The indicator adopted here is therefore constructed from the condition-specific battery, which is available in every wave. A respondent is classified as chronically ill if they report at least one of fourteen conditions: diabetes,

hypertension, myocardial infarction, angina or other heart disease, chronic bronchitis or emphysema, bronchial asthma, allergic conditions, malignant tumour, gastric or duodenal ulcer, gallstones, liver cirrhosis, kidney stones, arthrosis or arthritis, and osteoporosis. The neurological item is excluded because it is renamed from 2021 onwards and its definition altered at the same time, so that including it would introduce the very discontinuity the reconstruction is designed to avoid.

The battery records conditions in groups of four, with each item within a group carrying its own affirmative code rather than a value common to all items. Because a change in this convention would corrupt the indicator without generating an error, the construction is preceded in every wave by assertions verifying that each variable takes only its expected values. These checks pass in all six waves.

Where both measures are available they do not coincide, and the difference follows from the way the two questions are asked. In 2018 the summary variable classifies 31.5 per cent of the adult sample as chronically ill, against 47.5 per cent for the item-based indicator. The two agree on 74 per cent of observations, and the disagreement is asymmetric: 84.8 per cent of those identified as chronically ill by the summary variable are also identified by the item-based measure, whereas 30.6 per cent of those classified as not chronically ill by the summary variable nonetheless report at least one of the fourteen conditions.

The global question requires respondents both to apply the six-month duration criterion themselves and to judge whether their own situation amounts to a chronic illness. Conditions such as hypertension, allergic disorders or arthrosis are frequently not classified as such by those who have them, yet the same respondents report them when the conditions are named individually. The item-based indicator therefore captures a broader notion of morbidity. Being broader, it absorbs a larger share of observed inequality into the need component, so that the resulting inequity estimates are somewhat higher than those that would be obtained using the summary variable.

A further consideration supports this choice on grounds of data quality. The health section of the questionnaire is not compulsory, since it collects special categories of personal data within the meaning of Article 9 of the General Data Protection Regulation, and item non-response is correspondingly high: in 2023 the summary variable is missing for 10.3 per cent of respondents. By contrast, the share of respondents for whom the entire battery of fourteen conditions is unrecorded is 1.1 per cent in 2018. The reconstruction thus rests on a more completely observed set of items, in addition to being the only definition available for all six waves.

2.3.2 The 2022 Change in the Coding of Forgone Care

The variables recording forgone care, introduced in Section 2.2.6, are coded differently before and after 2022. Until 2021 the questionnaire follows a skip-pattern design: a single code identifies respondents who did not forgo the service, and those who did are routed directly to the questions asking why, so that forgoing is identified by the absence of that code rather than by a dedicated one. From 2022 an explicit code for forgoing is introduced alongside the code for not forgoing.

Table 2.2: Coding of the forgone care filter variables, raw files (per cent)

Wave	Diagnostic examinations		Specialist visits	
	No response	Explicit code	No response	Explicit code
2018	6.28	—	9.66	—
2019	6.20	—	9.04	—
2020	8.29	—	12.37	—
2021	8.79	—	14.18	—
2022	4.65	3.89	4.20	7.53
2023	9.52	4.82	2.92	8.09

Notes: Coding checks on the raw files, not population estimates of forgone care. “No response” denotes an unrecorded filter value; “Explicit code” denotes an affirmative forgone-care response. A dash indicates that no dedicated affirmative code exists in that wave, not a zero share. Under the pre-2022 skip pattern, absence of the negative code identifies forgone care but may also include item non-response; from 2022, an explicit affirmative code is used. The columns should not be read as a continuous series of forgone-care rates.

Source: Author’s calculations using ISTAT AVQ microdata, 2018–2023.

Table 2.2 reports the distribution of the two filter variables in the raw files. The change affects both batteries simultaneously, which identifies it as a revision of the questionnaire rather than an isolated recoding. The variables are therefore constructed according to the convention in force in each wave: as the absence of the negative code until 2021, and as the explicit affirmative code from 2022. Both conventions identify the same underlying event as intended by the questionnaire design in force at the time; neither is a proxy for the other.

What changes across the revision is the composition of the residual category, not the validity of either convention taken on its own terms. Until 2021, a respondent who skips the filter question for any reason is indistinguishable, within the filter variable itself, from one who skips it because they forwent care; from 2022 the explicit code removes this ambiguity at the filter stage. Section 2.3.3 shows that this has little practical effect on the pre-2022 waves: in 2018 and 2019, the large majority of respondents identified as forgoing care under the skip-pattern convention go on to report a specific reason, which would be an unlikely pattern if the filter were picking up substantial numbers of respondents who had not in fact forgone care.

The practical consequence of the revision is narrower than a difference in validity: the level of the forgone care rate is not comparable across the two conventions, since the pre-2022 residual and the post-2022 explicit code are not constructed to capture identical boundaries at the margin. For this reason the descriptive statistics on forgone care in Chapter 3 are presented in two blocks, 2018–2021 and 2022–2023, matching the two coding conventions, and comparisons of level are confined to within each block.

2.3.3 COVID-19 Coding and the Residual Reason Category

Respondents who report forgoing care are asked to indicate the reason. For each wave, a residual category can be constructed for respondents for whom none of the reason items available in the released microdata is recorded. The size of this residual varies sharply across waves (Table 2.3).

Table 2.3: Share of respondents forgoing care with no separately coded reason among the categories available in the released microdata (per cent)

	2018	2019	2020	2021	2022	2023
Specialist visits	9.1	8.1	45.8	15.0	3.9	0.9
Diagnostic examinations	16.4	19.6	50.9	24.8	5.4	2.8

Notes: The denominator is respondents identified as forgoing the respective service in each wave, using the filter convention described in Section 2.3.2. The residual identifies cases with no recorded reason among the items available in the released microdata; it does not imply that respondents had no reason for forgoing care. Dedicated COVID-19 reason items are unavailable in 2020 and included from 2021 onwards, so the 2020 residual is not directly comparable with that in 2021–2023.

Source: Author’s calculations using ISTAT AVQ microdata, 2018–2023.

In 2018 and 2019, the residual share is relatively low — between 8 and 20 per cent depending on service and wave — so that the large majority of respondents identified as forgoing care under the pre-2022 skip-pattern convention described in Section 2.3.2 go on to name a specific barrier.

The 2020 wave requires separate treatment. The released 2020 record layout contains the standard reason items up to the generic “other” category (VISRI7 and ACCRI7), but no dedicated COVID-19 items corresponding to the VISRI8 and ACCRI8 variables available from 2021 onwards. At the same time, ISTAT’s published aggregate statistics identify COVID-related reasons as a major component of forgone care in 2020 (Istituto Nazionale di Statistica (Istat), 2021). The 2020 residual should therefore not be interpreted as pure item non-response, nor as evidence that respondents had no reason for forgoing care. Rather, it is the share for which no reason can be separately identified from the coded categories available in the released microdata used here. The service-specific 2020 microdata do not permit COVID-related

reasons to be isolated as an autonomous category.

From 2021 onwards, VISRI8 and ACCRI8 explicitly identify COVID-19 as a reason for forgoing specialist visits and diagnostic examinations. COVID-19 is therefore retained as a separate category in all descriptive results from 2021 onwards. Because the set of separately coded reasons differs between 2020 and later pandemic waves, the composition of reasons in 2020 is not directly comparable with that in 2021–2023. The results chapter consequently reports the 2020 COVID entry as not separately identifiable rather than assigning these cases to a “no reason” category.

2.4 Measuring Inequity: Need Standardisation and the ENA Index

2.4.1 Horizontal Equity

The Erreygers index measures how utilisation is distributed across the socioeconomic distribution, but not all of that distribution is inequitable. If the poor are in worse health than the rich, greater utilisation among the poor is what a well-functioning health system should produce. Inequality in utilisation and inequity in access are therefore distinct quantities, and the first is not by itself informative about the second.

The criterion adopted here is horizontal equity: equal treatment for equal need, irrespective of socioeconomic position. A system satisfies it when two individuals of the same sex and age, with the same need for care, have the same probability of obtaining a service regardless of their economic circumstances. Departures from this criterion constitute the potentially avoidable component of inequality, and it is that component the analysis seeks to isolate (O’Donnell et al., 2008).

This is the same criterion introduced in Section 1.6, where it underlies the horizontal inequity index $HI = C_M - C_N$, the difference between the concentration index of actual utilisation and that of need-expected utilisation. The need-adjusted index defined in this chapter is a different implementation of the same underlying idea, expressed through the decomposition of a single index rather than through the difference of two. Where HI isolates inequity by comparing the concentration index of observed utilisation to that of a separately estimated need-expected counterfactual, ENA isolates it by subtracting, from the observed Erreygers index, only those components of its decomposition attributable to need (Section 2.5). The two approaches share the same object — the gap between the socioeconomic gradient in care actually received and the gradient that need alone would justify — and the

same sign convention, positive indicating inequity favouring the better-off.

2.4.2 Indirect Standardisation

Isolating it requires separating the part of observed inequality attributable to differences in need from the remainder. This is achieved through indirect standardisation, implemented via the decomposition of the concentration index (O'Donnell et al., 2008; Wagstaff et al., 2003).

The procedure begins from a model of utilisation. For each service, year and territorial level, the binary utilisation indicator is regressed on the socioeconomic status index, the four need variables, the non-need controls and the regional indicators, using a probit specification estimated by maximum likelihood. What the decomposition requires from this model is the marginal effect of each covariate on the probability of utilisation; these are obtained as marginal effects evaluated at the sample means of the covariates rather than as average marginal effects, a distinction that affects the magnitude of the individual contributions and is stated here for completeness.

Erreygers (2009) shows that the corrected index admits a decomposition directly analogous to that of the standard index. Writing β_k for the marginal effect of covariate x_k , μ_k for its mean and $C(x_k)$ for its own concentration index, computed as in equation (2.1) but with the covariate in place of the utilisation outcome, the contribution of that covariate to the Erreygers index is

$$c_k = 4 \beta_k \mu_k C(x_k), \quad (2.3)$$

and the sum of these contributions across all covariates, plus a residual term capturing the portion of inequality the model does not explain, reproduces $E(y)$. The factor of four and the absence of any division by the mean of the outcome follow from the definition $E(y) = 4\mu C(y)$ given in Section 2.2.2: the corresponding decomposition of the standard index divides each term by μ , and it is precisely the Erreygers rescaling that cancels it. Expressing contributions on the same scale as $E(y)$ is what allows them to be subtracted from it directly, as equation (2.4) below requires. Their interpretation, and the use of the decomposition as an analytical device in its own right, are taken up in Section 2.5.

2.4.3 The Need-Adjusted Index

The need-adjusted Erreygers index is obtained by removing from the observed index the contributions of the four need variables:

$$\text{ENA} = E(y) - \sum_{k \in \mathcal{N}} c_k, \quad (2.4)$$

where $\mathcal{N}$ denotes the set comprising age, sex, self-assessed health and chronic morbidity. The contributions of the socioeconomic status index, of the non-need controls and of the regional indicators are not subtracted: these represent inequality that is not justified by differences in need, and are therefore part of what the index is intended to capture.

The resulting quantity answers a counterfactual question. It measures the inequality in utilisation that would remain if need were distributed equally across the socioeconomic distribution, while the association between socioeconomic position and utilisation remained as observed. A positive value indicates that, at equal need, utilisation is concentrated among the better off; a negative value that it is concentrated among the worse off; a value of zero that the horizontal equity criterion is satisfied.

The relationship between the two indices is itself informative. Where need is concentrated among the socioeconomically disadvantaged, the need contributions are negative, so that subtracting them raises the index: in such cases the observed distribution understates the degree of inequity, because greater utilisation among the poor partly reflects their greater need rather than more favourable access. The gap between $E(y)$ and ENA thus measures the extent to which the distribution of need conceals the equity picture, and is reported alongside both indices in Chapter 3.

2.5 Decomposition of the Index

2.5.1 Purpose

The decomposition introduced in Section 2.4.2 is a prerequisite for defining the need-adjusted index, but it serves a second and independent purpose: it attributes the level of the index, and its variation over time, to the individual characteristics that generate it.

Two indices of equal magnitude can arise from entirely different configurations. A given degree of pro-rich inequity may be driven principally by the socioeconomic gradient itself, or principally by educational attainment, or by a combination that shifts from one wave to the next. The aggregate index cannot distinguish these cases; the decomposition can. Contributions are therefore reported separately for each service and each year in Chapter 3, rather than only as an intermediate step in constructing ENA.

2.5.2 Reading a Contribution

Each contribution defined by equation (2.3) is the product of three distinct quantities, and its interpretation requires all three.

The marginal effect β_k measures how strongly the characteristic predicts utilisation. The concentration index $C(x_k)$ measures how the characteristic itself is distributed across the socioeconomic ranking, taking a negative value when the characteristic is concentrated among the disadvantaged. The mean μ_k measures how widespread it is in the population, and scales the contribution accordingly.

The sign of a contribution follows from the product of the first two. A characteristic that raises the probability of utilisation and is concentrated among the better off contributes positively to measured inequality; so does one that lowers the probability of utilisation and is concentrated among the worse off. Contributions are negative when the two signs differ. A characteristic that is a strong predictor of utilisation but is evenly distributed across the socioeconomic ranking contributes nothing, however large its marginal effect; and the converse holds equally.

Because a contribution can differ across waves through any of its three components, interpretation requires attention to each component separately. A smaller need-contribution point estimate may be associated with a lower estimated marginal effect, a different socioeconomic distribution of the characteristic, or a different prevalence in the population, and these have different substantive readings.

2.5.3 Covariates Decomposed and the Residual

Contributions are computed for twelve covariates: the socioeconomic status index; the four need variables (age, sex, self-assessed health, chronic morbidity); and the seven non-need controls (three marital status indicators, three educational attainment indicators, and employment status).

The regional indicators are included in the utilisation model, so that the marginal effects of the remaining covariates are estimated net of territorial heterogeneity, but their contributions are not extracted individually. They are therefore absorbed, together with the portion of inequality the model does not explain, into the residual term. The residual reported in Chapter 3 is accordingly a composite quantity, and not a measure of unexplained inequality alone.

2.5.4 Decomposing Differences Across Waves

Since equation (2.4) is an identity, the difference in the need-adjusted index between two waves can be written as the difference in the observed index less the difference

in the need contributions:

$$\Delta\text{ENA} = \Delta E(y) - \sum_{k \in \mathcal{N}} \Delta c_k. \quad (2.5)$$

This identity provides an arithmetic accounting of differences in measured inequity across waves. A higher ENA point estimate may coincide with a higher unadjusted index, with less negative need contributions, or with some combination of the two. Likewise, a difference in a need contribution may arise through its estimated marginal effect, prevalence or socioeconomic concentration. These decompositions describe which components account arithmetically for differences in point estimates; they do not by themselves establish that cross-wave differences are statistically significant or identify the causal mechanism generating them. Equation (2.5) is used in Chapter 3 in this descriptive accounting sense.

2.6 Estimation and Inference

2.6.1 Design of the Analysis

The Erreygers and need-adjusted indices are estimated separately for each of the four services, for each of the six waves, and at two territorial levels: nationally, and for each of the twenty Italian Regions. This produces a complete matrix of estimates spanning service, year and territory, allowing the analysis to describe the overall degree of horizontal inequity in healthcare utilisation within the Italian SSN context, its annual variation across waves, and its heterogeneity across Regions. This interpretation concerns the healthcare system in which utilisation occurs and does not imply that every observed service was publicly delivered. The same combination of national and sub-national annual estimation is adopted by Di Novi (2025).

Each estimate is obtained on the sample relevant to it. Regional indices are computed on the regional subsample, so that the socioeconomic ranking, the utilisation model and the decomposition are all internal to the Region concerned. Regional estimates therefore describe inequity within each Region and not the position of one Region relative to another.

2.6.2 The Socioeconomic Ranking

Both the index and the bootstrap procedure rely on the fractional socioeconomic rank derived from the index described in Section 2.2.4. Once the index has been normalised to the $[0, 1]$ interval, each individual is assigned a fractional rank cor-

responding to their position in the socioeconomic distribution of the relevant estimation sample — national or regional according to the level at which the index is computed. Individuals sharing the same value of the underlying index are assigned the average rank of the tied group.

Quintiles of the same index are used for two further purposes: as strata in the resampling procedure described below, and as the basis for the descriptive tables in Chapter 3. Because the index is constructed from seven binary or discrete items, it takes a limited number of distinct values, and the quintiles are consequently not exactly equinumerous.

2.6.3 Standard Errors

Standard errors for both $E(y)$ and ENA are obtained by bootstrap with 100 replications, stratified by quintile of the socioeconomic status index so that resampling preserves the socioeconomic composition of the original sample within each replication. The procedure is applied identically at the national level and for each Region, for every wave and every service. All standard errors, z statistics and significance levels reported in Chapter 3 derive from this procedure rather than from the analytic variance formula. Throughout Chapter 3, one, two and three asterisks denote significance at the 10, 5 and 1 per cent level respectively.

Moving from the national sample to the regional subsamples reduces the number of observations available for each estimate. The AVQ is designed to be representative both nationally and regionally, but necessarily allocates fewer respondents to each individual Region than to the country as a whole. Smaller samples translate into greater sampling variability, which the bootstrap reflects in the associated standard errors. Regional estimates therefore carry standard errors several times larger than the national ones.

The resampling does not reproduce the two-stage structure of the AVQ design, in which municipalities constitute the primary sampling units. Standard errors are accordingly not adjusted for clustering at that level.

2.6.4 Reliability of the Regional Estimates

Because regional standard errors are substantially larger than national ones, part of the dispersion observed across the twenty Regions in any given wave is attributable to sampling error rather than to genuine differences in inequity. This can be quantified directly. Under classical measurement theory, the observed variance of a set of estimates decomposes into a component attributable to true between-unit variation

and a component attributable to measurement error:

$$\text{Var}(\hat{\theta}_r) = \text{Var}(\theta_r) + \text{Var}(\varepsilon_r), \quad (2.6)$$

where $\hat{\theta}_r$ is the ENA estimate for Region r , θ_r its unobserved true value, and ε_r the sampling error captured by the bootstrap standard error. The left-hand side is estimated directly as the sample variance of the twenty regional point estimates in a given wave; the error component is approximated by the average of the squared regional bootstrap standard errors, $\overline{\text{SE}^2}$. The *reliability ratio* reported in Chapter 3 is then

$$\rho = \frac{\max\left(\text{Var}(\hat{\theta}_r) - \overline{\text{SE}^2}, 0\right)}{\text{Var}(\hat{\theta}_r)}, \quad (2.7)$$

the share of observed between-Region variance not attributable to sampling error. A ratio of one indicates that regional differences are measured with enough precision that sampling error contributes negligibly to their observed dispersion; a ratio close to zero indicates that the twenty regional estimates would look much the same if measured on independent samples of the same size, so that little of their apparent dispersion should be read as a genuine territorial pattern. The maximum operator prevents the ratio from taking a negative value in waves where sampling error alone would be expected to exceed the observed dispersion. The ratio is computed separately for each service, averaging $\text{Var}(\hat{\theta}_r)$ and $\overline{\text{SE}^2}$ across the six waves to obtain a single summary figure.

2.6.5 Weighting

The AVQ assigns each respondent a sampling weight reproducing the resident population of the reference year. The same weighting principle is applied to all quantities entering the inequity analysis. The utilisation model is estimated using probability weights; the means of the outcome and of all covariates are weighted; and the covariance terms used to construct the covariate concentration indices are also computed with the sampling weight. The socioeconomic fractional rank is constructed from cumulative sampling weights, and the Erreygers index itself is evaluated using the same weight. Consequently, the raw Erreygers index and the need contributions subtracted from it to obtain the ENA refer to the same weighted target population.

Sampling weights are also used for the descriptive results reported in Chapter 3, including utilisation and forgone-care rates and the distribution of reported reasons for forgoing care. Unweighted tabulations used only as coding or diagnostic checks

are not interpreted as population estimates.

2.6.6 Software

All estimates were produced in Stata 19.5. The Erreygers index is computed using the routine written by Erreygers, dated 2007 in its header, which implements the regression-based estimator of the corrected index subsequently published as Erreygers (2009). The decomposition and bootstrap procedures were written around it for this analysis.

2.7 Forgone Care: Descriptive Analysis

2.7.1 Denominators and Breakdowns

Two distinct quantities are reported from the variables introduced in Section 2.2.6. The rate of forgone care is computed over the whole estimation sample and expresses the share of adults who report having forgone the service. The shares of reported reasons are computed over those who forwent the service, and therefore express the proportion of those forgoing care who cite each barrier. Since the question admits multiple responses, these shares sum to more than one hundred per cent. A further category records respondents who report forgoing care but indicate no reason at all; as set out in Section 2.3.3, it is reported explicitly rather than left implicit in the denominator.

Both quantities are reported separately for each wave, and disaggregated along three dimensions: quintile of the socioeconomic status index, age class, and macro-area. Age classes are defined as 18–44, 45–64, and 65 and over, the last coinciding with the age indicator entering the utilisation models. Macro-areas follow the five-category ISTAT territorial partition: North-West, North-East, Centre, South, and Islands.

To keep the tables legible, only the first, third and fifth quintiles are displayed; the shares are computed over all five, and the monotonicity of any gradient described in the text is assessed on the complete set. Cell sizes are reported alongside every disaggregated figure, since the number of respondents forgoing care in a given wave, quintile or macro-area can be small.

Chapter 3

Results

3.1 Sample and Descriptive Statistics

The sample composition follows the construction described in Section 2.1.3, yielding estimation samples of between 31,098 and 35,760 adults per wave (Table 2.1). The subsections below describe utilisation of the four healthcare services examined, the distribution of the need and socioeconomic variables entering the analysis, and the relationship between utilisation and socioeconomic position that motivates the standardisation carried out in Section 3.2.

3.1.1 Utilisation of Healthcare Services

Table 3.1: Utilisation rates by service and wave (per cent)

	2018	2019	2020	2021	2022	2023
Blood tests	58.9	57.5	53.2	51.7	55.7	52.2
Specialist visits	48.3	45.8	41.6	41.5	43.9	43.0
Diagnostic examinations	43.4	40.9	37.6	35.9	37.9	37.8
Emergency department	7.0	7.3	5.0	4.1	5.7	6.0

Notes: Sampling-weighted percentages in the adult common estimation sample of each wave. Utilisation refers to at least one use in the preceding twelve months, except emergency department attendance, which refers to three months. Public and private provision are not distinguished.

Source: Author’s calculations using ISTAT AVQ microdata, 2018–2023.

Table 3.1 reports utilisation rates by service and wave. The four services differ markedly in the share of the adult population using them, from around 7 per cent for emergency department access to around 60 per cent for blood tests, consistent with the distinction between programmable and non-programmable services set out in Section 2.2.3. The weighted utilisation point estimates are lower in 2020 than in 2019 for all four services, most notably for blood tests and specialist visits. The 2022–2023

point estimates are higher again for several services, although they remain below their 2018–2019 values. These are descriptive comparisons of annual point estimates rather than formally tested year-to-year differences. The pattern is examined in relation to the concentration indices in Section 3.2.

3.1.2 Need and Socioeconomic Variables

In the 2018 wave, 28.9 per cent of the estimation sample is aged 65 or over, and 47.5 per cent reports at least one of the fourteen conditions underlying the chronic morbidity indicator constructed in Section 2.3.1. The corresponding shares are of similar magnitude in the remaining waves, as reported alongside the national index estimates in Section 3.2.

The socioeconomic status index described in Section 2.2.4 is, by construction, built from a limited number of binary and discrete items and therefore takes a restricted set of distinct values. This is reflected in the distribution of the resulting quintiles, which are not equinumerous: in 2018 they contain, respectively, 21.6, 21.0, 17.8, 24.4 and 15.2 per cent of the estimation sample (Table 3.2). The pattern is similar in the other waves. This imbalance does not affect the concentration index or its decomposition, which are computed on the continuous fractional rank rather than on the quintile grouping; it is relevant only to the descriptive tables reported by quintile in this chapter, where cell sizes should be read with this in mind.

Table 3.2: Distribution of the SES quintiles, 2018

Quintile	N	Share (per cent)
Q1 (poorest)	7,398	21.6
Q2	7,174	21.0
Q3	6,098	17.8
Q4	8,329	24.4
Q5 (richest)	5,202	15.2

Notes: SES denotes socioeconomic status. N and the corresponding shares describe the unweighted composition of the 2018 common estimation sample. Ties in the discrete SES index produce unequal quintile sizes; Q1 and Q5 identify the lowest and highest SES groups, respectively.

Source: Author’s calculations using ISTAT AVQ microdata, 2018.

3.1.3 Utilisation by Socioeconomic Status Quintile

Table 3.3 reports utilisation rates by SES quintile for 2018, the first wave of the observation window. The four services display markedly different gradients. Specialist visits and diagnostic examinations rise from the poorest to the richest quintile, with a slight dip at the fourth quintile for specialist visits: from 44.3 to 51.8 per cent for

specialist visits and from 39.2 to 46.2 per cent for diagnostic examinations. Emergency department access falls, from 8.5 to 6.2 per cent. Blood tests, by contrast, show almost no gradient: utilisation ranges narrowly between 57.5 and 60.7 per cent across all five quintiles, with no monotonic pattern.

Table 3.3: Utilisation rate by SES quintile, 2018 (per cent)

	Q1	Q2	Q3	Q4	Q5
Blood tests	57.5	60.7	58.8	58.3	59.9
Specialist visits	44.3	47.4	50.1	49.4	51.8
Diagnostic examinations	39.2	43.0	45.0	45.0	46.2
Emergency department	8.5	7.3	6.9	5.9	6.2

Notes: Sampling-weighted utilisation rates within each socioeconomic status (SES) quintile of the adult common sample. Q1 is the lowest and Q5 the highest SES group. Recall periods are twelve months for the three programmable services and three months for emergency department attendance.

Source: Author’s calculations using ISTAT AVQ microdata, 2018.

This asymmetry across services is the descriptive counterpart of the standardisation carried out in Section 2.4. Where the raw gradient is pronounced, as for specialist visits and diagnostic examinations, the unadjusted concentration index already signals a pro-rich distribution of utilisation. Where the raw gradient is flat, as for blood tests, the unadjusted index is close to zero and offers little indication of inequity by itself; whether the flat distribution is equitable depends on whether need for the service is itself distributed evenly across the socioeconomic ranking, a question the standardised index addresses directly. The two readings are brought together in Section 3.2.

Table 3.4 reports the gap between the richest and poorest quintile for each service in every wave, confirming that the 2018 pattern is not specific to that year. The gap is positive and of similar magnitude for specialist visits and diagnostic examinations in all six waves, negative for the emergency department in all six, and consistently small — positive throughout, and never exceeding 4.0 percentage points — for blood tests.

3.1.4 Utilisation by Age Class

Table 3.5 reports the gap in utilisation between respondents aged 65 and over and those aged 18–44, the two age classes defined in Section 2.7.1. For the three programmable services the gap is large and stable across all six waves: older respondents are between 17 and 23 percentage points more likely to use specialist visits, between 16 and 22 points more likely to use diagnostic examinations, and between 24 and 31 points more likely to use blood tests. For the emergency department the gap is

Table 3.4: Utilisation gap between the richest and poorest SES quintile (Q5–Q1, percentage points)

	2018	2019	2020	2021	2022	2023
Specialist visits	+7.5	+5.5	+9.6	+8.1	+6.7	+6.7
Diagnostic examinations	+7.0	+4.9	+8.9	+7.7	+7.6	+8.1
Blood tests	+2.4	+1.6	+3.8	+1.6	+2.7	+4.0
Emergency department	−2.3	−2.6	−0.4	−1.9	−0.5	−0.7

Notes: Differences between sampling-weighted utilisation rates in the highest and lowest socioeconomic status quintiles (Q5 minus Q1) of the adult common sample. Positive values indicate higher utilisation in Q5. Gaps are descriptive percentage-point differences, not need-adjusted indices or tests of differences between groups.

Source: Author’s calculations using ISTAT AVQ microdata, 2018–2023.

positive but an order of magnitude smaller, between 2 and 4 points, consistent with the much smaller role need plays in accessing this service (Section 3.3.3).

Table 3.5: Utilisation gap between respondents aged 65+ and 18–44 (percentage points)

	2018	2019	2020	2021	2022	2023
Specialist visits	+23.4	+23.3	+19.3	+17.1	+22.3	+19.3
Diagnostic examinations	+21.6	+21.4	+17.8	+16.0	+20.7	+17.5
Blood tests	+30.7	+28.5	+25.9	+24.4	+28.5	+24.7
Emergency department	+2.9	+3.7	+2.1	+1.9	+2.3	+2.5

Notes: Sampling-weighted utilisation rate among respondents aged 65 and over minus that among respondents aged 18–44, within each wave’s common sample. Positive values indicate higher utilisation among older respondents. These are descriptive, unadjusted percentage-point gaps.

Source: Author’s calculations using ISTAT AVQ microdata, 2018–2023.

3.1.5 Utilisation by Macro-Area

Table 3.6 reports the gap in utilisation between North-East and South, the conventional contrast in analyses of regional disparities in the Italian SSN and, in most of the six waves examined, also the macro-areas recording the highest and the lowest utilisation of specialist visits and diagnostic examinations. The gap is large and stable for both services, between 13 and 18 percentage points throughout. Blood tests and the emergency department show a smaller and more variable North-East figure over South, and their highest-utilisation area is not consistently North-East: for blood tests it alternates between North-East and Centre, while the lowest-utilisation area alternates between South and the Islands.

Table 3.6: Utilisation gap between North-East and South (percentage points)

	2018	2019	2020	2021	2022	2023
Specialist visits	+18.3	+14.6	+16.3	+16.8	+17.1	+15.9
Diagnostic examinations	+15.0	+13.1	+13.5	+14.2	+16.1	+16.4
Blood tests	+6.6	+7.4	+8.8	+7.5	+6.5	+12.1
Emergency department	+2.3	+2.5	+2.9	+2.3	+2.6	+3.6

Notes: Sampling-weighted utilisation rate in the North-East minus that in the South, using the ISTAT macro-area classification and the adult common sample. The South excludes the Islands. Positive values indicate higher utilisation in the North-East; gaps are descriptive and unadjusted for differences in need.

Source: Author's calculations using ISTAT AVQ microdata, 2018–2023.

3.1.6 Sensitivity to the 2023 Sample Definition

The unusually large number of missing values on the 2023 emergency-department item raises the possibility that the common-sample rule described in Section 2.1.3 could affect estimates for observations that remain usable for other outcomes. As a sensitivity check, the four 2023 models were therefore re-estimated on service-specific complete-case samples. The full comparison between the common and service-specific specifications is reported in the Appendix, Table 4.1.

For the three programmable services, the outcome-specific samples recover just over one thousand observations each: the sample rises from 31,098 to 32,210 for specialist visits, 32,188 for blood tests and 32,205 for diagnostic examinations. The corresponding ENA estimates change only marginally, from 0.1243 to 0.1231 for specialist visits, from 0.0932 to 0.0920 for blood tests, and from 0.1241 to 0.1237 for diagnostic examinations. Their sign, substantive magnitude and statistical significance are unchanged.

For the emergency department, both specifications must exclude the 2,777 observations for which `PROSOC` itself is missing. The service-specific specification therefore recovers only the additional 85 observations that were excluded from the common sample because of missing values on other healthcare-use items, raising the sample from 31,098 to 31,183. The need-adjusted estimate remains close to zero and statistically indistinguishable from zero: ENA changes from -0.0005 ($p = 0.903$) in the common sample to -0.0026 ($p = 0.500$) in the service-specific sample. The unadjusted Erreygers index is somewhat more sensitive, changing from -0.0072 ($p = 0.077$) to -0.0094 ($p = 0.016$). The central conclusion on horizontal inequity in 2023 is therefore robust to the sample definition, while the inference for the unadjusted concentration index is more sensitive.

3.2 National Indices: EI and ENA

3.2.1 Levels and Ordering Across Services

Each estimate reported in this section measures the same underlying quantity: the extent to which two individuals with identical need for a given healthcare service, but different socioeconomic position, have a different probability of obtaining it, the horizontal equity criterion set out in Section 2.4.1. Table 3.7 reports the Erreygers concentration index $E(y)$ and the need-adjusted index ENA at the national level for each of the four services and six waves, together with bootstrap standard errors for ENA. All ENA estimates are statistically distinguishable from zero at the 5 per cent level except for the emergency department, whose ENA is statistically distinguishable from zero in 2018, 2019 and 2021 but not in 2020, 2022 or 2023, as discussed in Section 3.2.4.

Table 3.7: National concentration indices, 2018–2023

	2018	2019	2020	2021	2022	2023
<i>Specialist visits</i>						
$E(y)$	0.055	0.040	0.077	0.071	0.045	0.058
ENA	0.134	0.123	0.137	0.137	0.113	0.124
(SE)	(0.007)	(0.007)	(0.008)	(0.007)	(0.008)	(0.009)
<i>Diagnostic examinations</i>						
$E(y)$	0.057	0.039	0.075	0.072	0.051	0.068
ENA	0.130	0.110	0.130	0.131	0.113	0.124
(SE)	(0.008)	(0.008)	(0.007)	(0.008)	(0.007)	(0.007)
<i>Blood tests</i>						
$E(y)$	0.007	0.013	0.031	0.022	0.013	0.027
ENA	0.080	0.087	0.092	0.093	0.087	0.093
(SE)	(0.007)	(0.007)	(0.009)	(0.008)	(0.007)	(0.007)
<i>Emergency department</i>						
$E(y)$	−0.021	−0.024	−0.005	−0.012	−0.009	−0.007
ENA	−0.013	−0.014	−0.000	−0.007	−0.003	−0.001
(SE)	(0.004)	(0.004)	(0.004)	(0.002)	(0.004)	(0.004)

Notes: Sampling-weighted estimates for the adult common sample. $E(y)$ is the unadjusted Erreygers concentration index; ENA is its need-adjusted counterpart. Positive values indicate pro-rich concentration, negative values pro-poor concentration. Parentheses report ENA standard errors (SE) from 100 bootstrap replications stratified by socioeconomic status quintile, without municipality-level clustering. Values are rounded to three decimals; −0.000 denotes a negative estimate rounded to zero.

Source: Author’s calculations using ISTAT AVQ microdata, 2018–2023.

Across all six waves, the broad grouping of services is stable. Specialist visits and diagnostic examinations form the highest-inequity group, blood tests display a substantially smaller positive ENA, and the emergency department remains close to zero and negative in sign. No substantive ranking is inferred between specialist visits

and diagnostic examinations. In particular, the very small differences between their ENA point estimates in 2022 and 2023 are minor relative to the bootstrap standard errors reported in Table 3.7 and are not interpreted as substantively meaningful differences between the two services. The relevant substantive contrast is instead between programmable services, which show pro-rich inequity, and emergency access, which does not display the same pattern.

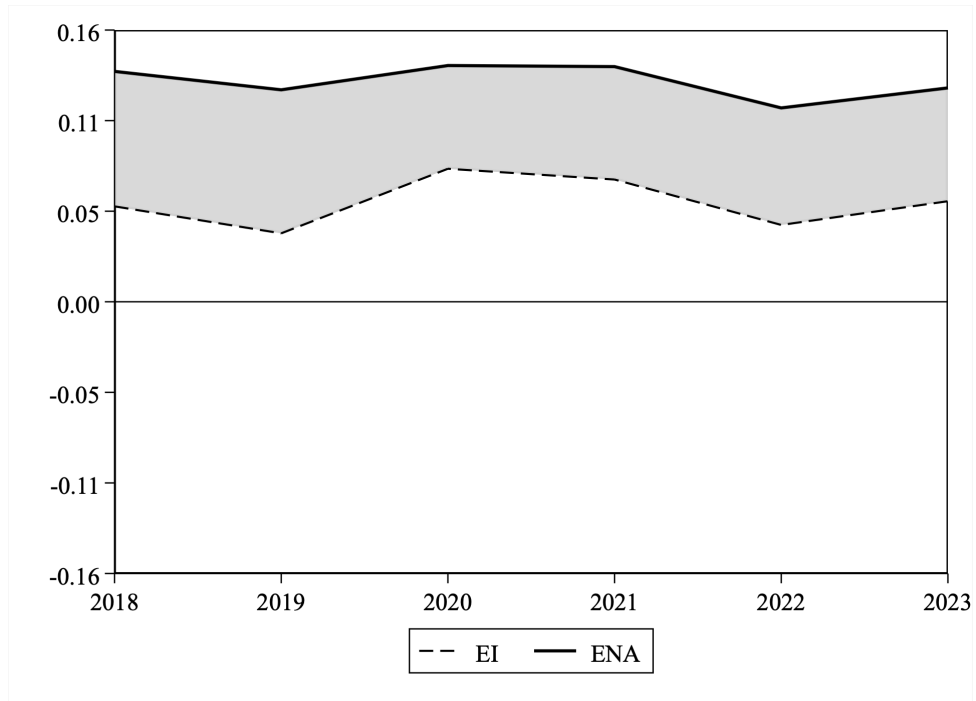
3.2.2 The Gap Between $E(y)$ and ENA

For the three programmable services, need standardisation raises the measured index substantially above its unadjusted level in every wave. Figure 3.1 plots $E(y)$ and ENA together for each service across the six waves. In 2018 the adjustment more than doubles the index for specialist visits and diagnostic examinations, and increases it more than tenfold for blood tests, from 0.007 to 0.080. As set out in Section 2.4.3, this gap arises because need is itself concentrated among the socioeconomically disadvantaged: subtracting the resulting negative need contributions from $E(y)$ removes a component that was masking, rather than justifying, the pro-rich distribution of utilisation.

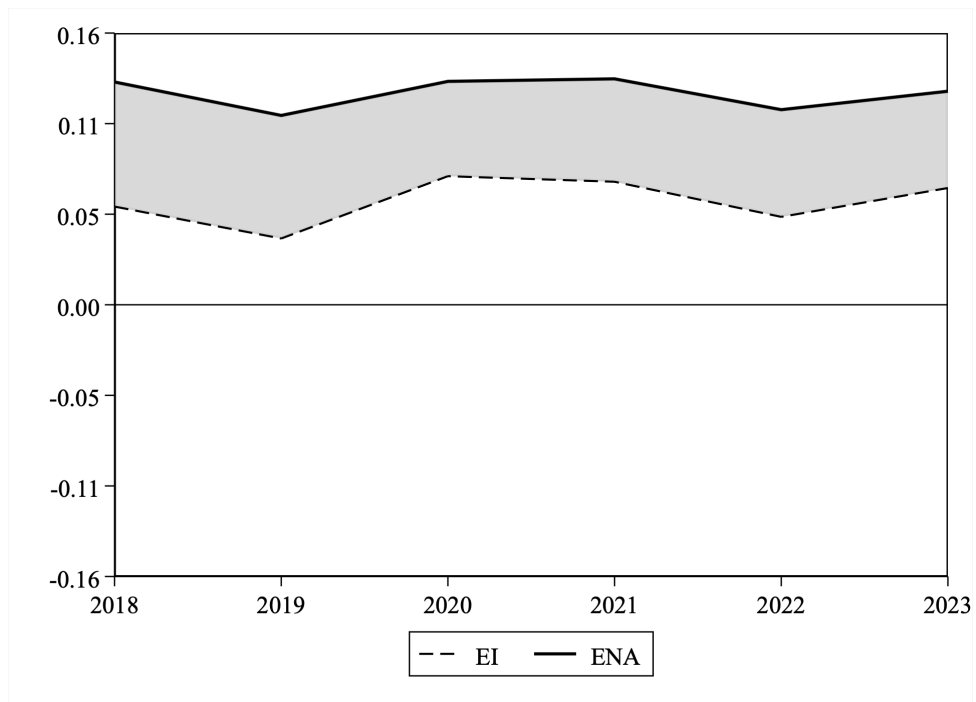
The gap is largest, in both absolute and relative terms, for blood tests. This follows directly from the descriptive pattern established in Section 3.1.3: because raw utilisation of blood tests is almost flat across the socioeconomic distribution, the unadjusted index is close to zero and gives little indication of the underlying inequity. Once utilisation is standardised for need — concentrated, as for the other services, among older and less healthy individuals who are themselves disproportionately represented among the socioeconomically disadvantaged — the flat raw distribution is revealed to conceal a pro-rich pattern of comparable magnitude to that observed for specialist visits and diagnostic examinations in several waves. A service that appears equitable on the basis of raw utilisation alone is not necessarily equitable once need is taken into account; this is the clearest illustration in the data of why the standardisation described in Section 2.4.2 is necessary rather than a technical refinement of secondary importance.

3.2.3 Trajectory, 2018–2023

Figure 3.2 plots the ENA series for the four services with 95 per cent confidence bands. For the three programmable services, the ENA point estimates are higher in 2020–2021 than in 2019 and lower again in 2022–2023. These movements are descriptive: the bootstrap inference reported here tests each annual estimate against



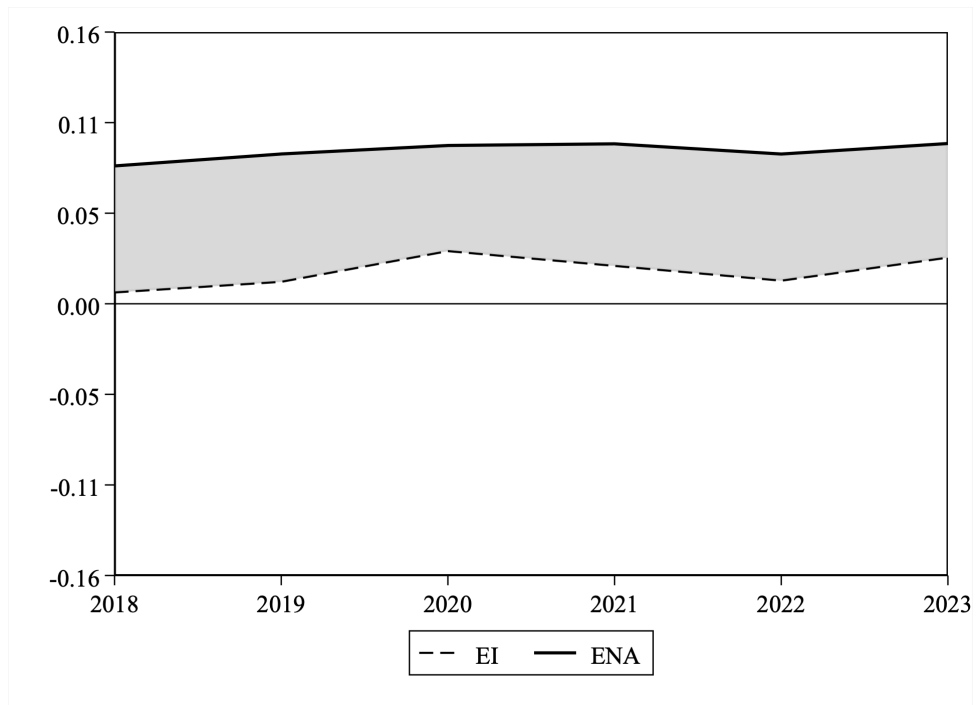
(a) Specialist visits



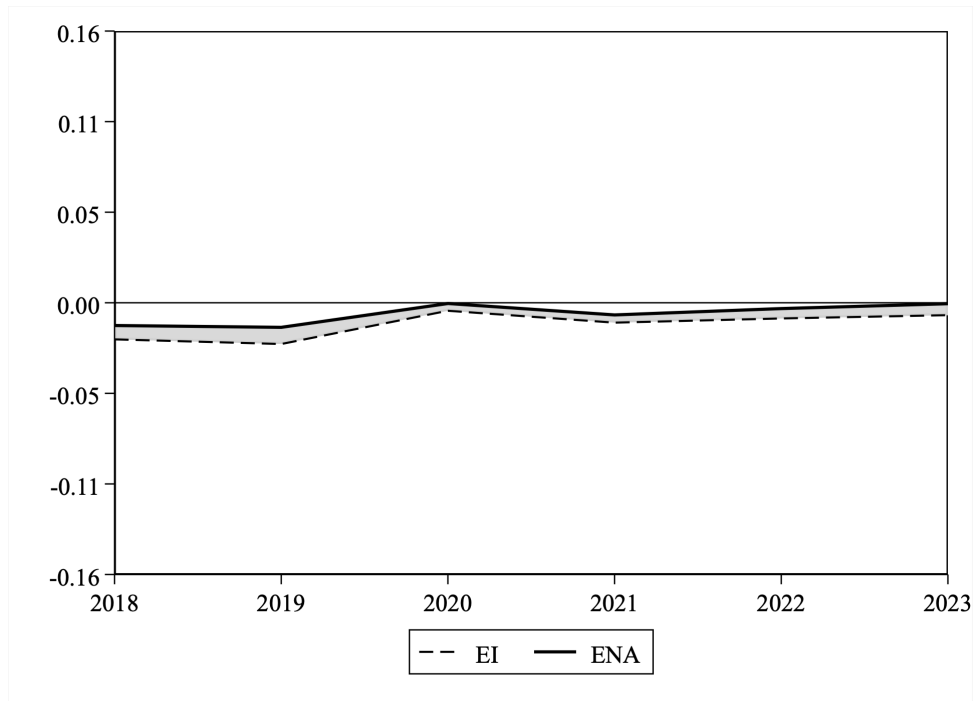
(b) Diagnostic examinations

Figure 3.1: Standard (EI) and need-adjusted (ENA) concentration index, by service, 2018–2023. Shaded areas mark the gap between the two indices. Panels (a) and (b); panels (c) and (d) follow on the next page.

Source: Author's calculations using ISTAT AVQ microdata, 2018–2023.



(c) Blood tests



(d) Emergency department

Figure 3.1: Standard (EI) and need-adjusted (ENA) concentration index, 2018–2023 (continued): panels (c) and (d). Shaded areas mark the gap between the two indices.

Source: Author's calculations using ISTAT AVQ microdata, 2018–2023.

zero, not the difference between estimates from two different waves, and no formal year-to-year difference test is conducted. The higher 2020 point estimates coincide with the lower raw-utilisation point estimates documented in Section 3.1.1; the two patterns are examined jointly, through the decomposition of the index, in Section 3.3.

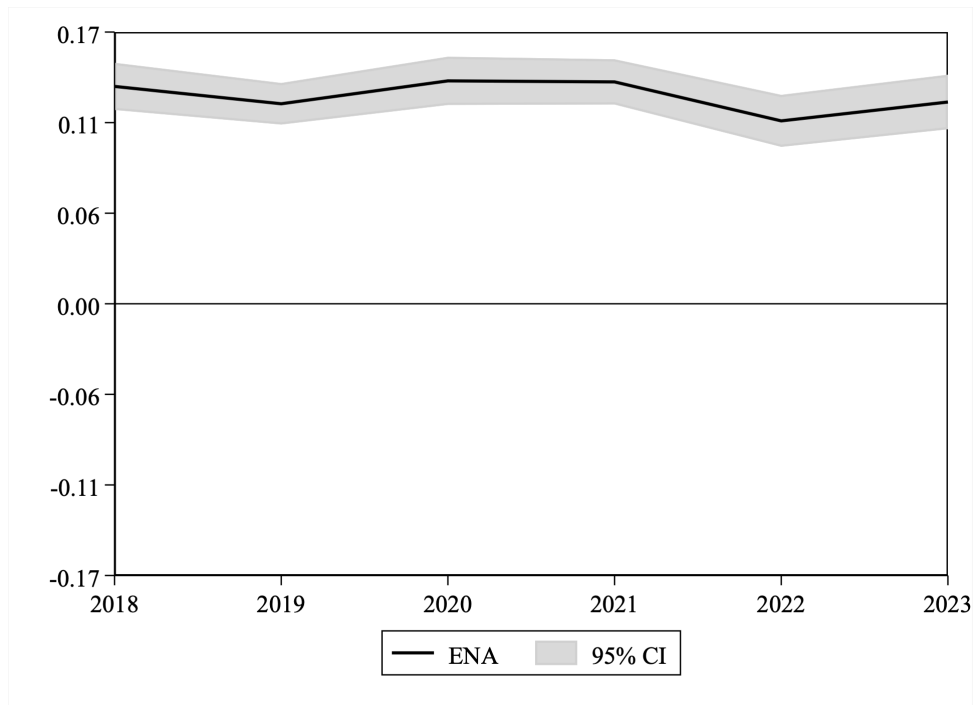
3.2.4 The Emergency Department

The emergency department is the only service for which ENA is not consistently distinguishable from zero: it is significant at the 1 per cent level in 2018, 2019 and 2021, and not significant in 2020, 2022 or 2023. The point estimate is negative in every wave and remains close to zero relative to the other three services. This is consistent with emergency access having a different institutional profile from the programmable services, which are scheduled in advance and may involve co-payments or direct private payment.

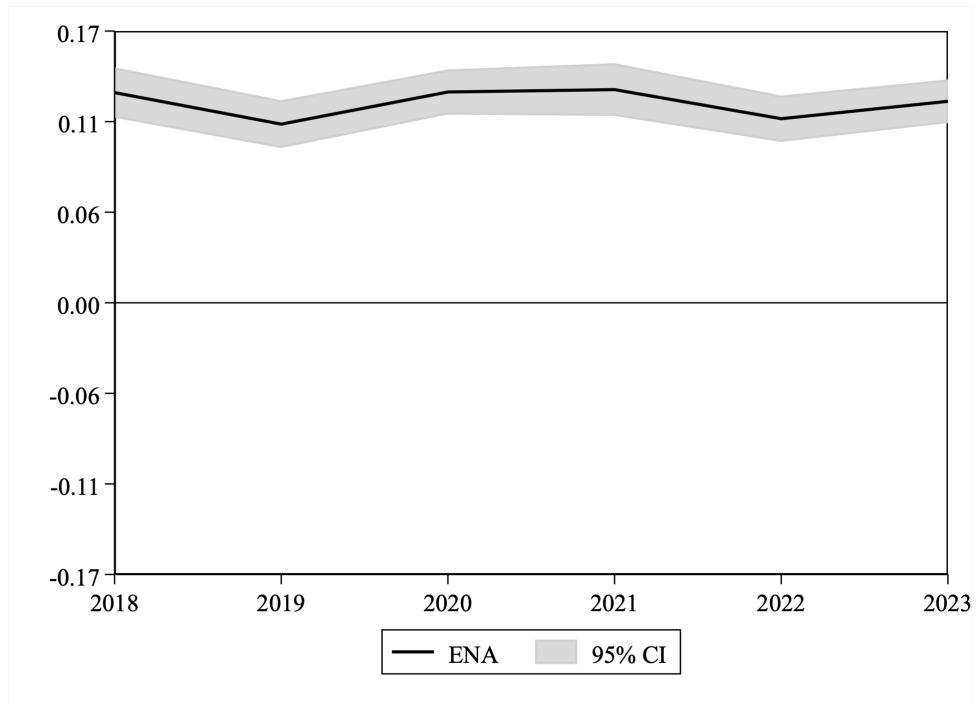
More notable than the point estimates is the changing significance status rather than any change in sign. In 2018, 2019 and 2021 the index is negative and significant, indicating a pro-poor distribution of emergency-department use. One possible interpretation is that some lower-income individuals use emergency care as a substitute for programmable care that is harder to obtain, but this mechanism is not identified by the data and is discussed only as a candidate interpretation in Chapter 4. In 2020, 2022 and 2023 the estimate is not reliably distinguishable from zero, although it remains negative. For 2023 this conclusion is unchanged when the emergency-department model is re-estimated on its service-specific sample, although the unadjusted index is more sensitive to the sample definition (Section 3.1.6).

3.3 Decomposition of the ENA Index

The decomposition set out in equation (2.3) is applied to each of the four services in every wave. Since ENA is the object of interest throughout this thesis, the decomposition presented here is anchored to ENA rather than to the unadjusted $E(y)$: the socioeconomic status contribution, the non-need contribution and the residual are reported as the three components that sum to ENA directly, following the definition in equation (2.4). The need contribution — age, sex, self-assessed health and chronic morbidity combined — is the block subtracted in moving from $E(y)$ to ENA; it is reported alongside as a memorandum item, since its own magnitude and composition are informative in their own right and are central to the mechanism examined in Section 3.3.2. This section presents the full decomposition for specialist visits, summary results for diagnostic examinations and blood tests, and a detailed



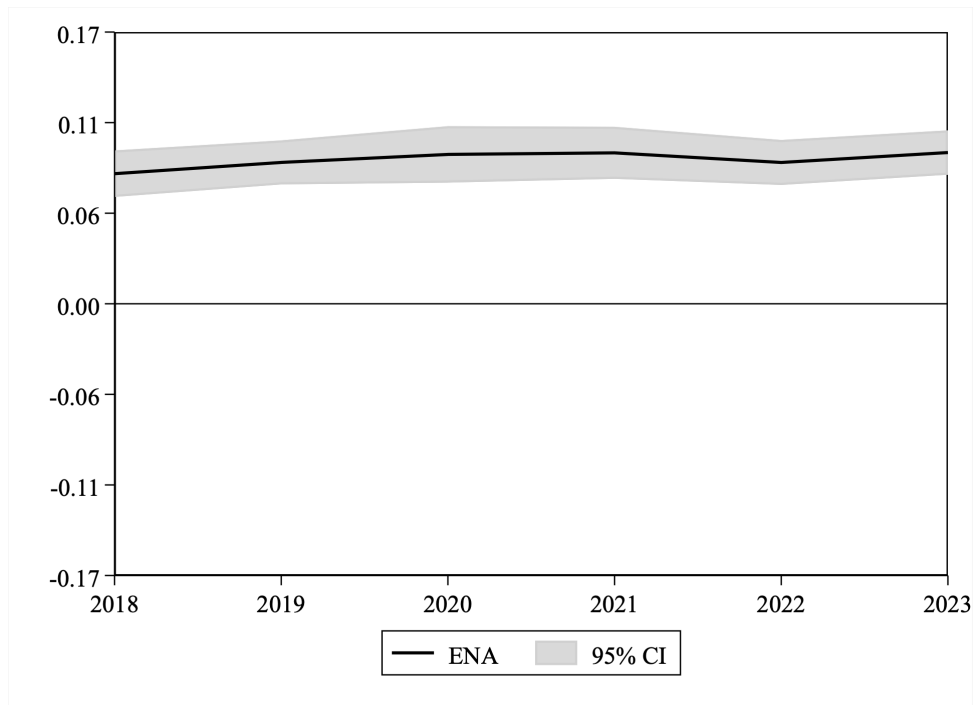
(a) Specialist visits



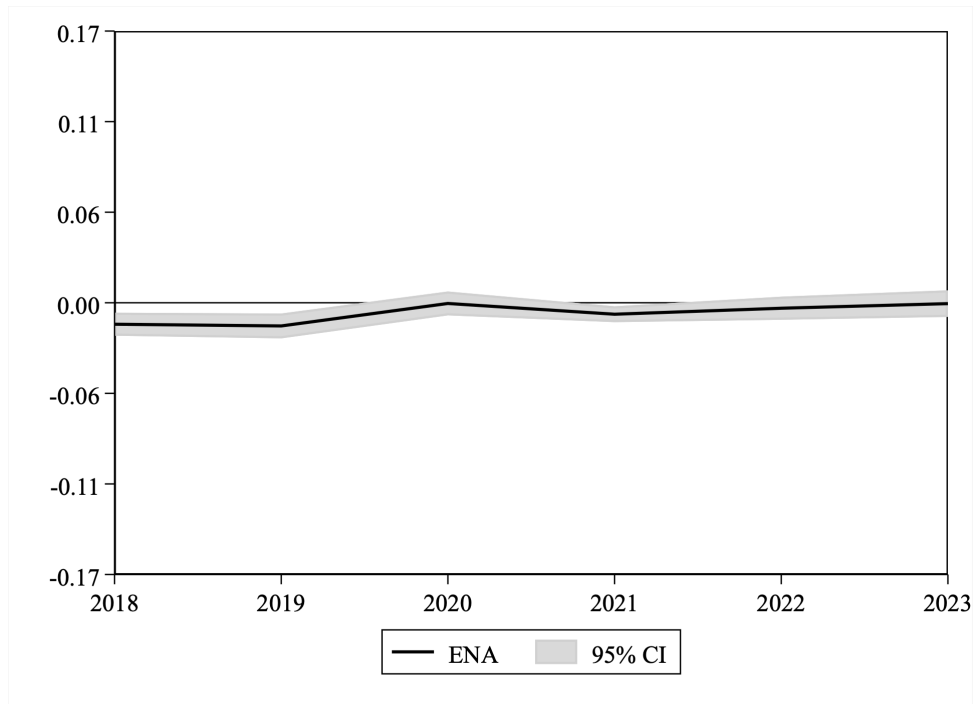
(b) Diagnostic examinations

Figure 3.2: Need-adjusted concentration index (ENA), with 95 per cent confidence bands, 2018–2023. Panels (a) and (b); panels (c) and (d) follow on the next page.

Source: Author's calculations using ISTAT AVQ microdata, 2018–2023.



(c) Blood tests



(d) Emergency department

Figure 3.2: Need-adjusted concentration index (ENA), with 95 per cent confidence bands, 2018–2023 (continued): panels (c) and (d).

Source: Author's calculations using ISTAT AVQ microdata, 2018–2023.

comparison with the emergency department, whose composition differs qualitatively from the other three.

3.3.1 What Generates the Level of ENA

Figure 3.3 decomposes ENA for each of the four services into the socioeconomic status contribution, the non-need contribution and the residual, with the need contribution shown as the memorandum item subtracted from $E(y)$ to obtain ENA. For specialist visits, the SES contribution is positive and of comparable magnitude to the non-need contribution in every wave, together accounting for the great majority of ENA; the residual is smaller and comparatively stable. The need contribution, shown for reference, is negative throughout and is discussed in detail below.

Table 3.8: Decomposition of ENA, specialist visits (national)

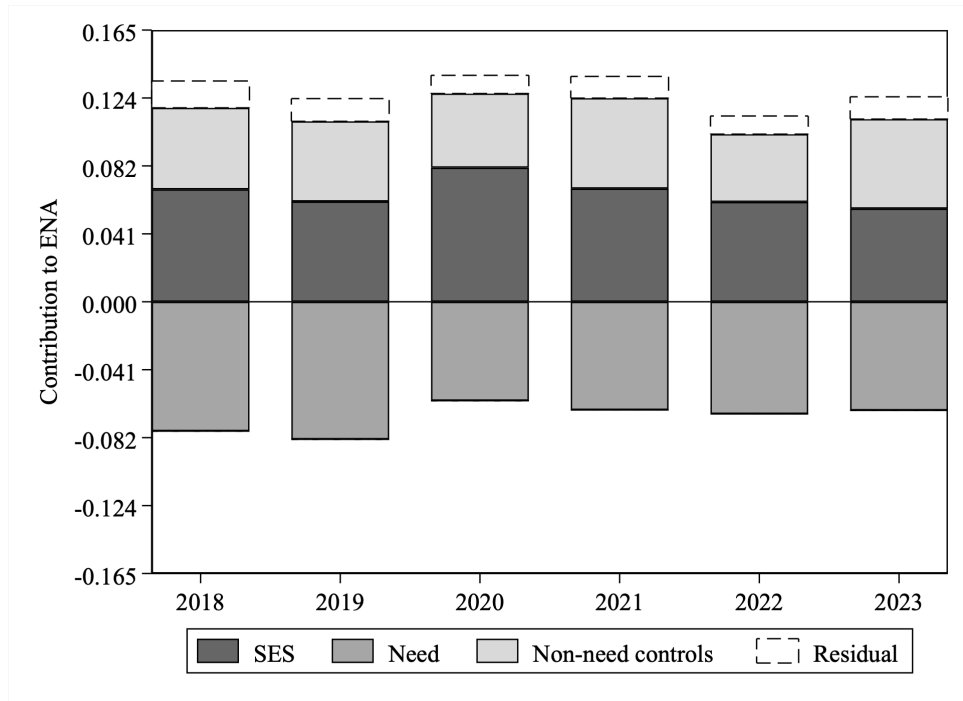
	2018	2019	2020	2021	2022	2023
SES	0.068	0.061	0.081	0.069	0.061	0.056
Non-need controls	0.049	0.049	0.045	0.055	0.041	0.054
Residual	0.016	0.014	0.011	0.013	0.011	0.014
ENA	0.134	0.123	0.137	0.137	0.113	0.124
<i>Memo: Need (subtracted)</i>	−0.078	−0.083	−0.060	−0.066	−0.068	−0.066
Age (65+)	−0.031	−0.034	−0.022	−0.021	−0.026	−0.024
Sex	−0.005	−0.006	−0.004	−0.005	−0.004	−0.004
SAH	−0.022	−0.023	−0.018	−0.021	−0.020	−0.019
Chronic	−0.019	−0.021	−0.016	−0.018	−0.018	−0.019
<i>Memo: E(y) (unadjusted)</i>	0.055	0.040	0.077	0.071	0.045	0.058

Notes: Sampling-weighted contributions in Erreygers index units, not percentage shares. SES denotes socioeconomic status and SAH self-assessed health. ENA is the need-adjusted index: the need block is subtracted from $E(y)$, not added to ENA. The residual includes regional-indicator contributions and unexplained inequality. Totals may differ due to rounding; contributions describe associations, not causal effects.

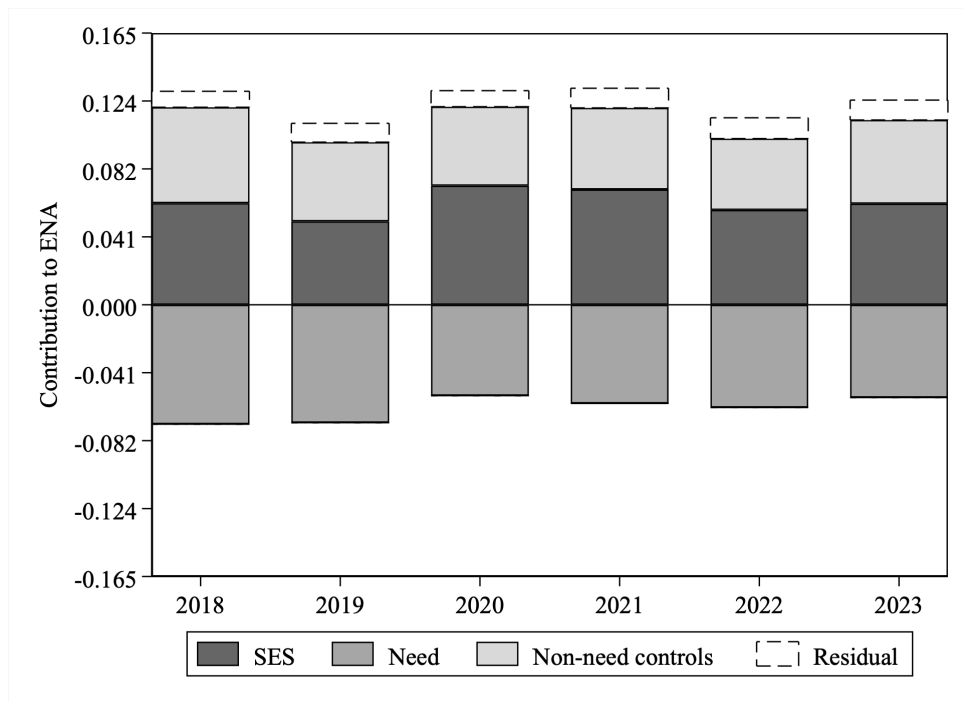
Source: Author's calculations using ISTAT AVQ microdata, 2018–2023.

Within the need block, self-assessed health and chronic morbidity make the largest and most stable contributions across the six waves. Age contributes consistently but with more variation, examined in Section 3.3.2. Sex contributes little in every wave, reflecting a weak association between being female and utilisation once the other need variables are held fixed, combined with a socioeconomic distribution of sex that is close to even; as set out in Section 2.5.2, both conditions must hold for a need variable to contribute little, and here they do.

Diagnostic examinations and blood tests display the same qualitative pattern: a positive and dominant SES contribution, a positive non-need contribution, and a



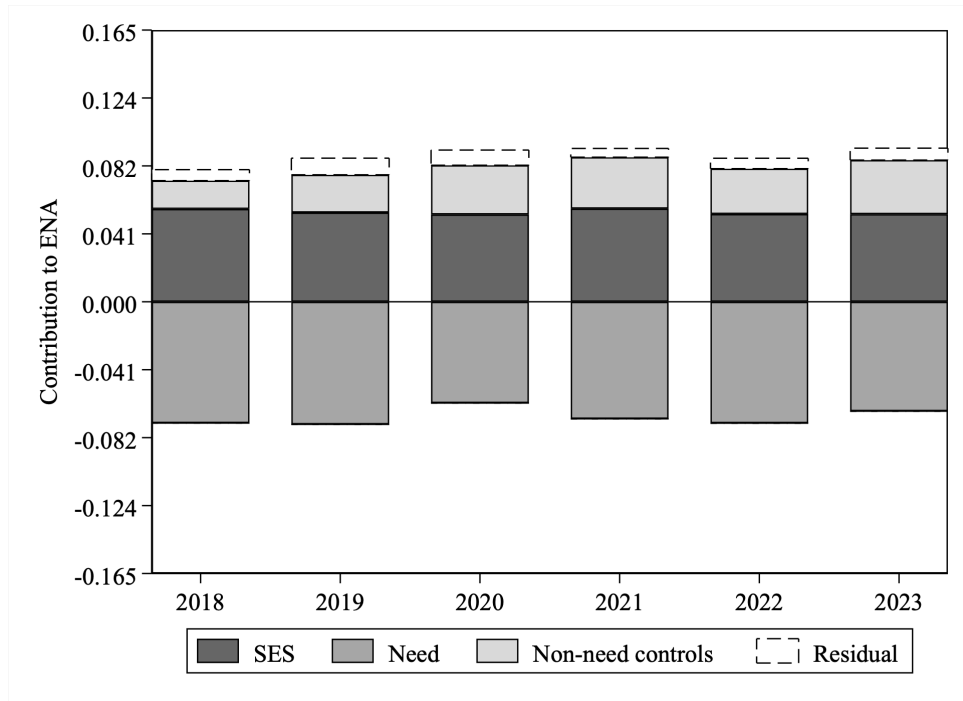
(a) Specialist visits



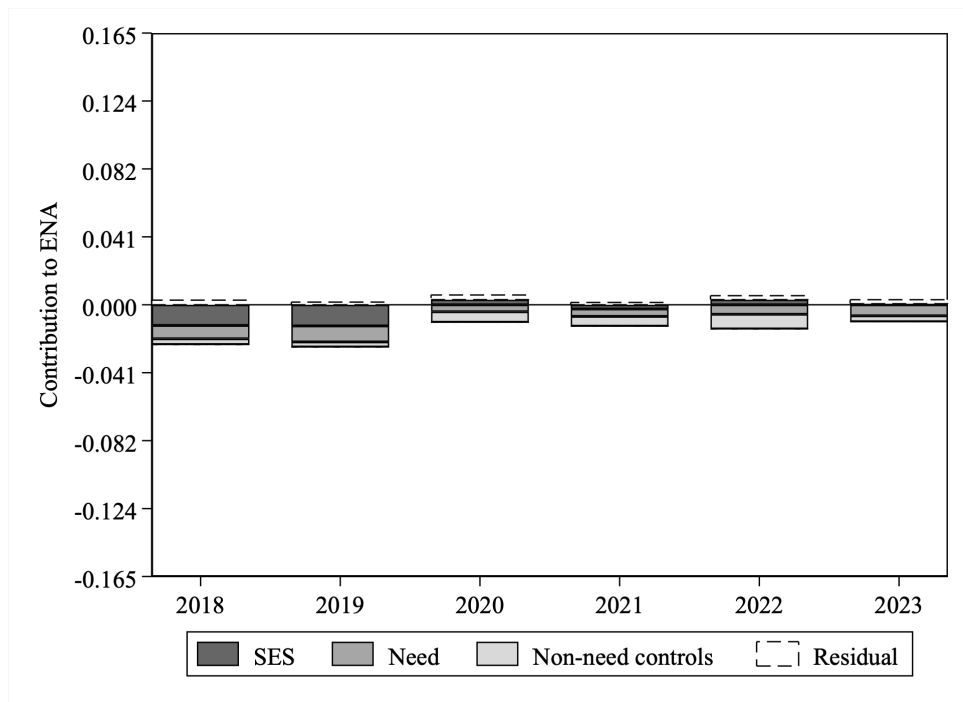
(b) Diagnostic examinations

Figure 3.3: Decomposition of ENA into SES, non-need and residual contributions, with the need contribution shown for reference, by service, 2018–2023. Panels (a) and (b); panels (c) and (d) follow on the next page.

Source: Author's calculations using ISTAT AVQ microdata, 2018–2023.



(c) Blood tests



(d) Emergency department

Figure 3.3: Decomposition of ENA into SES, non-need and residual contributions, with the need contribution shown for reference, 2018–2023 (continued): panels (c) and (d).

Source: Author's calculations using ISTAT AVQ microdata, 2018–2023.

need contribution, shown as a memorandum item, of comparable magnitude to that of specialist visits (Table 3.9). The main quantitative difference is in the residual, which is smaller and more stable for diagnostic examinations than for the other two services, and in the non-need contribution for blood tests, which is consistently the smallest of the three programmable services — a service whose raw distribution across quintiles, established in Section 3.1.3, is nearly flat is one where non-need socioeconomic characteristics other than the ranking variable itself add comparatively little.

Table 3.9: Decomposition of ENA: block totals, diagnostic examinations and blood tests (national)

		2018	2019	2020	2021	2022	2023
Diagnostic exams	SES	0.062	0.051	0.072	0.070	0.058	0.061
	Non-need	0.058	0.048	0.048	0.049	0.043	0.051
	Residual	0.010	0.011	0.010	0.012	0.013	0.012
	<i>ENA</i>	<i>0.130</i>	<i>0.110</i>	<i>0.130</i>	<i>0.131</i>	<i>0.113</i>	<i>0.124</i>
	<i>Memo: Need</i>	−0.072	−0.072	−0.055	−0.060	−0.062	−0.056
Blood tests	SES	0.056	0.054	0.053	0.056	0.053	0.053
	Non-need	0.017	0.023	0.030	0.031	0.028	0.033
	Residual	0.007	0.010	0.009	0.005	0.006	0.007
	<i>ENA</i>	<i>0.080</i>	<i>0.087</i>	<i>0.092</i>	<i>0.093</i>	<i>0.087</i>	<i>0.093</i>
	<i>Memo: Need</i>	−0.074	−0.074	−0.061	−0.071	−0.074	−0.066

Notes: Sampling-weighted contributions in Erreygers index units. SES denotes socioeconomic status; ENA is the need-adjusted index. SES, non-need and residual contributions sum to ENA, subject to rounding. Need is the block subtracted from the unadjusted index and is shown for reference only. The residual includes regional-indicator contributions and unexplained inequality; individual need components are not reported here.

Source: Author’s calculations using ISTAT AVQ microdata, 2018–2023.

3.3.2 Accounting for the 2020 Point-Estimate Difference

The higher ENA point estimate in 2020 documented in Section 3.2.3 is not associated with a uniform movement in every contribution. For specialist visits, the SES contribution point estimate is higher, at 0.081 in 2020 compared with 0.061 in 2019, while the age contribution point estimate is less negative, at −0.022 compared with −0.034.

As stressed in Section 2.5.2, a difference in a contribution across waves cannot by itself be interpreted as a difference in the predictive role of the underlying covariate, because the contribution combines the marginal effect, the covariate mean and its concentration index. Table 3.10 therefore reports these three components separately for the 2019–2020 comparison in specialist visits.

Table 3.10: Components of the age (65+) contribution, specialist visits, 2019–2020

	2019	2020
Marginal effect of age 65+	0.134	0.091
Weighted mean of age 65+	0.277	0.284
Concentration index of age 65+	−0.230	−0.212
Contribution	−0.034	−0.022

Notes: Components from the sampling-weighted specialist-visit model. Age 65+ is a binary indicator; its mean is a proportion and its concentration index describes its distribution across the socioeconomic ranking. The contribution to the unadjusted Erreygers index is four times the product of the marginal effect, mean and concentration index (Section 2.4.2). Cross-wave differences are descriptive, not formally tested.

Source: Author’s calculations using ISTAT AVQ microdata, 2019–2020.

The arithmetic difference is associated primarily with the lower marginal-effect point estimate for being aged 65 or over, 0.091 in 2020 compared with 0.134 in 2019. The weighted prevalence of the 65+ group is similar in the two waves and is slightly higher in 2020 (0.284 versus 0.277), so it cannot account for the smaller contribution. The age concentration index is also moderately less negative in 2020 (−0.212 versus −0.230), which contributes arithmetically to the smaller absolute contribution. No formal test of the 2019–2020 difference in these component estimates is conducted, so this is an accounting of the difference in point estimates rather than evidence of a statistically significant structural change in the role of age.

For diagnostic examinations and blood tests, Table 3.9 reports only the aggregate need block. Its point estimate is less negative in 2020 than in 2019, moving from −0.072 to −0.055 for diagnostic examinations and from −0.074 to −0.061 for blood tests. Since the individual need components are not reported for these two services, this movement is not attributed specifically to age.

Because the need contribution is subtracted in constructing ENA (equation (2.4)), the less negative age contribution for specialist visits mechanically raises the need-adjusted index. It therefore contributes arithmetically to the higher 2020 ENA point estimate for this service, alongside the higher SES contribution, without by itself identifying a causal mechanism.

The residual reported in Tables 3.8 and 3.9 comprises the contribution of the regional indicators, included in the utilisation model but not decomposed individually, together with the portion of inequality the model does not explain (Section 2.5.3). It is modest and stable across the six waves for all three programmable services and is not interpreted further here.

Table 3.11: Decomposition of ENA, emergency department (national)

	2018	2019	2020	2021	2022	2023
SES	−0.013	−0.013	0.003	−0.003	0.003	0.001
Non-need controls	−0.003	−0.003	−0.006	−0.006	−0.009	−0.004
Residual	0.003	0.002	0.003	0.001	0.002	0.002
ENA	−0.013	−0.014	−0.000	−0.007	−0.003	−0.001
<i>Memo: Need (subtracted)</i>	−0.008	−0.010	−0.004	−0.004	−0.006	−0.007
<i>Memo: E(y) (unadjusted)</i>	−0.021	−0.024	−0.005	−0.012	−0.009	−0.007

Notes: Sampling-weighted contributions in Erreygers index units. SES denotes socioeconomic status; ENA is the need-adjusted index. The need block is subtracted from $E(y)$ and shown for reference only. The residual includes regional-indicator contributions and unexplained inequality. Totals may differ due to rounding; −0.000 denotes a negative estimate rounded to zero. Component significance is not reported in this table.

Source: Author’s calculations using ISTAT AVQ microdata, 2018–2023.

3.3.3 The Emergency Department: a Different Composition

The composition for the emergency department differs from that of the other three services in a way that is consistent with the sign pattern documented in Section 3.2.4. The SES contribution, positive for the programmable services in every wave, is negative or negligible here throughout the period. In 2021, the wave for which contribution-level standard errors are examined in detail, the SES contribution is −0.0048 with a bootstrap standard error of 0.0025 ($p = 0.059$), and the self-assessed health contribution is −0.0050 (SE 0.0005, $p < 0.001$); age and sex contribute negligibly and are not distinguishable from zero (point estimates of 0.0003 and 0.0002 respectively, against standard errors of 0.0009 and 0.0002). The pattern that emerges is one in which self-assessed health, rather than age, is the principal need-related driver of the pro-poor association for this service, while the SES contribution itself is comparatively imprecise and only marginally significant. As in the aggregate index, this reflects a matter of degree rather than a different underlying mechanism: needier and less well-off individuals remain, on average, associated with somewhat higher emergency department use throughout the period, but several of the individual contributions that make up this association are small relative to their sampling variability, which is consistent with ENA itself losing statistical significance in three of the six waves.

3.4 Regional Estimates

3.4.1 Precision at the Regional Level

Regional estimates carry substantially wider standard errors than the national series discussed in Section 3.2, a consequence of the reduced sample available for each Region relative to the country as a whole (Section 2.6.3). The gap in precision can be quantified directly using the reliability ratio defined in Section 2.6.4: the share of the variance observed across the twenty Regions in a given wave that is not attributable to sampling error, averaged over the six waves for each service.

Table 3.12: Reliability of regional ENA estimates

	SD across Regions	Average SE	Reliability
Specialist visits	0.039	0.033	0.30
Diagnostic examinations	0.033	0.032	0.04
Blood tests	0.042	0.033	0.39
Emergency department	0.018	0.016	0.16

Notes: ENA is the need-adjusted Erreygers index; SD denotes standard deviation and SE bootstrap standard error. Statistics summarise the twenty regional estimates over six waves. Reliability is the estimated share of between-Region variance remaining after subtracting sampling-error variance, bounded below by zero. It uses the six-wave averages of variances and squared SEs, as defined in Section 2.6.4, not the squared average SE shown here.

Source: Author’s calculations using ISTAT AVQ microdata, 2018–2023.

The reliability ratio ranges from 0.04 for diagnostic examinations to 0.39 for blood tests. Under the reliability approximation, only around 4 per cent of the dispersion visible across the twenty Regions for diagnostic examinations is estimated to reflect between-Region variation beyond sampling error; the remaining share is consistent with sampling variability. Even for blood tests, the service with the highest reliability ratio, a substantial share of the observed dispersion cannot be distinguished from sampling noise. This has a direct consequence for how the estimates that follow should be read: individual regional rankings, and differences between specific pairs of Regions, are not in general statistically robust, and the discussion below is confined to patterns that are stable across multiple waves or that involve differences large relative to the standard errors reported.

3.4.2 Regional Rankings and the North–South Question

Table 3.13 reports ENA for specialist visits by Region in 2018, 2021 and 2023, the three waves selected in Section 2.6.1 to span the observation window. The individual regional point estimates do not display a clear North–South ordering in the three waves shown. For example, in 2018 Basilicata (0.164) and Calabria

(0.162) — both Southern — record the two highest point estimates in the country, with Toscana (0.149) third, while other Southern Regions lie closer to the national estimate. Sicilia’s point estimate also changes its relative position across the reported waves. These rankings are descriptive only: they are not a formal comparison of average inequity across macro-areas and cannot establish that an underlying North–South gradient is absent.

Table 3.13: ENA by Region, specialist visits, 2018, 2021, 2023

Region	2018	2021	2023
Piemonte	0.139***	0.107***	0.131***
Valle d’Aosta	−0.013	0.052	0.000
Lombardia	0.119***	0.145***	0.110***
Trentino-Alto Adige	0.106***	0.111***	0.065**
Veneto	0.076***	0.093***	0.048*
Friuli Venezia Giulia	0.144***	0.199***	0.069*
Liguria	0.089***	0.095***	0.071**
Emilia-Romagna	0.095***	0.117***	0.093***
Toscana	0.149***	0.091***	0.134***
Umbria	0.076	0.031	0.076
Marche	0.053	0.123***	0.090***
Lazio	0.139***	0.081***	0.119***
Abruzzo	0.064*	0.044	0.047
Molise	0.126***	0.042	0.010
Campania	0.093***	0.153***	0.083***
Puglia	0.106***	0.134***	0.143***
Basilicata	0.164***	0.144***	0.083**
Calabria	0.162***	0.124***	0.216***
Sicilia	0.082***	0.133***	0.123***
Sardegna	0.092***	0.125***	0.125***

Notes: Sampling-weighted need-adjusted Erreygers indices (ENA), with socioeconomic ranks defined within each Region. Positive values indicate pro-rich inequity. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ test each estimate against zero using 100 bootstrap replications stratified by socioeconomic status quintile, without municipality-level clustering. Stars do not test differences between Regions or years.

Source: Author’s calculations using ISTAT AVQ microdata, 2018, 2021 and 2023.

Table 3.14 reports the same three waves for diagnostic examinations. As for specialist visits, no Region is above the national estimate in all three waves, and the relative positions of the regional point estimates vary considerably across the three reported waves. Friuli Venezia Giulia records the highest point estimate in the country in both 2018 (0.196) and 2021 (0.209), whereas its 2023 point estimate is 0.095 — thirteenth of twenty Regions. Sardegna shows the opposite descriptive ordering, from eighteenth in 2018 (0.069) to the highest point estimate in the country in 2023 (0.160). As with specialist visits, the individual-region point estimates do

not show a clear territorial ordering in the three waves reported; no formal macro-area comparison is performed.

Table 3.14: ENA by Region, diagnostic examinations, 2018, 2021, 2023

Region	2018	2021	2023
Piemonte	0.128***	0.115***	0.105***
Valle d'Aosta	0.056	0.009	0.047
Lombardia	0.099***	0.147***	0.118***
Trentino-Alto Adige	0.062**	0.062**	0.092***
Veneto	0.117***	0.103***	0.057**
Friuli Venezia Giulia	0.196***	0.209***	0.095**
Liguria	0.099***	0.090**	0.095***
Emilia-Romagna	0.130***	0.094***	0.115***
Toscana	0.142***	0.093***	0.133***
Umbria	0.105**	0.099*	0.052
Marche	0.076**	0.063*	0.115***
Lazio	0.122***	0.084***	0.129***
Abruzzo	0.062**	0.089**	0.029
Molise	0.127***	0.043	0.060
Campania	0.084***	0.123***	0.105***
Puglia	0.122***	0.147***	0.110***
Basilicata	0.102**	0.077**	0.036
Calabria	0.152***	0.133***	0.105***
Sicilia	0.079***	0.123***	0.097***
Sardegna	0.069**	0.111***	0.160***

Notes: Sampling-weighted need-adjusted Erreygers indices (ENA), with socioeconomic ranks defined within each Region. Positive values indicate pro-rich inequity. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ test each estimate against zero using 100 bootstrap replications stratified by socioeconomic status quintile, without municipality-level clustering. Stars do not test differences between Regions or years.

Source: Author's calculations using ISTAT AVQ microdata, 2018, 2021 and 2023.

3.4.3 Stability and Volatility Across Waves

Reading across the three columns of Table 3.13 illustrates the consequence of the low reliability documented in Section 3.4.1. No Region is above the national estimate in all three waves shown, and several — including Veneto, Liguria, Emilia-Romagna, Umbria and Marche, spanning both Northern and Central Italy — are below it in every wave. Individual regional point estimates vary considerably from one wave to the next. Calabria ranges from 0.124 in 2021 to 0.216 in 2023, the highest point estimate recorded anywhere in the table; Molise is 0.123 in 2018 and a non-significant 0.021 in 2023. Given that Calabria and Molise are, respectively, a mid-sized and one of the smallest Regions by sample, this volatility is consistent with the low reliability

quantified above rather than with a genuine reversal of regional position, and neither trajectory is interpreted as a substantive finding on its own.

Valle d'Aosta records estimates that are never statistically significant in any of the three waves shown, the only Region for which this is the case. Given its small regional sample and the associated uncertainty, this should not be interpreted as evidence of an absence of inequity.

3.4.4 Consistency with the National Pattern

Despite the volatility documented above, the regional estimates are broadly consistent with the national series. The great majority of significant regional estimates are positive, matching the pro-rich distribution found at the national level in Table 3.7, and no Region records a significant negative estimate for specialist visits in any of the three waves shown. The regional dimension therefore adds evidence on the heterogeneity and imprecision surrounding the national estimate, rather than qualifying its sign or general magnitude.

3.4.5 Blood Tests and the Emergency Department

The regional pattern for blood tests follows that of specialist visits and diagnostic examinations, while emergency department access does not. Averaging across the three waves reported in Tables 3.13 and 3.14, blood-test point estimates range from close to zero in Valle d'Aosta and Basilicata to 0.114 in Calabria. The ranking does not display an obvious North–South ordering, but this descriptive observation is not a test of a macro-area gradient. It is also subject to the same caveat regarding reliability discussed above, given the 0.39 ratio reported in Table 3.12 for this service.

Emergency department access differs. Regional estimates cluster narrowly around zero, mostly between -0.03 and 0.00 , consistent with the near-neutral national estimate documented in Section 3.2.4; Valle d'Aosta is the only Region with an average estimate below -0.04 , and Trentino-Alto Adige is the only one with an average at or above zero. Given the reliability ratio of 0.16 for this service, the estimates provide little evidence that any individual Region departs sharply from the national pattern and should not be used to infer fine-grained regional differences.

Figures 3.4 and 3.5 present the specialist-visit and diagnostic-examination estimates graphically, with 95 per cent confidence intervals, allowing the overlap between Regions to be assessed directly rather than inferred from the significance stars in Tables 3.13 and 3.14 alone.

Read together, the results in this section show positive pro-rich ENA estimates

for specialist visits and diagnostic examinations in many Regions, although the precision of individual regional estimates is limited. The regional rankings themselves do not reveal a clear North–South gradient in the waves reported. This should not be interpreted as evidence that such a gradient does not exist: establishing or rejecting a territorial gradient would require a formal comparison of average inequity across macro-areas, which is not undertaken here. The regional evidence is therefore used to describe heterogeneity in point estimates, not to make a formal claim about the existence or absence of a North–South divide.

3.5 Forgone Care

The concentration indices reported in the preceding sections measure inequity in care that is actually received; forgone care is the more direct counterpart, recording need that goes unmet altogether rather than need that is met unequally. The AVQ records forgone care only for specialist visits and diagnostic examinations (Section 2.2.6). This section reports the rate of forgone care for both services across the six waves, the reasons respondents give for forgoing care, and how these reasons vary by socioeconomic quintile, age class and macro-area.

3.5.1 Rates of Forgone Care

Table 3.15: Forgone care rate, by wave (per cent of the estimation sample)

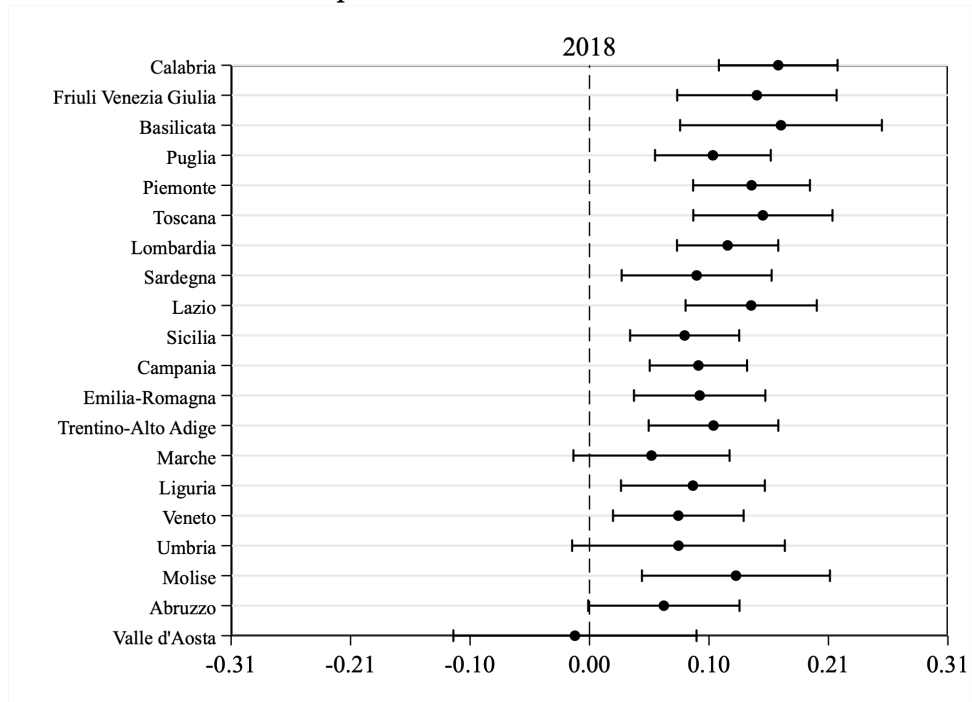
	2018	2019	2020	2021	2022	2023
Specialist visits	9.96	8.97	11.93	14.15	8.99	9.61
Diagnostic examinations	5.98	5.49	7.30	8.01	4.66	5.76

Notes: Sampling-weighted percentages of the adult common estimation sample reporting forgone care, not percentages of service users. The 2022 filter revision limits level comparisons to within 2018–2021 and within 2022–2023 (Section 2.3.2). Cross-wave differences are descriptive and have not been formally tested.

Source: Author’s calculations using ISTAT AVQ microdata, 2018–2023.

The weighted forgone-care point estimates are higher in 2020–2021 than in 2019 for both services, reaching around 14 and 8 per cent respectively in 2021, and are lower again from 2022. These are descriptive comparisons rather than formally tested year-to-year differences. The pattern is examined together with the reasons reported for forgoing care in the following subsection, since the higher 2020–2021 estimates coincide with a larger residual share of respondents for whom no separately coded reason is identified.

(a) Regional ENA estimates
Specialist visits — 2018



(b) Regional ENA estimates
Specialist visits — 2021

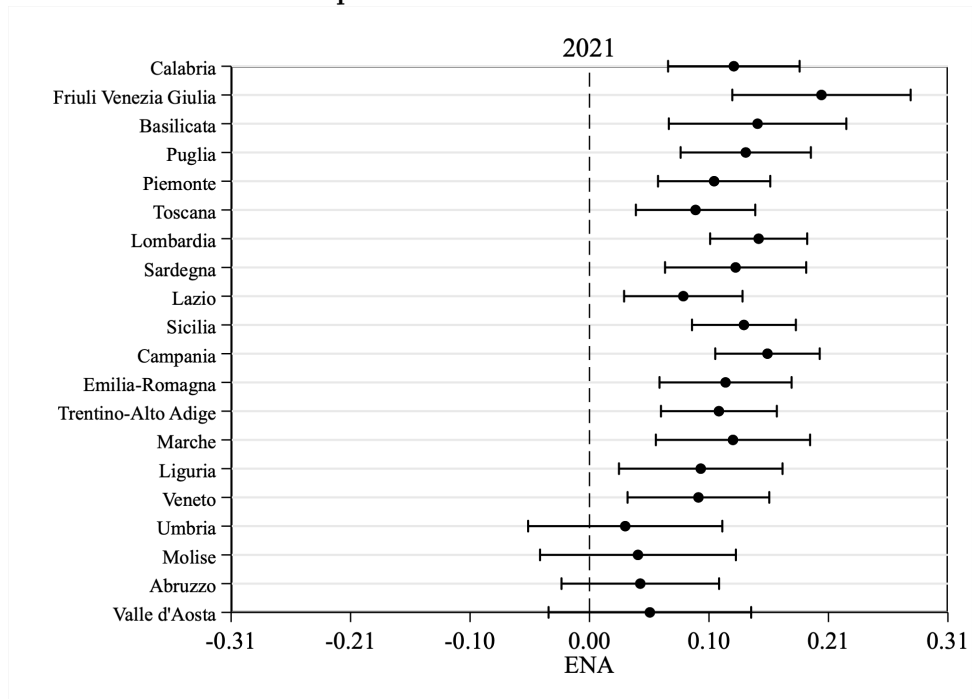


Figure 3.4: Regional ENA for specialist visits with 95 per cent confidence intervals, 2018, 2021 and 2023. Panels (a) and (b); panel (c) follows on the next page.

Source: Author's calculations using ISTAT AVQ microdata, 2018, 2021 and 2023.

(c) Regional ENA estimates
Specialist visits — 2023

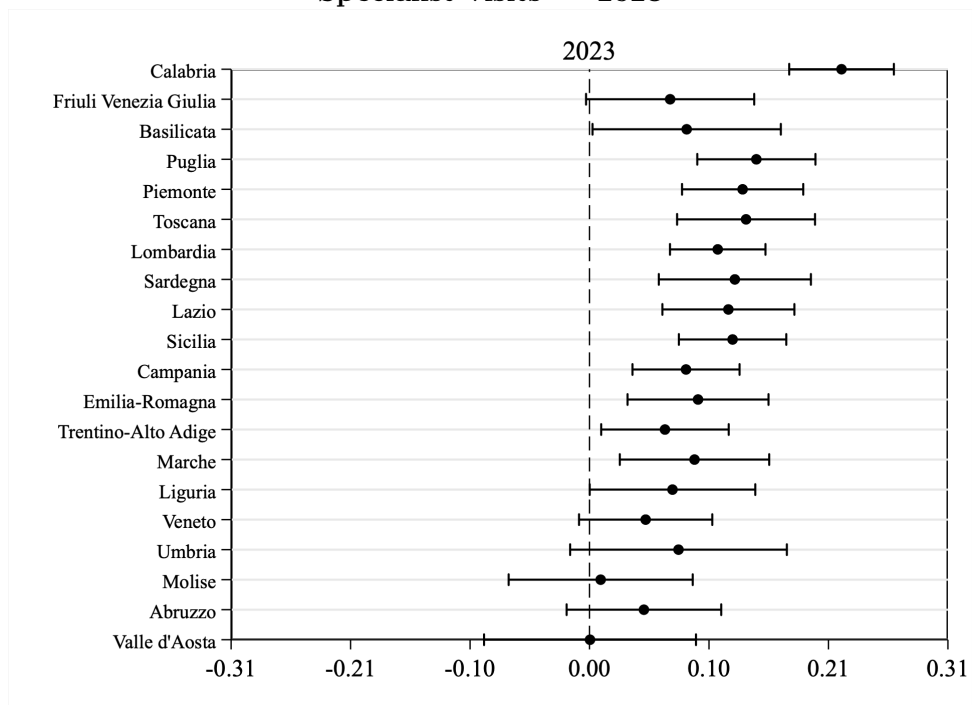
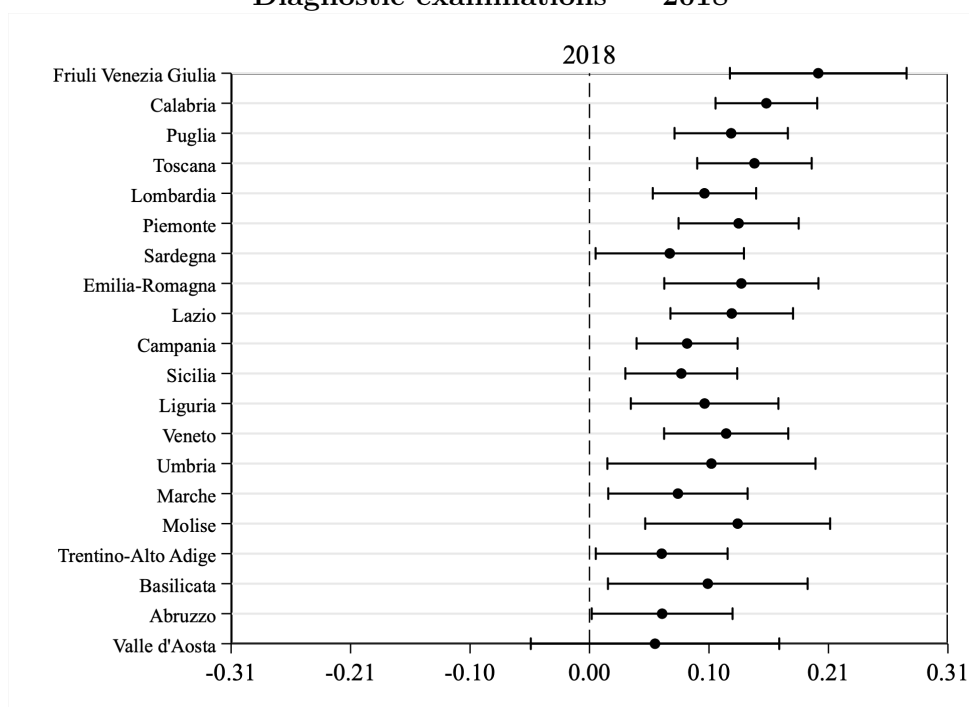


Figure 3.4: Regional ENA for specialist visits with 95 per cent confidence intervals (continued): panel (c), 2023.

Source: Author's calculations using ISTAT AVQ microdata, 2018, 2021 and 2023.

(a) Regional ENA estimates
Diagnostic examinations — 2018



(b) Regional ENA estimates
Diagnostic examinations — 2021

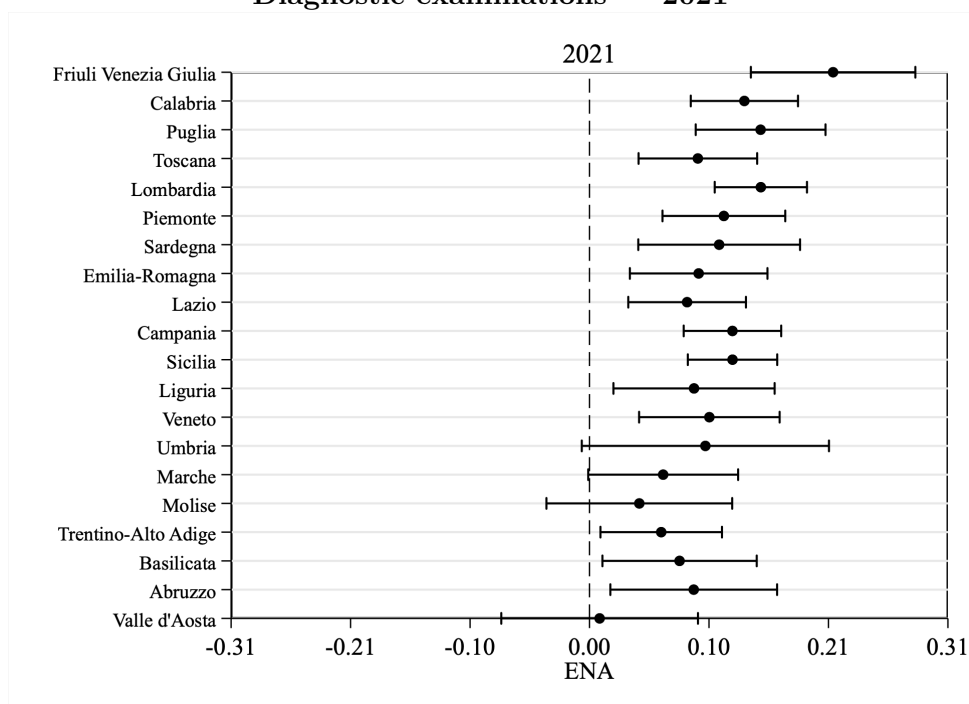


Figure 3.5: Regional ENA for diagnostic examinations with 95 per cent confidence intervals, 2018, 2021 and 2023. Panels (a) and (b); panel (c) follows on the next page.

Source: Author's calculations using ISTAT AVQ microdata, 2018, 2021 and 2023.

(c) Regional ENA estimates
Diagnostic examinations — 2023

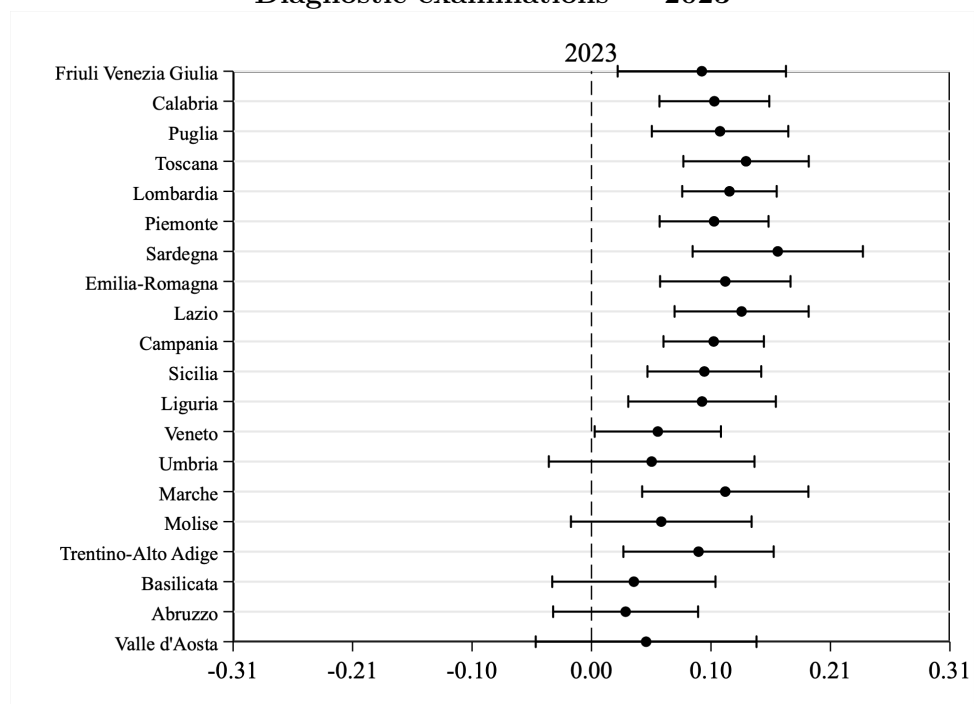


Figure 3.5: Regional ENA for diagnostic examinations with 95 per cent confidence intervals (continued): panel (c), 2023.

Source: Author's calculations using ISTAT AVQ microdata, 2018, 2021 and 2023.

3.5.2 Forgone Care by Quintile, Age and Macro-Area

The rate of forgoing care, and not only the reasons reported for it, varies with socioeconomic position, age and location.

Table 3.16 reports the gap in the forgone care rate between the poorest and richest SES quintile for each of the six waves. The gap is negative throughout — the poorest quintile forgoes care at a higher rate than the richest — for both services in every wave, mirroring at the level of the rate itself the utilisation gradient documented in Section 3.1.3.

Table 3.16: Forgone care rate gap between the richest and poorest SES quintile (Q5–Q1, percentage points)

	2018	2019	2020	2021	2022	2023
Specialist visits	–4.0	–3.7	–1.6	–2.9	–1.9	–4.0
Diagnostic examinations	–1.6	–2.2	–0.3	–1.4	–1.4	–2.2

Notes: Sampling-weighted forgone-care rate in the highest socioeconomic status quintile (Q5) minus that in the lowest (Q1), using all common-sample respondents in each group as the denominator. Negative values indicate higher forgone care in Q1. Gaps are descriptive; the 2022 filter revision limits comparisons across coding periods.

Source: Author’s calculations using ISTAT AVQ microdata, 2018–2023.

Table 3.17 reports the forgone care rate by age class. In four of the six waves — 2018, 2019, 2022 and 2023 — the highest rate is recorded by the 45–64 age class, not by respondents aged 65 and over, whose rate is closer to that of the youngest class. In 2020 and 2021 this ordering changes: the rate among respondents aged 65 and over rises above that of the 45–64 class and becomes the highest of the three, reaching 14.7 and 17.4 per cent for specialist visits. This descriptive pattern is consistent with unusual pandemic-period access and utilisation conditions among older respondents, but it is not used to infer that age became a weaker predictor of utilisation. The decomposition evidence in Section 3.3.2 concerns the 2019–2020 change in specialist visits and shows explicitly that the age contribution depends jointly on the age marginal effect, its weighted prevalence and its socioeconomic concentration.

Table 3.18 reports the gap in the forgone care rate between the Islands and North-East. Unlike the utilisation gap reported in Section 3.1.5, where South records the lowest utilisation, it is the Islands rather than the South that most frequently records the highest forgone care rate over the six waves examined, indicating that the geography of forgoing care does not simply mirror the geography of utilisation.

Table 3.17: Forgone care rate by age class (per cent)

	2018	2019	2020	2021	2022	2023
<i>Specialist visits</i>						
18–44	8.5	7.1	8.2	9.4	6.2	7.4
45–64	12.2	11.0	13.3	15.1	10.2	11.3
65+	8.5	8.1	14.7	17.4	10.1	10.0
<i>Diagnostic examinations</i>						
18–44	4.7	4.0	4.3	4.9	3.0	4.1
45–64	7.2	6.7	8.4	9.0	5.2	6.7
65+	5.7	5.1	9.3	9.7	5.4	6.0

Notes: Sampling-weighted percentages reporting forgone care among all common-sample respondents in each age class and wave. Age is measured in completed years; 65+ includes respondents aged 65 and over. Differences are descriptive. Levels are comparable within, not across, the 2018–2021 and 2022–2023 coding periods (Section 2.3.2).

Source: Author’s calculations using ISTAT AVQ microdata, 2018–2023.

Table 3.18: Forgone care rate gap between the Islands and North-East (percentage points)

	2018	2019	2020	2021	2022	2023
Specialist visits	+3.0	+3.7	+0.1	+1.4	+2.1	+1.6
Diagnostic examinations	+2.2	+1.8	0.0	+1.2	+1.9	+2.0

Notes: Sampling-weighted forgone-care rate in the Islands (Sicilia and Sardegna) minus that in the North-East, using all common-sample respondents in each macro-area as the denominator. Positive values indicate higher forgone care in the Islands. Gaps are descriptive; the 2022 filter revision limits comparisons across coding periods.

Source: Author’s calculations using ISTAT AVQ microdata, 2018–2023.

Table 3.19: Reasons for forgoing care, national, by wave (per cent of those forgoing care; rows sum to more than 100 as multiple reasons may be reported)

	2018	2019	2020	2021	2022	2023
<i>Specialist visits</i>						
Cost	52.7	53.7	26.0	21.7	35.9	47.4
Waiting time	30.0	31.8	22.3	23.4	45.0	49.2
Inconvenience	8.6	8.4	5.0	5.5	9.3	9.8
Other	24.7	23.2	14.9	11.6	21.0	22.4
COVID-19	—	—	n.a.	37.0	10.9	1.4
No separately coded reason	9.1	8.1	45.8	15.0	3.9	0.9
<i>Diagnostic examinations</i>						
Cost	47.8	45.7	24.2	18.8	35.8	47.7
Waiting time	32.6	31.1	21.4	23.6	49.6	52.5
Inconvenience	7.7	8.3	4.0	5.3	10.5	9.3
Other	21.6	19.5	12.5	9.7	17.3	17.5
COVID-19	—	—	n.a.	31.2	9.3	1.3
No separately coded reason	16.4	19.6	50.9	24.8	5.4	2.8

Notes: Sampling-weighted shares among respondents forgoing the respective service in each wave. Multiple reasons may be selected, so shares across reasons within a service and year may exceed 100 per cent. COVID-19 dashes denote pre-pandemic waves; n.a. means not separately identifiable in the released 2020 microdata, not zero. The residual means no available reason item is recorded, not absence of a reason. Coding changes limit comparisons across years (Sections 2.3.2 and 2.3.3).

Source: Author’s calculations using ISTAT AVQ microdata, 2018–2023.

3.5.3 COVID-19 Coding and the Residual Reason Category

The 2020 wave cannot be treated in the same way as the later pandemic waves. As documented in Section 2.3.3, the released 2020 microdata contain no dedicated COVID-19 reason item for either specialist visits or diagnostic examinations. The COVID-19 entry is therefore reported as not separately identifiable (n.a.) rather than as zero. The large 2020 residual — 45.8 per cent for specialist visits and 50.9 per cent for diagnostic examinations — is correspondingly labelled “no separately coded reason”: it cannot be interpreted as pure item non-response or as the absence of a substantive reason. This limitation is material because official ISTAT aggregate reporting identifies COVID-related reasons as an important component of forgone care already in 2020 (Istituto Nazionale di Statistica (Istat), 2021), although that aggregate information cannot be decomposed into the two service-specific categories used in this table.

From 2021 onwards, COVID-19 is explicitly coded in the released microdata. It is reported by 37.0 per cent of those forgoing a specialist visit and 31.2 per cent of those forgoing a diagnostic examination in 2021, falling to roughly one in ten in 2022 and to residual levels by 2023. The 2021–2023 figures therefore treat COVID-19 as

an autonomous reason rather than merging it into the residual category. Because this dedicated item is unavailable in 2020, the composition of reasons in that wave is not directly comparable with the later pandemic waves, and the 2020 shares for cost, waiting time, inconvenience and other reasons should be read with that qualification.

3.5.4 The Changing Relative Importance of Cost and Waiting Time

The weighted descriptive results show a clear change in the relative ordering of cost and waiting time across the two internally more comparable endpoint periods. In 2018 and 2019, cost is reported more often than waiting time for both services: 52.7 versus 30.0 per cent for specialist visits in 2018, and 47.8 versus 32.6 per cent for diagnostic examinations. In 2022–2023, the ordering is the opposite for both services. In 2023, waiting time is reported by 49.2 per cent of those forgoing a specialist visit against 47.4 per cent citing cost, and by 52.5 against 47.7 per cent for diagnostic examinations.

The two pandemic-transition waves are not used to establish this comparison. In 2020 a COVID-specific reason cannot be separately identified in the released microdata, while in 2021 COVID-19 is explicitly coded and accounts for a large share of reported forgone care. The interpretation therefore rests on 2018–2019 versus 2022–2023, where the reason indicators are more internally comparable. On this basis, the weighted AVQ microdata are consistent with a change in the reported composition of barriers, from a stronger role for economic cost before the pandemic to a stronger role for waiting time by 2022–2023. This is a descriptive change in the ordering of reported reasons rather than a formal test of a structural break.

The same endpoint ordering is also visible in official ISTAT population-level statistics published in the Report on Equitable and Sustainable Well-being (Istituto Nazionale di Statistica (Istat), 2021, 2024). In those statistics, the share of the entire population — rather than only of those forgoing care — citing economic difficulty as a reason for forgone specialist visits or diagnostic examinations is 4.3 per cent in 2019 and 4.2 per cent in 2023, while the share citing long waiting lists rises from 2.8 to 4.5 per cent. The two sets of indicators have different denominators and are therefore not combined quantitatively. Their common change in ordering nevertheless provides an external consistency check on the pattern observed in the weighted AVQ reason shares.

3.5.5 Cost and Waiting Time by Macro-Area

Figure 3.6 reports the same two reasons by ISTAT macro-area (North-West, North-East, Centre, South, Islands) for 2018, 2021 and 2023. At the national level, cost is the more frequently reported barrier in 2018 and waiting time in 2023 (Section 3.5.4); disaggregating by macro-area shows that the change in ordering is not uniform across the country.

In 2018, cost is the more frequently reported reason in every macro-area, though its margin over waiting time varies considerably: from 9 percentage points in the Centre to 32 in the South (Table 3.20). By 2023, waiting time is more frequently reported than cost in the North-West, North-East and Islands, matching the national ordering; in the Centre and the South, cost remains the more frequently reported barrier, at 52.2 against 48.8 per cent for specialist visits in the Centre and 59.3 against 39.5 per cent in the South, with a comparable pattern for diagnostic examinations in the South (61.7 against 41.8 per cent). The Centre and the South are therefore the two macro-areas in which cost remains more frequent than waiting time in 2023.

Table 3.20: Reasons for forgoing a specialist visit by macro-area (per cent of those forgoing care)

Macro-area	2018		2021		2023	
	Cost	Waiting	Cost	Waiting	Cost	Waiting
North-West	49.7	26.2	16.6	21.9	38.5	53.1
North-East	44.2	25.3	16.9	22.3	40.5	51.0
Centre	47.4	38.1	24.2	24.5	52.2	48.8
South	61.9	29.9	26.9	25.3	59.3	39.5
Islands	61.3	29.6	27.6	23.8	47.2	56.5

Notes: Sampling-weighted shares among respondents forgoing a specialist visit within each ISTAT macro-area and wave, not shares of the whole population. Cost and waiting time are non-exclusive reported reasons. Other reasons, including COVID-19 in 2021, remain in the denominator. Cross-period comparisons concern the ordering of reasons rather than directly comparable levels (Section 2.3.2).

Source: Author's calculations using ISTAT AVQ microdata, 2018, 2021 and 2023.

In every wave and macro-area, waiting time shows less variation than cost: its share ranges from about 22 to 57 per cent across the fifteen macro-area-wave combinations shown in Table 3.20, against a range of roughly 17 to 62 per cent for cost. The South is consistently among the macro-areas with the highest share citing cost in all three waves shown.

(a) Reasons for forgone care by macro-area — Specialist visits

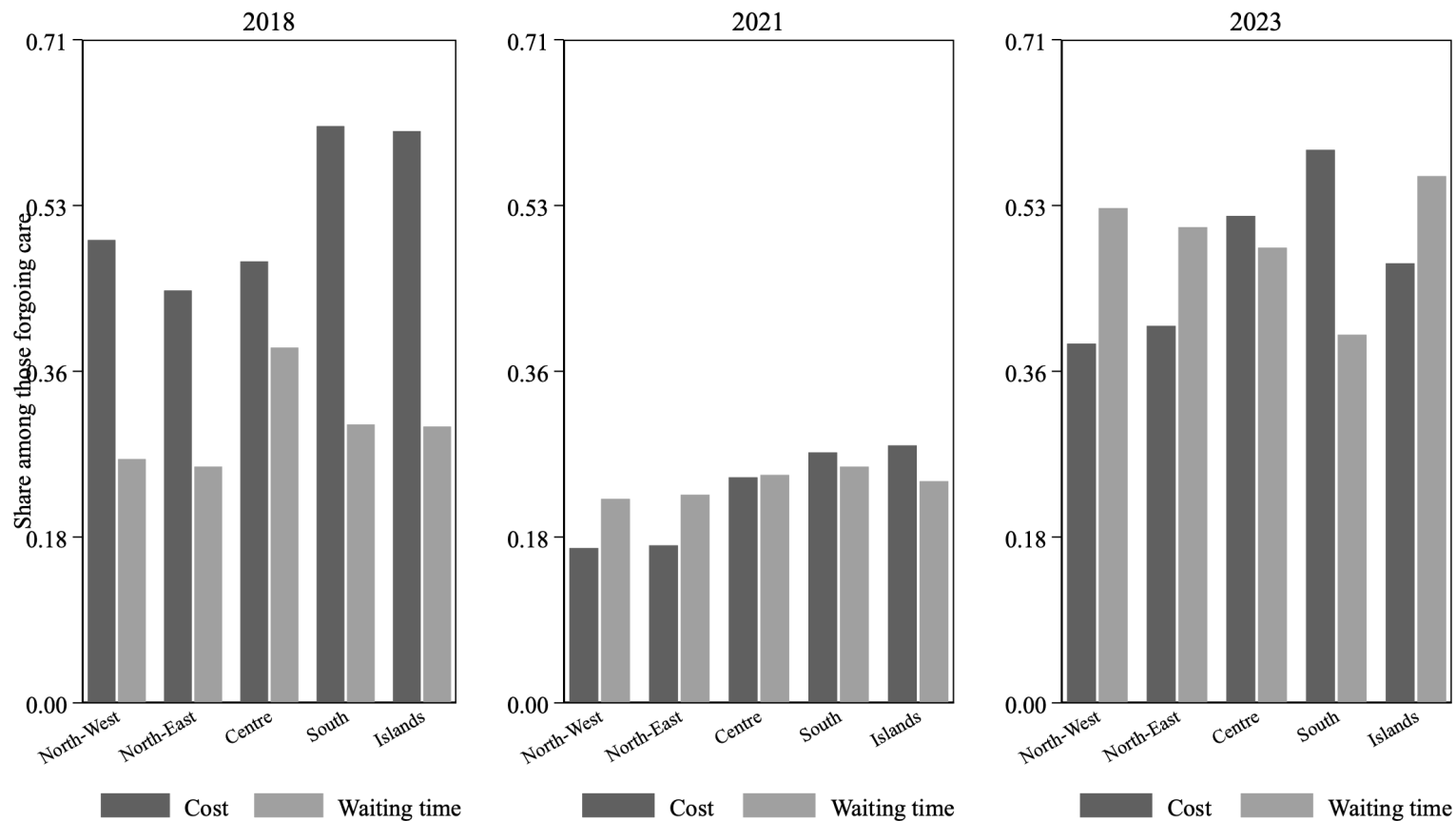


Figure 3.6: Reasons for forgone care by macro-area and service, 2018, 2021 and 2023. Panel (a); panel (b) follows on the next page.

Source: Author's calculations using ISTAT AVQ microdata, 2018, 2021 and 2023.

(b) Reasons for forgone care by macro-area — Diagnostic examinations

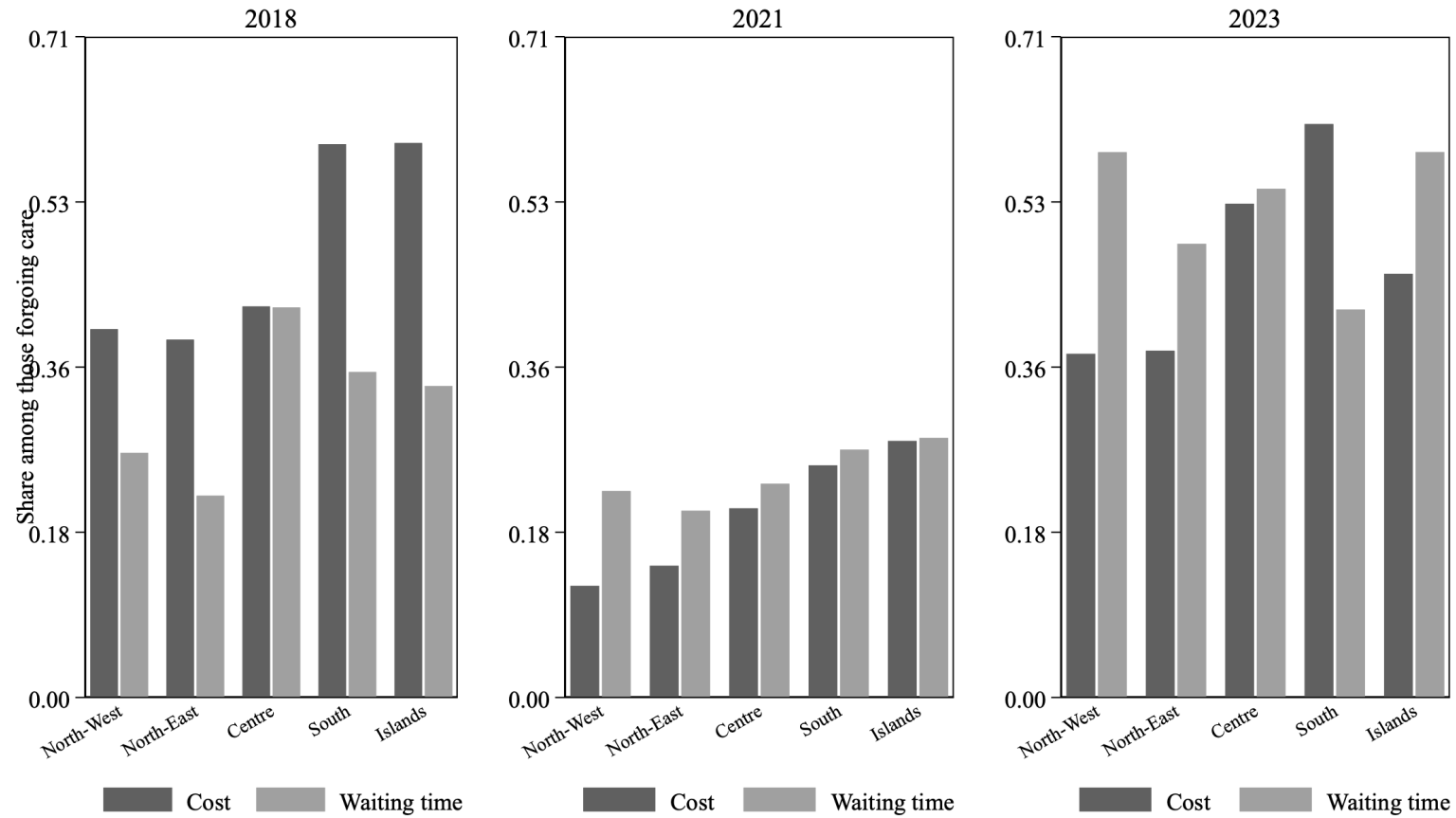


Figure 3.6: Reasons for forgone care by macro-area and service, 2018, 2021 and 2023 (continued): panel (b).

Source: Author's calculations using ISTAT AVQ microdata, 2018, 2021 and 2023.

3.5.6 The Socioeconomic Gradient in Reported Barriers

Table 3.21 reports the share of respondents forgoing a specialist visit who cite cost or waiting time, by SES quintile, for the three waves examined at the regional level in Section 3.4. In every wave, the share citing cost falls from the poorest to the richest quintile: from 60.5 to 44.5 per cent in 2018, from 28.1 to 15.5 per cent in 2021, and from 57.0 to 39.0 per cent in 2023. The share citing waiting time shows no comparable gradient in any of the three waves, ranging narrowly and non-monotonically across quintiles.

Table 3.21: Reasons for forgoing a specialist visit by SES quintile (per cent of those forgoing care)

Quintile	2018		2021		2023	
	Cost	Waiting	Cost	Waiting	Cost	Waiting
Q1 (poorest)	60.5	26.6	28.1	24.2	57.0	46.3
Q2	51.8	33.5	21.2	25.3	42.6	50.4
Q3	51.3	29.2	21.6	24.7	48.3	48.1
Q4	49.8	30.6	20.2	23.3	44.4	50.4
Q5 (richest)	44.5	32.4	15.5	18.6	39.0	52.7

Notes: Sampling-weighted shares among respondents forgoing a specialist visit within each socioeconomic status (SES) quintile and wave. Q1 is the lowest and Q5 the highest SES group. Cost and waiting time are non-exclusive; respondents citing other reasons remain in the denominator. Differences are descriptive, and the 2022 coding revision limits direct cross-period comparisons of levels.

Source: Author's calculations using ISTAT AVQ microdata, 2018, 2021 and 2023.

This is consistent with the change in ordering documented in Section 3.5.4. Cost is reported disproportionately by respondents in the lower socioeconomic groups, whereas the corresponding descriptive gradient is much less evident for waiting time.

3.5.7 Summary

Two findings from this section are carried forward to the discussion in Chapter 4: the change in the relative ordering of the two main reported barriers, with cost more frequent in 2018–2019 and waiting time more frequent in 2022–2023, and the socioeconomic gradient that is more evident for cost than for waiting time.

Chapter 4

Discussion

4.1 Summary of Main Findings

Chapter 3 established five results. Need standardisation raises the measured concentration index for the three programmable services in every wave, and most sharply for blood tests, where the unadjusted index is close to zero throughout (Section 3.2.2). For the three programmable services, the need-adjusted index has higher point estimates in 2020–2021 than in 2019, followed by lower point estimates in 2022–2023 (Section 3.2.3). In the decomposition, the 2019–2020 point-estimate difference is accompanied by a higher SES contribution for specialist visits and diagnostic examinations, while the corresponding contribution for blood tests is essentially unchanged; for specialist visits, the age contribution also becomes less negative. For diagnostic examinations and blood tests only the aggregate need block is reported (Section 3.3.2). For the emergency department, the point estimate is negative in every wave; it is statistically distinguishable from zero in 2018, 2019 and 2021, but not in 2020, 2022 or 2023 (Section 3.2.4). The individual regional point estimates do not display a clear North–South ordering in the reported waves (Section 3.4.2), but no formal macro-area comparison is conducted; this descriptive pattern therefore does not establish that a North–South gradient is absent (Section 3.4.1). Finally, the weighted descriptive results show a change in the relative ordering of the two principal reported barriers: cost is more frequently reported in 2018–2019, while waiting time is more frequently reported in 2022–2023. The same endpoint ordering is visible in official ISTAT population-level indicators, although the denominators differ (Sections 3.5.4–3.5.6).

The remainder of this chapter interprets these results, situates them against the existing literature on equity in the Italian SSN, and sets out the principal limitations of the analysis.

4.2 The Raw Index Conceals Inequity: the Case of Blood Tests

The gap between $E(y)$ and ENA documented in Section 3.2.2 is not merely a technical artefact of the standardisation procedure: it carries a substantive implication for how equity in access is assessed in practice. Blood tests illustrate this most clearly. Raw utilisation is distributed almost evenly across the socioeconomic ranking (Section 3.1.3), and the unadjusted index is correspondingly close to zero throughout the period (Table 3.7). Judged on utilisation data alone, blood tests would appear to be the service, among the four examined, closest to the horizontal equity benchmark set out in Section 2.4.1. The need-adjusted index tells a different story: once the greater need for blood tests among older, less healthy and more chronically ill individuals is taken into account, the service displays a pro-rich pattern of a magnitude comparable to specialist visits and diagnostic examinations in several waves.

The mechanism behind this is not specific to blood tests as a service, but follows directly from the algebra of the decomposition set out in equation (2.3). A flat raw distribution can arise for reasons that have opposite implications for equity: because access is genuinely unrelated to socioeconomic position at any level of need, or because greater need among the disadvantaged is being offset by comparatively worse access conditional on that need. The unadjusted index cannot distinguish between these two cases; the need-adjusted index resolves the ambiguity by asking, in effect, what the distribution of utilisation would look like if need were held equal across the socioeconomic ranking. For blood tests, the answer is that the apparent equality conceals a pro-rich pattern in access conditional on need.

This has a direct implication for how routine monitoring of access to care is interpreted. A service whose raw utilisation appears evenly distributed across income or wealth groups is sometimes read, in policy discussion, as evidence that the service poses no equity concern. The results for blood tests indicate that this reading can be mistaken whenever the need for a service is itself unevenly distributed across the population — which, for most health services, it plausibly is. Monitoring based on raw utilisation alone is consequently insufficient to establish equity of access, and the case for routinely reporting need-adjusted rather than raw indicators, at least for services where need is known to vary systematically with age, health status or chronic morbidity, follows from this result rather than from a general methodological preference.

A qualification is in order. The finding that blood tests display substantial need-adjusted inequity does not by itself indicate that the service is administered

inequitably in a narrow sense — for instance, through differential co-payment enforcement or scheduling practices that disadvantage lower-income patients. The decomposition in Section 3.3.1 shows that the SES contribution for blood tests, while smaller in absolute terms than for specialist visits, is positive and substantial throughout the period; distinguishing whether this reflects differential access to the same nominal service, differential propensity to seek testing conditional on a physician’s recommendation, or some combination of the two is not possible with the utilisation data used here, and would require information on referral patterns that the AVQ does not collect.

4.3 The Higher 2020 Inequity Point Estimates and Their Subsequent Pattern

4.3.1 Timing and Magnitude of the Point-Estimate Difference

The most visible movement in the national point estimates occurs between the 2019 and 2020 waves. Raw-utilisation point estimates are lower in 2020 for all four services (Section 3.1.1), most notably for blood tests and specialist visits, while ENA point estimates are higher for the three programmable services (Section 3.2.3). For the emergency department the point estimate remains negative, but the 2020 estimate is no longer statistically distinguishable from zero (Section 3.2.4). The 2021 estimates remain closer to the 2020 than to the 2019 values for several of these outcomes, while the 2022–2023 estimates move back toward the earlier range to varying degrees.

It should be stressed that the significance tests reported in Chapter 3 establish only whether each wave’s estimate is distinguishable from zero. They do not test whether, for example, the 2020 ENA estimate is statistically different from the 2019 estimate. No formal cross-wave difference test is conducted in this thesis. All temporal comparisons in this section are therefore interpreted as differences in point estimates and their descriptive pattern, not as statistically ascertained year-to-year changes or discontinuities.

The higher 2020 point estimate coincides with the onset of the COVID-19 pandemic and with documented disruption in access to care, although the repeated cross-sectional design does not permit a causal attribution. The 2020 wave should therefore be interpreted as a pandemic-affected observation: ISTAT’s official BES reporting documents a marked increase in forgone care in 2020 and identifies COVID-related reasons as an important component of that increase (Istituto Nazionale di

Statistica (Istat), 2021). The interpretation adopted here is therefore contextual rather than causal, and the thesis does not estimate the counterfactual path that would have prevailed in the absence of the pandemic.

4.3.2 Changes in the Age and Socioeconomic Contributions

The decomposition offers a more specific reading of the difference in point estimates. The SES contribution point estimate is higher in 2020 than in 2019 for specialist visits and diagnostic examinations, while the two point estimates are essentially the same for blood tests. For specialist visits, the age contribution point estimate is less negative in 2020, at -0.022 compared with -0.034 in 2019 (Section 3.3.2). The component-level calculation shows that this difference is associated primarily with a lower marginal-effect point estimate for being aged 65 or over, from 0.134 in 2019 to 0.091 in 2020. The weighted prevalence of the 65+ group changes little, from 0.277 to 0.284 , while its concentration index becomes moderately less negative, from -0.230 to -0.212 . The smaller age contribution is therefore not a mechanical consequence of there being fewer older respondents; arithmetically, it is associated mainly with the lower estimated conditional association between age and specialist-visit utilisation, with a smaller contribution from the difference in the socioeconomic distribution of older respondents. No formal test of the cross-wave difference in these component estimates is conducted, so the comparison remains descriptive.

For diagnostic examinations and blood tests, the available decomposition table reports only the aggregate need block. Its 2020 point estimate is less negative than its 2019 point estimate, but this descriptive difference cannot be assigned specifically to age without the individual need components.

A plausible interpretation of the specialist-visit result is that pandemic-period changes in care-seeking or access altered the conditional association between older age and utilisation. Fear of infection, shielding behaviour and the suspension or postponement of services are possible channels, but the AVQ utilisation data cannot distinguish changes in demand from changes in supply or access. The component decomposition therefore supports a narrower statement about how the estimated age contribution changed; it does not identify the behavioural or institutional mechanism that generated that change.

4.3.3 Post-2019 Patterns and Reported Barriers

For specialist visits, the post-2019 age-contribution point estimates are less negative than the two pre-2020 point estimates, while the SES contribution is positive

throughout the period (Table 3.8). For diagnostic examinations and blood tests, only the aggregate need contribution is available in the summary decomposition, so no claim is made about the persistence of an age-specific component. These decomposition patterns coincide in time with the change in the relative ordering of reported barriers to access (Sections 3.5.4—3.5.6). In the weighted descriptive results, cost is the more frequently reported barrier in 2018–2019, whereas waiting time is the more frequently reported barrier in 2022–2023. The 2020–2021 waves are not used to establish this comparison because the COVID-related reason is not coded comparably across those two years. Cost is also disproportionately cited by the lowest socioeconomic quintile, while the descriptive gradient for waiting time is much less evident.

The two sets of results are consistent with one another in their timing, but the AVQ’s repeated cross-sectional design does not permit them to be linked at the individual level, and no causal relationship between the two is proposed here. What can be said is that the persistence of the socioeconomic gradient in utilisation after 2019 does not coincide with a comparable persistence in the specific barrier — cost — most clearly associated with that gradient in the descriptive data; whether, and how, the two are substantively connected is a question this analysis is not designed to answer.

4.4 The Emergency Department: A Small and Unstable Pattern

The emergency-department ENA point estimate is negative in every wave, but its statistical significance is not stable across the period. It is significantly pro-poor in 2018, 2019 and 2021, and not statistically distinguishable from zero in 2020, 2022 and 2023 (Section 3.2.4). This pattern invites two different readings, which the data available here cannot fully separate.

One reading treats 2018–2019 as the informative baseline and the later waves as a departure from it. On this reading, the pro-poor distribution observed before 2020 could reflect lower-income individuals substituting emergency access for programmable care they found comparatively harder to obtain — a substitution consistent with the higher SES and cost barriers to programmable services documented in Sections 3.3.1 and 3.5.4. If the capacity or willingness to use emergency departments as a substitute channel changed after 2019 — for instance, because the same disruption affecting programmable services also affected emergency access, or because individuals became less willing to attend emergency departments during the

pandemic period — the pattern from 2020 onward, alternating between a significant pro-poor estimate in 2021 and non-significant estimates in the surrounding waves, would be consistent with an intermittent or weakened form of this substitution mechanism rather than its clean interruption.

A second, more cautious reading gives less weight to the changing significance status across waves. Three of the six annual estimates are significant at conventional levels, but the sequence — significant, significant, non-significant, significant, non-significant, non-significant — does not describe a single clean break between two regimes; a pattern that switches on and off across adjacent waves is at least as consistent with substantial year-to-year sampling variability around a true value close to zero as with a genuine substitution mechanism that appears and disappears. The point estimates remain small relative to the other three services in every wave (Table 3.7), and the emergency department decomposition shows contributions that are frequently indistinguishable from zero at the individual level (Section 3.3.3). On this reading, the non-monotonic pattern of significance may reflect noise around a small and unstable parameter rather than a reversal, disruption, or recovery of an underlying substitution mechanism.

Distinguishing between these two readings is not possible with the estimates presented here. Both are consistent with the observed sequence of significant, significant, non-significant, significant, non-significant, non-significant estimates; neither can be ruled out on the basis of the point estimates and standard errors alone. What can be said without ambiguity is that the emergency-department point estimate remains negative and small in every wave, while its significance status varies across the period. This variation occurs over the same broad 2019–2020 transition examined in Section 4.3 for the programmable services, but no formal cross-wave difference is inferred.

This finding sits uneasily alongside part of the literature reviewed in Chapter 1. Ferré et al. (2014) describe specialist and emergency care as the two services with the most pronounced departures from equity in the earlier Italian literature, with emergency care in some Regions used as a surrogate for specialist care that is harder to obtain (Section 1.6). The results in this thesis point in the opposite direction for the emergency department specifically: across all six waves examined here, it is the service with by far the smallest departure from horizontal equity, pro-poor or neutral rather than pro-rich, and never approaching the magnitude documented for specialist visits or diagnostic examinations (Table 3.7). Three explanations are possible, and the data available here cannot distinguish between them. One possibility is that the underlying pattern differs from that documented for the earlier period

studied in the literature, during which healthcare organisation and reported access barriers also differed. The present data do not establish that such a change occurred. Alternatively, the substitution mechanism Ferré et al. (2014) describe — emergency care used as a surrogate for specialist care — may still operate, but may manifest as a pro-poor rather than a pro-rich signal in a concentration-index framework, since substitution by the disadvantaged toward a less desirable but more accessible channel would tend to produce exactly the pattern found in 2018 and 2019 before losing significance from 2020. A third possibility is more mundane: differences in data source, need adjustment and time period between this thesis and the literature synthesised by Ferré et al. (2014) may be sufficient on their own to account for the divergence, without requiring any substantive change in the underlying pattern of access. Adjudicating between these readings would require data this thesis does not have, in particular a direct measure of whether emergency department use is associated with prior difficulty in obtaining programmable care at the individual level.

4.5 Regional Estimates and the North–South Question

The regional point estimates for specialist visits, diagnostic examinations and blood tests do not display a clear North–South ordering in the waves reported in Chapter 3. For specialist visits, for example, Southern Regions appear both among relatively high and more moderate point estimates, while Northern and Central Regions also occupy different positions in the ranking across waves. A similarly mixed descriptive ordering is visible for diagnostic examinations and blood tests (Sections 3.4.2 and 3.4.5).

This visual and rank-based evidence is not sufficient to establish the absence of a territorial gradient. Regional estimates are imprecise, as shown by the reliability analysis in Section 3.4.1, and instability in individual-region rankings does not constitute a formal test of differences between macro-areas. A comparison of average ENA between, for example, Northern and Southern macro-areas would be required to test that hypothesis directly, and such a test is not undertaken in this thesis.

The appropriate conclusion is therefore narrower: *no clear North–South gradient emerges from the individual regional point estimates reported here*. This statement describes the pattern visible in the data; it does not imply that a North–South gradient is absent in the underlying population. The contrast with the well-documented territorial gradients in health-system resources, performance and patient mobility

discussed in Chapter 1 should accordingly be treated as a question for further research rather than as a discrepancy explained by the present analysis.

4.6 Comparison with the Existing Literature

4.6.1 Levels Relative to Di Novi (2025)

The estimates reported in this thesis are broadly comparable in sign and in their broad grouping across services to those reported by Di Novi (2025) for the Italian SSN, but not identical in level. Where a direct comparison is possible, the estimates obtained here for 2023 are somewhat higher: for specialist visits, diagnostic examinations and blood tests, the need-adjusted index reported in Table 3.7 exceeds the corresponding published estimate. In both studies, specialist visits and diagnostic examinations are the two services with the largest pro-rich inequity, blood tests are lower, and emergency department access is approximately neutral. The relative levels of specialist visits and diagnostic examinations are not treated as a substantive contribution of the present thesis: the differences are very small in the present series, and Di Novi (2025) already reports the corresponding 2023 estimates (0.104 for specialist visits and 0.114 for diagnostic examinations).

As set out in Section 2.3.1, part of this discrepancy may be related to the definition of chronic morbidity. The item-based indicator constructed for this thesis, necessitated by the unavailability of the ISTAT summary variable in two of the six waves, captures a broader notion of morbidity than the single global question underlying that summary variable: it classifies 47.5 per cent of the 2018 sample as chronically ill, against 31.5 per cent under the ISTAT variable. A broader morbidity indicator can absorb a larger share of observed inequality into the need component of the decomposition and thereby affect the resulting need-adjusted index. The direction of the discrepancy documented here is consistent with this explanation, but the two studies also differ in sample construction and other implementation details; the chronicity measure should therefore be read as a plausible contributor rather than as a demonstrated explanation of the entire difference.

The 2023 macro-area pattern is also broadly consistent with Di Novi (2025): waiting time is more frequently reported than cost in the North-West, North-East and Islands, whereas cost remains more frequent in the Centre and South (Section 3.5.5). This geographic ordering should not be projected backward across the entire six-wave period. In 2018, cost is more frequently reported in every macro-area, and the 2020–2021 reason shares are not used to establish a continuous trend because pandemic-related reasons are not coded comparably across those waves.

4.6.2 Positioning Relative to Di Novi, Piacenza, Robone and Turati (2019)

Di Novi et al. (2019) address a different question about the same institutional setting, using a different outcome and a different design. Exploiting cross-regional variation in per capita GDP as a continuous measure of exposure to the 1998 fiscal decentralisation reform in a difference-in-differences framework, they ask whether greater regional fiscal autonomy affected disparities in self-assessed health, and find no effect on disparities between Regions but a significant reduction in disparities within Regions, concentrated in wealthier ones (Section 1.7). Their outcome is a median-based inequality index of self-assessed health computed at the Region-year level, and their design is built to isolate the causal effect of a specific reform on that outcome; neither the measurement approach nor the object of inference coincides with the one used here.

The two studies are complementary rather than overlapping. Di Novi et al. (2019) establish that a specific institutional reform narrowed within-Region disparities in health status, concentrated among wealthier Regions, over the period 1994–2007. The present thesis documents, without attributing the pattern to any single reform, annual estimates of the socioeconomic gradient in utilisation of four specific healthcare services within the Italian SSN context over 2018–2023, extending the ENA methodology of Di Novi (2025) discussed in Section 4.6.1 along the temporal dimension. Read together, the two studies point to more than one channel through which regional institutional arrangements bear on health equity in Italy — the distribution of health status within Regions, examined by Di Novi et al. (2019), and the socioeconomic gradient in access to specific services, examined here — without either study establishing a direct link between the two.

4.6.3 What This Thesis Adds

The main contribution of this thesis is the extension of a previously cross-sectional ENA analysis into a six-wave series that can be examined jointly with decomposition and forgone-care evidence. The relative ordering of specialist visits and diagnostic examinations is not itself treated as a new contribution, since their point estimates are very similar and the 2023 levels are already reported by Di Novi (2025).

Within that multi-wave design, three results are emphasised. First, a service with an approximately flat raw utilisation gradient — blood tests — can nonetheless display substantial need-adjusted inequity, illustrating the importance of standardising for need. Second, the decomposition-based account of the 2019–2020 point-estimate

difference separates the higher SES contribution from the less negative age contribution observed for specialist visits, and the component-level analysis further distinguishes differences in the estimated age marginal effect, prevalence and concentration rather than treating the aggregate contribution as a direct measure of predictive strength. Third, the documented change in the relative ordering of cost and waiting time as reported barriers to access is examined alongside the socioeconomic gradient in utilisation and the distribution of reported barriers.

4.7 Limitations

4.7.1 Precision of the Estimates

The regional estimates presented in Section 3.4 carry substantially more sampling uncertainty than the national series. The reliability ratios reported in Table 3.12 indicate that between 4 and 39 per cent of the observed dispersion across the twenty Regions reflects genuine differences in inequity, with the remainder attributable to sampling variability. For diagnostic examinations, the service with the lowest reliability, this means that the regional pattern documented in this thesis should be interpreted with particular caution: individual regional rankings, and comparisons between specific pairs of Regions, are not in general statistically robust, even where the point estimates differ considerably.

This limitation follows from the structure of the AVQ rather than from a specific choice made in this analysis. The survey is designed to be representative both nationally and regionally, but necessarily allocates substantially fewer respondents to each individual Region than to the country as a whole, and the resulting regional samples — particularly for smaller Regions such as Valle d’Aosta, Molise and Basilicata — are not large enough to estimate a bootstrapped concentration index with precision comparable to the national estimate. A related limitation, noted in Section 2.6.3, is that the bootstrap procedure resamples individuals directly and does not reproduce the two-stage structure of the AVQ sampling design, in which municipalities constitute the primary sampling unit; standard errors throughout this thesis do not account for clustering at that level and may therefore understate sampling uncertainty. Both limitations apply with greater force to the regional than to the national estimates, since clustering effects are diluted less by sample size at the regional level.

4.7.2 Construction of the Data

Three aspects of the underlying data constrain what the analysis can establish. The first concerns the chronic morbidity indicator. As discussed in Section 2.3.1 and again in Section 4.6.1, the item-based indicator adopted here was necessitated by the unavailability of the ISTAT summary variable in two of the six waves, and captures a broader notion of morbidity than that variable does. This choice was required for the six-wave series to be internally comparable, but it also means that the resulting need-adjusted estimates are not directly comparable, in level, to studies using the ISTAT summary variable, and that some part of the discrepancy documented in Section 4.6.1 cannot be separated, with the data available here, from other sources of difference between the two studies.

The second concerns the 2022 revision of the questionnaire items recording forgone care, discussed in Section 2.3.2. Because the revision changes what a non-response on these items represents, the forgone care rates and reason shares reported in Section 3.5 are only comparable within, and not across, the two resulting periods. This constrains the forgone-care analysis to qualitative comparisons of internally comparable periods — such as the different ordering of cost and waiting time in 2018–2019 versus 2022–2023 — rather than to a continuous six-wave series of the kind presented for the concentration indices themselves. This limitation arises from changes in the source questionnaire and in the coding of pandemic-related reasons rather than from a methodological choice made in this thesis.

A related limitation concerns the coding of reasons for forgone care during the pandemic. The released 2020 microdata do not contain the dedicated COVID-19 reason items that are available from 2021 onwards. Official ISTAT aggregate statistics confirm that COVID-related reasons were already important in 2020 (Istituto Nazionale di Statistica (Istat), 2021), but that aggregate information cannot be used to identify which respondents in the service-specific microdata reported COVID-related renunciation. The 2020 composition of reasons is therefore incomplete with respect to COVID-19 and is not treated as directly comparable with 2021–2023.

A third limitation concerns the emergency-department sample specifically in 2023. After the general sample restrictions, 2,777 observations have no recorded value for PROSOC, a much larger number than in the other waves (Section 2.1.3). Diagnostic checks show that the same observations are almost always missing on the adjacent questions on out-of-hours care, home care and hospitalisation, while all retain a valid self-assessed-health response and roughly 37 per cent also answer the later questions on specialist visits, blood tests and diagnostic examinations. This makes general survey dropout an implausible explanation and points instead

to concentrated non-recording within part of the healthcare-use block. The available documentation does not, however, identify a routing rule that would justify classifying these cases as structural missing values.

The missingness rate does not differ detectably by sex or age category, but it does vary significantly across Regions, from about 4.8 to 13.8 per cent. Consequently, the possibility of territorial selection in the 2023 emergency-department sample cannot be ruled out. Re-estimating the 2023 models on service-specific complete-case samples leaves the need-adjusted conclusions essentially unchanged for all four services, including the near-zero and non-significant emergency-department ENA (Section 3.1.6). This sensitivity analysis shows that the additional common-sample exclusions are not driving the main result, but it cannot recover or reveal emergency-department utilisation for the 2,777 respondents whose PROSOC value is itself unobserved. The 2023 emergency-department estimate should therefore still be interpreted with particular caution.

4.8 Policy Implications

The results in this thesis point toward two considerations for policy that follow reasonably directly from the empirical findings, and a third that follows only tentatively.

The first concerns the different distributional profiles of the reported barriers. In 2018–2019 the barrier most frequently reported by those forgoing specialist visits or diagnostic examinations was cost, and cost was disproportionately reported by the lowest socioeconomic quintile in every wave examined (Section 3.5.6). By 2022–2023, waiting time had become the more frequently reported barrier, and it shows no comparable socioeconomic gradient among those who report it (Section 3.5.4). Two qualifications limit how far these patterns can be read as policy guidance. The AVQ does not distinguish whether the cost reported is a public co-payment or a private-sector fee (Section 2.2.3), so an intervention aimed at this barrier cannot be specified more precisely, from these data alone, than as one targeting the affordability of care in general. And the absence of a socioeconomic gradient among those who report waiting time as a barrier does not establish that reducing waiting times would leave the socioeconomic distribution of access unaffected: if wealthier respondents are disproportionately able to avoid public waiting lists by paying for care privately, they would not appear in this statistic at all, and the observed absence of a gradient could coexist with a waiting-time barrier that is itself unequally distributed before any such avoidance takes place. Read as hypotheses rather than as demonstrated effects,

the results are consistent with cost-reducing interventions addressing a barrier that remains disproportionately borne by the disadvantaged throughout the period, and with the distributional consequences of waiting-time interventions requiring evidence beyond what this thesis provides before they can be anticipated.

The second concerns what the blood-test result implies for equity monitoring. The result for blood tests in Section 4.2 indicates that a service can appear equitable on the basis of routinely available raw utilisation data while displaying substantial need-adjusted inequity. Where monitoring of equity in healthcare utilisation within the Italian health system relies on raw utilisation indicators — as much routine reporting does, given that need-adjusted indices require additional data and modelling that raw utilisation rates do not — this result motivates closer examination of services with a known and unevenly distributed need profile, such as blood tests, even when raw indicators show no cause for concern.

The third consideration is more tentative. The fact that several post-2019 ENA point estimates remain above their 2019 values (Section 3.2.3) raises the question of whether the distributional pattern changed only temporarily or may have persisted beyond the immediate pandemic shock. Because no formal cross-wave difference test is conducted, the six waves observed here cannot establish such persistence statistically. If the same pattern were confirmed by future waves, it would motivate examining whether recovery in utilisation levels is accompanied by recovery in the socioeconomic distribution of utilisation rather than assuming that the two necessarily move together. This consideration is offered with appropriate caution, given the limitations on causal attribution to the pandemic discussed in Section 4.3.1 and the general constraint that repeated cross-sectional data of the kind used throughout this thesis cannot establish causal mechanisms.

Conclusion

This thesis has extended recent evidence on horizontal equity in healthcare utilisation in Italy from a single cross-section to six consecutive annual waves of the ISTAT *Aspetti della Vita Quotidiana* survey covering 2018–2023. Using the need-adjusted Erreygers concentration index at both the national and regional level, the analysis places the 2023 estimates reported by Di Novi (2025) within a broader temporal sequence. Because the AVQ consists of independent repeated cross-sections, the purpose is not to identify individual-level changes or causal effects, but to document how annual point estimates and their decomposition vary across successive representative samples.

Taken together, the results show a persistent pro-rich pattern for the three programmable services examined — blood tests, specialist visits and specialist diagnostic examinations — while emergency department access follows a different distributional pattern. The magnitude and composition of the estimated inequity are not identical across waves, but the analysis does not formally test every year-to-year difference. Temporal comparisons are therefore interpreted as differences in annual point estimates rather than as statistically established structural breaks.

Summary of Findings

Several findings emerge consistently from the analysis. First, need standardisation materially changes the assessment of equity, particularly for blood tests. Their raw utilisation gradient across socioeconomic groups is weak, yet the need-adjusted index is positive in every wave. This shows that an apparently flat raw distribution of utilisation can coexist with pro-rich horizontal inequity once the greater health need of socioeconomically disadvantaged respondents is taken into account. Specialist visits and diagnostic examinations display the largest positive ENA estimates across the period, while blood tests show smaller but persistent positive values. The small differences between specialist visits and diagnostic examinations are not interpreted as a substantively meaningful ranking between the two services.

Second, the multi-wave decomposition clarifies the 2019–2020 difference in point estimates for specialist visits. The age contribution becomes less negative, from approximately -0.034 in 2019 to -0.022 in 2020, while the SES contribution point estimate is higher in 2020. Examining the age contribution in its component parts shows that the arithmetic difference is associated mainly with a lower estimated marginal effect of being aged 65 or over; the weighted prevalence of older respondents changes little, while the concentration index of age becomes only moderately less negative. For diagnostic examinations and blood tests, the available summary decomposition reports the aggregate need block rather than the individual age component, so the corresponding changes are not attributed specifically to age. Since no formal test of the 2019–2020 difference in these component estimates is conducted, the interpretation remains descriptive.

The emergency department displays a different pattern from the three programmable services. Its ENA point estimate is negative in every wave and small in magnitude. It is statistically distinguishable from zero in 2018, 2019 and 2021, but not in 2020, 2022 or 2023. The relevant feature is therefore the instability of statistical significance around a small negative estimate, not a change in sign.

At the regional level, the individual point estimates for specialist visits, diagnostic examinations and blood tests do not display a clear North–South ordering in the waves reported. This descriptive result should not be read as evidence that an underlying territorial gradient is absent. Regional estimates are relatively imprecise, and no formal comparison of average ENA across macro-areas is undertaken in this thesis.

Finally, the weighted descriptive evidence on forgone care points to a change in the relative ordering of reported barriers across the more internally comparable endpoint periods. Cost is more frequently reported than waiting time in 2018–2019, whereas waiting time is more frequently reported in 2022–2023. The 2020–2021 waves are not used to establish this comparison because pandemic-related reasons are not coded comparably across the two years. Cost also displays a clearer socioeconomic gradient than waiting time, being reported more frequently by respondents in lower socioeconomic groups. Official ISTAT population-level indicators show the same change in ordering between 2019 and 2023, although their denominator differs from that used in the AVQ conditional reason shares.

Contribution

The principal contribution of this thesis is to extend a recent cross-sectional application of the need-adjusted Erreygers index into a six-wave repeated cross-sectional analysis. This makes it possible to examine whether the broad pattern of horizontal inequity observed in 2023 is also visible in earlier waves, to compare the composition of the index across years, and to place the regional estimates within a multi-year rather than a single-wave framework.

The thesis also combines the concentration-index analysis with weighted descriptive evidence on forgone care. This allows the distribution of utilisation and the reported barriers to obtaining care to be considered together, while keeping the distinction between them explicit. The resulting contribution is therefore not a claim that any single year represents a statistically established break or that the small ranking differences between programmable services are substantively meaningful. It is instead a more detailed account of how pro-rich horizontal inequity and reported barriers to access are distributed across services, socioeconomic groups, Regions and successive annual waves.

Limitations and Directions for Future Research

The limitations of the analysis are discussed in detail in Section 4.7. Several are particularly important for interpreting the conclusions. First, the AVQ is a repeated cross-section, so the same individuals are not followed over time and the observed differences between waves cannot be interpreted as individual-level transitions or causal effects. Relatedly, the thesis does not formally test every year-to-year difference in ENA or in its decomposition components.

Second, regional estimates carry substantial sampling uncertainty. The individual regional rankings are therefore used descriptively, and the thesis does not formally test average differences in inequity between Northern and Southern macro-areas. Third, harmonisation across waves requires an item-based chronic morbidity indicator, while changes in the questionnaire and the coding of forgone-care reasons limit the comparability of that module across the full 2018–2023 period. In particular, a COVID-specific reason is not separately identifiable in 2020. The 2023 emergency-department outcome also has substantially more missing observations than the other utilisation outcomes; sensitivity analysis reduces concern about the need-adjusted result, but the raw emergency-department estimate is more sample-sensitive.

A further limitation is that the utilisation indicators do not distinguish whether

care was obtained from a directly public provider, an accredited private provider or a fully private provider paid for out of pocket. Accordingly, the results describe inequity in healthcare utilisation within the Italian SSN context rather than inequity in publicly delivered care alone.

These limitations suggest several directions for future research. Extending the analysis to later AVQ waves would allow the post-2023 pattern to be examined and, ideally, formal pooled tests could be used to assess whether differences between annual ENA estimates are statistically distinguishable. A second extension would be to test territorial differences directly by estimating and comparing average inequity across macro-areas rather than inferring a North–South pattern from individual regional rankings. Administrative data on interregional patient mobility could complement this analysis by showing how mobility among residents is related to regional patterns of access; because the AVQ classifies respondents by Region of residence, care received elsewhere remains attributed to the respondent’s Region of origin and cannot be interpreted as inequity being “exported” to the Region of treatment. Finally, panel or linked administrative data would be required to distinguish changes in care-seeking from changes in supply and access, and to evaluate more directly the candidate mechanisms associated with the pandemic period.

The Servizio Sanitario Nazionale was founded on the principle that access to healthcare should depend on need rather than socioeconomic position. The evidence assembled in this thesis indicates that this principle is not fully reflected in the distribution of the programmable healthcare services examined over 2018–2023: after adjustment for need, utilisation remains systematically concentrated among socioeconomically advantaged individuals. At the same time, the magnitude and composition of the estimated inequity, as well as the reported barriers to care, differ across waves. Documenting these patterns while distinguishing robust findings from descriptive cross-wave differences is the contribution this thesis has sought to make.

Bibliography

- Bifulco, L., & Neri, S. (2022). The Italian National Health Service: Universalism, marketization and the fading of territorialization. *Forum for Social Economics*, 51(2), 192–206. <https://doi.org/10.1080/07360932.2022.2036625>
- Brenna, E. (2011). Quasi-market and cost-containment in Beveridge systems: The Lombardy model of Italy. *Health Policy*, 103(2–3), 209–218. <https://doi.org/10.1016/j.healthpol.2011.10.003>
- Brenna, E., & Spandonaro, F. (2015). Regional incentives and patient cross-border mobility: Evidence from the Italian experience. *International Journal of Health Policy and Management*, 4(6), 363–372. <https://doi.org/10.15171/ijhpm.2015.65>
- Di Novi, C. (2025). Disuguaglianze di accesso ai servizi sanitari. In P. Berta, M. Castelli, G. Vittadini, A. Zangrandi, & C. Zocchetti (Eds.), *Sussidiarietà e... salute. Rapporto sulla Sussidiarietà 2025/2026* (pp. 107–132). Fondazione per la Sussidiarietà. Retrieved July 31, 2026, from <https://www.sussidiarieta.net/cn4604/servizio-sanitario-criticita-e-riforme-perche-resti-universale.html>
- Di Novi, C., Piacenza, M., Robone, S., & Turati, G. (2019). Does fiscal decentralization affect regional disparities in health? Quasi-experimental evidence from Italy. *Regional Science and Urban Economics*, 78, 103465. <https://doi.org/10.1016/j.regsciurbeco.2019.103465>
- Erreygers, G. (2009). Correcting the concentration index. *Journal of Health Economics*, 28(2), 504–515. <https://doi.org/10.1016/j.jhealeco.2008.02.003>
- Erreygers, G., & Van Ourti, T. (2011). Measuring socioeconomic inequality in health, health care and health financing by means of rank-dependent indices: A recipe for good practice. *Journal of Health Economics*, 30(4), 685–694. <https://doi.org/10.1016/j.jhealeco.2011.04.004>
- Ferré, F., de Belvis, A. G., Valerio, L., Longhi, S., Lazzari, A., Fattore, G., Ricciardi, W., & Maresso, A. (2014). Italy: Health system review [PMID: 25471543]. *Health Systems in Transition*, 16(4), 1–168. Retrieved July 31, 2026, from

- <https://eurohealthobservatory.who.int/publications/i/italy-health-system-review-2014>
- Giorgi, C. (2023). A history of Italy's health policy from the Republic to the new century. *Modern Italy*, 28(1), 1–17. <https://doi.org/10.1017/mit.2022.50>
- Glorioso, V., & Subramanian, S. V. (2014). Equity in access to health care services in Italy. *Health Services Research*, 49(3), 950–970. <https://doi.org/10.1111/1475-6773.12128>
- Istituto Nazionale di Statistica (Istat). (2021). *Rapporto bes 2020: Il benessere equo e sostenibile in italia — capitolo 12, qualità dei servizi* (tech. rep.). Istat. <https://www.istat.it/wp-content/uploads/2021/03/12.pdf>
- Istituto Nazionale di Statistica (Istat). (2024). *Rapporto bes 2023: Il benessere equo e sostenibile in italia — capitolo 12, qualità dei servizi* (tech. rep.). Istat. <https://www.istat.it/wp-content/uploads/2024/04/12.pdf>
- Lispi, L., Nuzzo, A., Di Fonzo, G., Gallo, M., Giordani, C., Ceccolini, C., Rossi, A., Campolongo, A., De Vito, T., Mennini, F. S., & Cicchetti, A. (2025). *Monitoraggio dei LEA attraverso il Nuovo Sistema di Garanzia – Relazione 2023* (Relazione) (Dipartimento della programmazione, dei dispositivi medici, del farmaco e delle politiche in favore del Servizio sanitario nazionale). Ministero della Salute. Roma. Retrieved July 31, 2026, from https://www.salute.gov.it/new/sites/default/files/2025-08/Relazione%5C%20NSG_31_07_2025_per_portale_copertine%5C%20%5C%281%5C%29.pdf
- Masseria, C., & Giannoni, M. (2010). Equity in access to health care in Italy: A disease-based approach. *European Journal of Public Health*, 20(5), 504–510. <https://doi.org/10.1093/eurpub/ckq029>
- O'Donnell, O., van Doorslaer, E., Wagstaff, A., & Lindelow, M. (2008). *Analyzing health equity using household survey data: A guide to techniques and their implementation*. The World Bank. <https://doi.org/10.1596/978-0-8213-6933-3>
- OECD & European Observatory on Health Systems and Policies. (2025). *Italy: Country Health Profile 2025* (tech. rep.). OECD Publishing, European Observatory on Health Systems, and Policies. Paris, Brussels. Retrieved July 31, 2026, from <https://eurohealthobservatory.who.int/publications/m/italy-country-health-profile-2025>
- Olsson, U. (1979). Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika*, 44(4), 443–460. <https://doi.org/10.1007/BF02296207>
- Signorelli, C., Odone, A., Oradini-Alacreu, A., & Pelissero, G. (2020). Universal Health Coverage in Italy: Lights and shades of the Italian National Health

- Service which celebrated its 40th anniversary. *Health Policy*, 124(1), 69–74. <https://doi.org/10.1016/j.healthpol.2019.11.002>
- Toth, F. (2016). The Italian NHS, the public/private sector mix and the disparities in access to healthcare. *Global Social Welfare*, 3(3), 171–178. <https://doi.org/10.1007/s40609-016-0072-2>
- Vyas, S., & Kumaranayake, L. (2006). Constructing socio-economic status indices: How to use principal components analysis. *Health Policy and Planning*, 21(6), 459–468. <https://doi.org/10.1093/heapol/czl029>
- Wagstaff, A. (2005). The bounds of the concentration index when the variable of interest is binary, with an application to immunization inequality. *Health Economics*, 14(4), 429–432. <https://doi.org/10.1002/hec.953>
- Wagstaff, A., Paci, P., & van Doorslaer, E. (1991). On the measurement of inequalities in health. *Social Science & Medicine*, 33(5), 545–557. [https://doi.org/10.1016/0277-9536\(91\)90212-U](https://doi.org/10.1016/0277-9536(91)90212-U)
- Wagstaff, A., & van Doorslaer, E. (2000). Measuring and testing for inequity in the delivery of health care. *Journal of Human Resources*, 35(4), 716–733. <https://doi.org/10.2307/146369>
- Wagstaff, A., van Doorslaer, E., & Watanabe, N. (2003). On decomposing the causes of health sector inequalities with an application to malnutrition inequalities in vietnam. *Journal of Econometrics*, 112(1), 207–223. [https://doi.org/10.1016/S0304-4076\(02\)00161-6](https://doi.org/10.1016/S0304-4076(02)00161-6)

List of Legislation

Constitution of the Italian Republic, Article 32 (1948). *Costituzione della Repubblica Italiana, art. 32.*

Law No. 833 of 23 December 1978. *Istituzione del Servizio sanitario nazionale.* . *Gazzetta Ufficiale* No. 360, 28 December 1978.
urn:nir:stato:legge:1978-12-23;833

Legislative Decree No. 502 of 30 December 1992. *Riordino della disciplina in materia sanitaria, a norma dell'articolo 1 della legge 23 ottobre 1992, n. 421.* urn:nir:stato:decreto.legislativo:1992-12-30;502

Legislative Decree No. 517 of 7 December 1993. *Modificazioni al decreto legislativo 30 dicembre 1992, n. 502, recante riordino della disciplina in materia sanitaria, a norma dell'articolo 1 della legge 23 ottobre 1992, n. 421.*
urn:nir:stato:decreto.legislativo:1993-12-07;517

Regional Law of Lombardy No. 31 of 1997. *Norme per il riordino del Servizio sanitario regionale e sua integrazione con le attività dei servizi sociali.*

Constitutional Law No. 3 of 18 October 2001. *Modifiche al titolo V della parte seconda della Costituzione.*
urn:nir:stato:legge.costituzionale:2001-10-18;3

Law No. 191 of 23 December 2009, Article 2, paragraph 68. *Disposizioni per la formazione del bilancio annuale e pluriennale dello Stato (legge finanziaria 2010).* urn:nir:stato:legge:2009-12-23;191

Ministerial Decree of 12 March 2019. *Nuovo sistema di garanzia per il monitoraggio dell'assistenza sanitaria.*

Legislative Decree No. 446 of 15 December 1997. *Istituzione dell'imposta regionale sulle attività produttive, revisione degli scaglioni, delle aliquote e delle*

detrazioni dell'Irpef e istituzione di una addizionale regionale a tale imposta.
urn:nir:stato:decreto.legislativo:1997-12-15;446

Legislative Decree No. 229 of 19 June 1999. *Norme per la razionalizzazione del Servizio sanitario nazionale, a norma dell'articolo 1 della legge 30 novembre 1998, n. 419..* urn:nir:stato:decreto.legislativo:1999-06-19;229

Legislative Decree No. 56 of 18 February 2000. *Disposizioni in materia di federalismo fiscale, a norma dell'articolo 10 della legge 13 maggio 1999, n. 133.* urn:nir:stato:decreto.legislativo:2000-02-18;56

Law No. 311 of 30 December 2004, Article 1, paragraph 180. *Disposizioni per la formazione del bilancio annuale e pluriennale dello Stato (legge finanziaria 2005).* urn:nir:stato:legge:2004-12-30;311

Law No. 42 of 5 May 2009. *Delega al Governo in materia di federalismo fiscale, in attuazione dell'articolo 119 della Costituzione.*
urn:nir:stato:legge:2009-05-05;42

Appendix

Robustness Check: Service-Specific Estimation Samples in 2023

The 2023 wave contains an unusually large number of missing observations for the emergency-department outcome. As discussed in Section 2.1.3, the main analysis retains a common complete-case sample within each wave so that the population underlying the four national service estimates is identical. In 2023 this common sample contains 31,098 observations. As a robustness check, the four national models are re-estimated using service-specific complete-case samples: the common demographic and covariate eligibility restrictions are retained, but an observation is no longer excluded solely because the outcome for another healthcare service is missing.

Table 4.1 compares the 2023 national Erreygers index, $E(y)$, and the need-adjusted index, ENA, under the main common-sample definition and the service-specific alternative.

Table 4.1: 2023 national indices: common and service-specific estimation samples

Service	N	$E(y)$ (SE)	p -value	ENA (SE)	p -value
<i>Panel A: Main common sample</i>					
Specialist visits	31,098	0.0585 (0.0080)	< 0.001	0.1243 (0.0086)	< 0.001
Diagnostic examinations	31,098	0.0679 (0.0069)	< 0.001	0.1241 (0.0068)	< 0.001
Blood tests	31,098	0.0268 (0.0070)	< 0.001	0.0932 (0.0070)	< 0.001
Emergency department	31,098	−0.0072 (0.0041)	0.077	−0.0005 (0.0042)	0.903
<i>Panel B: Service-specific samples</i>					
Specialist visits	32,210	0.0570 (0.0071)	< 0.001	0.1231 (0.0069)	< 0.001
Diagnostic examinations	32,205	0.0674 (0.0067)	< 0.001	0.1237 (0.0066)	< 0.001
Blood tests	32,188	0.0248 (0.0076)	0.001	0.0920 (0.0069)	< 0.001
Emergency department	31,183	−0.0094 (0.0039)	0.016	−0.0026 (0.0038)	0.500

Notes: $E(y)$ denotes the unadjusted Erreygers concentration index and ENA the need-adjusted Erreygers index. Standard errors are bootstrap standard errors based on 100 replications stratified by socioeconomic-status quintile; reported p -values are the corresponding normal-based bootstrap p -values. Sampling weights are applied consistently in the estimation. The main common sample requires complete information for the variables used across the four utilisation analyses, whereas the service-specific specification imposes the outcome-specific complete-case restriction separately for each service.

For specialist visits, diagnostic examinations and blood tests, moving from the common sample to the service-specific sample recovers slightly more than one thousand observations per service. The resulting estimates are nearly unchanged. Specialist-visit ENA moves from 0.1243 to 0.1231, diagnostic- examination ENA from 0.1241 to 0.1237, and blood-test ENA from 0.0932 to 0.0920. The sign, substantive magnitude and statistical significance of the need-adjusted estimates are therefore unaffected by the sample definition.

For the emergency department, the service-specific specification recovers only 85 observations because both samples must exclude respondents for whom the emergency-department item itself is missing. The need-adjusted estimate remains close to zero and statistically indistinguishable from zero, changing from -0.0005 ($p = 0.903$) in the common sample to -0.0026 ($p = 0.500$) in the service-specific sample. The unadjusted index is more sensitive: it moves from -0.0072 ($p = 0.077$) to -0.0094 ($p = 0.016$). Thus, the substantive 2023 conclusion for horizontal inequity in emergency-department utilisation is robust to the alternative sample definition, while inference for the unadjusted concentration index is more sample-sensitive.