

防止返贫动态监测及监测状态变换成因探究

——基于多维贝叶斯网络分类器模型的分析

平卫英 郭玉帑 黄 斐

摘要：立足现阶段防止返贫动态监测和帮扶的迫切需要，本文基于江西省2021—2023年的三期追踪调查数据，借助多维贝叶斯网络分类器模型预测农户返贫致贫风险。在此基础上，本文采用标准下行半偏差调整后的概率测量农户多维贫困脆弱性水平，更有效地甄别2020年后返贫治理的监测对象。进一步，本文利用有序Logit模型考察了农户防止返贫监测状态动态变换的风险诱因。研究表明：多维贝叶斯网络分类器模型能够较好地估算农户返贫致贫风险，对农户监测状态的整体预测精度达90%以上。而且，该模型可以根据实时数据动态更新农户的监测状态，为制定帮扶政策提供坚实的决策支持。人力资本和金融资本的减少是农户难以摆脱返贫风险状态的重要因素，而参与社会组织、改善住房条件等能够有效降低返贫致贫风险。此外，本文对防止返贫监测和完善帮扶机制提出了相关建议。

关键词：防止返贫动态监测 多维贝叶斯网络分类器 下行风险 动态变换

中图分类号：F328 **文献标识码：**A

一、引言

习近平在全国脱贫攻坚总结表彰大会上庄严宣告中国脱贫攻坚战取得全面胜利，中国完成了消除绝对贫困的艰巨任务。2024年12月召开的中央农村工作会议强调，要提升防止返贫致贫监测帮扶效能。然而，在部分经济发展基础较为薄弱的脱贫地区，农户容易返贫，部分边缘户和部分因突发事件使生活遭受严重冲击的农户也面临潜在的致贫风险。上述群体的生活状况具有不稳定性 and 脆弱性，需持续监测与帮扶。在2019年的重庆考察中，习近平主持召开解决“两不愁三保障”突出问题座谈会

【资助项目】 国家社会科学基金重大项目“后扶贫时代中国城乡相对贫困统计测度与治理机制研究”（编号：20&ZD131）；江西省宣传思想文化领域高层次人才专题项目“中国农村低收入群体识别、监测与治理机制研究”（编号：23ZXRC13）；江西省研究生创新专项资金项目“防止返贫动态监测及动态转换成因研究”（编号：YC2024-B166）。

【作者信息】 平卫英，江西财经大学统计与数据科学学院、江西财经大学应用统计研究中心；郭玉帑（通讯作者）、黄斐，江西财经大学统计与数据科学学院，电子邮箱：tarevachine19@163.com。

并发表重要讲话，强调要把防止返贫摆在重要位置，适时组织对脱贫人口开展“回头看”^①。2021年，习近平在全国脱贫攻坚总结表彰大会上指出，对易返贫致贫人口要加强监测，做到早发现、早干预、早帮扶^②，为稳固脱贫攻坚成果提供了明确的方向。第十四届全国人民代表大会第二次会议《政府工作报告》亦明确指出，要加强防止返贫监测和帮扶工作，确保不发生规模性返贫^③。由此可见，“动态清零”是中国打赢脱贫攻坚战的宝贵经验，在巩固拓展脱贫攻坚成果同乡村振兴有效衔接的阶段，构建高效且可靠的返贫风险监测与预警机制显得尤为重要，对防范规模性返贫、巩固拓展脱贫攻坚成果具有重大意义。

近年来，政府发布了一系列政策文件，明确提出了加强防止返贫动态监测的具体要求，围绕坚决守住不发生规模性返贫的底线进行了全方位、多角度的工作部署。不过，开展防止返贫监测工作仍面临诸多难题，政策的实施仍需进一步深化。一方面，以往对帮扶对象的识别与帮扶机制大多基于过去和当前的福利水平，主要侧重对已知福利被剥夺的事后应对（贾男和王赫，2022）。然而，防止规模性返贫的重点在于预防和杜绝：预防是对规模性返贫的诱导因素进行事前干预，实现源头控制；杜绝是及时识别风险并进行有效干预，遏制贫困再生。因此，防止返贫监测应当在风险发生前、中、后各个阶段充分发挥监测预警功能（左停和李泽峰，2022），而不能仅局限于事后评估。此外，目前的监测方法并不一致。国外学者用脆弱性概念测度返贫风险（Ligon and Schechter, 2003; Mahanta and Das, 2021）。Gallardo 多年致力于对贫困脆弱性的研究，将贫困脆弱性从一维测量扩展到多维（Gallardo, 2018），并运用多维贝叶斯网络分类器估计多维贫困的条件概率，将标准下行半偏差作为风险参数来识别拉丁美洲的弱势家庭（Gallardo et al., 2024）。国内学者多基于 VEP 测度（吕光明等，2021）、综合评价分析（王兆林和王洁仪，2024）等方法测度返贫风险。复杂多元的致贫风险会冲击不同的脆弱群体，对精准识别返贫风险提出了更高要求，学者试图引入综合学习算法以对规模性返贫识别、风险预测等提供数据与方法上的新思路和新的技术手段，随机森林（Hu et al., 2022）、神经网络（Zhang et al., 2022）等机器学习方法可以减少人为主观性。在扶贫工作与防止返贫监测实践中，地方政府普遍遵循“先监测、后预警、再帮扶”的思路开展实际工作。各级农业农村部门承担防止返贫监测的主要责任，依据当地的监测标准，通过一系列程序与环节识别监测对象并将其纳入动态监测系统。此项认定工作要求，从农户申报、干部排查、部门预警发现风险线索到完成监测对象识别认定一般不超过15天。然而，帮扶政策落到实处存在一定滞后性，监测和预警的逻辑关系尚未清晰界定，引发了资源过度投入、操作效率不高等问题（左停和李泽峰，2022）。有研究强调，应进一步缩短返贫风险识别与帮扶措施落地的时间差（章文光等，2020），提高防止返贫监测工作的精准性。

另一方面，防止返贫监测预警具有过程性和动态性（左停和李泽峰，2022）。多数学者一致认为，

^①参见《习近平在重庆考察并主持召开解决“两不愁三保障”突出问题座谈会》，https://www.gov.cn/xinwen/2019-04/17/content_5383915.htm。

^②参见《在全国脱贫攻坚总结表彰大会上的讲话》，<http://politics.people.com.cn/n1/2021/0226/c1024-32037098.html>。

^③参见《政府工作报告》，https://www.gov.cn/yaowen/liebiao/202403/content_6939153.htm。

返贫是一个持续的过程，在时间维度具有动态变换的特性，反映的是原贫困人口在得到帮扶脱离贫困后又因为种种原因再次陷入贫困的现象（李博等，2018；张全红等，2019）。为避免这种恶性循环，需要在返贫识别过程中准确把握不同时期个人或家庭的具体情况，依据返贫户与其他农户的特征差异针对性地进行帮扶（胡世文和曹亚雄，2021）。多数学者利用系统聚类、Tribble-Hurdle 模型等方法或基于贫困脆弱性、可持续生计理论框架等对返贫风险进行了大量研究（汪三贵和周园翔，2022；赵普等，2022；孟婷，2023；刘胜强和赵雪燕，2024），认为影响返贫风险的因素包括人力资本和物质资本不足、自然灾害、政策调整和市场变化等。大量文献评估了精准扶贫的政策效果（刘子宁等，2019；李芳华等，2020；黄薇和曹杨，2022）。在最新研究中，兰宇和张鹏（2024）探讨了不同类型的帮扶手段如何有效防止返贫，张鹏龙等（2024）通过模糊断点回归方法考察了防止返贫帮扶政策的劳动收入效应。不过，上述研究均未从动态视角探讨影响防止返贫监测状态变换的因素。

从以上文献可以发现，在研究内容上，现有对返贫的研究往往集中在返贫现象后期治理层面，对返贫的事前监测与阻断机制的关注较少。随着脱贫攻坚战的全面胜利，当前的工作重心已从消除绝对贫困转向防止返贫。尽管部分研究开始关注脱贫户的返贫风险，但相关研究成果不够丰富，对政策调整优化方案的系统性探讨较少。在研究方法上，现有研究多基于 VEP 测度方法、综合评价分析等方法测度返贫风险。这些方法对数据分布的要求过高、主观性较强，无法有效识别复杂多样的返贫致贫风险并评估其严重程度。而且，部分机器学习方法对变量间依赖关系的解释性不强，对数据量要求较高。在研究视角上，现有研究大多衡量家庭当前的返贫状态，未考虑家庭监测状态动态变换的影响因素，不利于系统性、前瞻性、针对性地制定帮扶政策，对帮扶政策减贫效应的评估也难以开展。

本文的边际贡献有两个方面：第一，所用监测方法可动态更新农户的监测状态。本文借助多维贝叶斯网络分类器直观描绘各监测变量的依赖关系。该方法能利用实时数据动态更新贝叶斯网络条件概率值，可以实现对监测对象的滚动预测，与风险评价预警流程的匹配度较高。而且，本文采用经过标准下行半偏差调整后的返贫致贫风险判断监测对象，避免因无法充分考量尾部风险而低估农户遭遇极端不利情况的可能，有效提升监测工作的效率与精准度。第二，从监测状态动态变换角度探究返贫致贫因素。本文关注农户监测状态随时间的变化，探究生计资本对农户多期监测状态的可能影响，并结合家庭结构特征提供监测状态动态变换的经验证据，进一步明确返贫致贫风险的预防措施。

二、防止返贫监测机制与“脆弱性—可持续生计”的理论分析

现阶段，中国的防止返贫监测机制强调依托大数据平台、充分发挥信息化手段并发动多部门协作和信息共享，形成常态化比对、动态化监测、精准化帮扶的防止返贫长效机制。然而，从现有研究和防止返贫监测的实地调研可以发现，当前的防止返贫监测机制仍存在问题：其一，部分地区存在形式主义问题，重复排查、过度填表等增加了基层工作负担，影响监测效率（左停和李泽峰，2022）。其二，部分帮扶监测工作仍存在“应纳未纳”问题，监测和帮扶工作的执行力和精准度仍需提升（谢治菊，2023）。其三，部分地方帮扶措施过于单一，缺乏针对性和灵活性，未能完全满足脱贫家庭的多元需求（李玉山等，2021）。因此，如何突破防止返贫监测的“瓶颈”，保证农户返贫致贫风险监控

测的持续性，明确防止增加返贫致贫风险应采取的措施，精准匹配帮扶需求是本文研究的关键问题。

脆弱性分析框架是探究脆弱性的成因及其影响因素的基本工具。在贫困研究领域，学者多从“风险—应对能力”两个维度出发将脆弱性界定为“家庭面临某些风险的可能，并且是由于遭遇风险而导致财富损失或生活质量下降到某一社会公认的水平之下的可能”（韩峥，2004）。脆弱性程度取决于风险的特点和家庭的风险应对能力，社会服务体系、公共政策等对提升家庭风险应对能力有重要作用。脆弱性分析框架经历了从二维的“外部—内部”“敏感—恢复力”（Chambers, 1989; Moser, 1998）等理论框架到三维的“暴露—敏感—适应”（Polsky et al., 2007）理论框架的转变，对脆弱性的理论分解与风险因素紧密相关。Dercon（2001）提出的“风险—脆弱性”分析框架明确探讨了风险、脆弱性与贫困的相互关联（见表1）。该框架以家庭为基本分析单位，根据可持续生计分析框架对资产的重视，将家庭福利水平提升与家庭能力增强的过程视为资产向收入转化和收入向“福利—能力”转化两个相互衔接的过程，这个过程中的任意环节都可能受到各类风险的冲击。“风险—脆弱性”分析框架对农户生计过程的风险进行了清晰地阐释，是政府及相关部门实施帮扶计划的重要工具之一。

表1 Dercon（2001）的脆弱性评估分析框架

资产	收入	福利—能力
<ul style="list-style-type: none"> ● 人力资本、劳动力 ● 物质或金融资本 ● 公共资源和公共物品 ● 社会资本 	<ul style="list-style-type: none"> ● 活动和资产的回报 ● 资产处置的收益 ● 储蓄、信贷和投资回报 ● 转移性收入 	<ul style="list-style-type: none"> ● 消费 ● 营养 ● 健康 ● 教育
风险案例（a）	风险案例（b）	风险案例（c）
<ul style="list-style-type: none"> ● 由于健康问题而失去技能或失业 ● 土地产权不安全，其他资产权益的不确定，气候、战争或灾难造成的资产损失 ● 通货膨胀、股票市场和汇率崩溃等导致的金融资产贬值 ● 	<ul style="list-style-type: none"> ● 气候冲击、疾病、冲突导致的产出风险 ● 收入和资产价格的变化 ● 储蓄和投资的回报风险 ● 非正式安排效力的不确定，例如转移性支付不能兑现 ● 获取机会的相关信息和知识的不完备 ● 政策环境风险 ● 	<ul style="list-style-type: none"> ● 食品市场价格风险 ● 食品可及性和供给风险 ● 在健康和领域公共供给质量的不稳定 ● 健康和计划的不确定 ● 健康和营养知识的不全面 ●

注：表中左右箭头表示资产、收入、福利—能力三者之间的转化方向，向上的箭头表示风险对各环节的冲击。

防止返贫监测预警的本质是预先对风险进行研判（左停和李泽峰，2022），而脆弱性理论不仅关注贫困的实然状态，还可以预测家庭在面对未来不确定性时的潜在风险，为预测农户的返贫致贫风险提供理论支撑。此外，脆弱性是具有返贫风险的主体缺乏基本抗贫能力的表现（严小燕，2022），脆弱性需从多维度深入探讨，返贫也受多重因素的影响（蒋南平和郑万军，2017）。因此，本文借鉴脆弱性理论与可持续生计理论的分析框架，从“收入—福利—能力”角度出发，结合现行返贫动态监测和帮扶机制的监测内容，通过收入、教育、健康、生活条件、工作与社会保障五个维度测度农户的多维贫困脆弱性，将脆弱性转化为具体的、可量化的指标以衡量农户的返贫致贫风险，进而根据风险程

度判断农户是否为监测对象^①。为深入理解监测状态变换的频率、不稳定程度的演变规律和影响因素，参考可持续生计理论（DFID，2000），本文从人力资本、金融资本、物质资本、社会资本、自然资本五个维度分析农户生计稳定性、适应性和恢复力。同时，本文增加了户主和家庭特征变量，从侧面反映农户的收入、福利、能力，以探究相关特征变量对监测状态变换的异质性影响，捕捉监测状态变换过程中农户的多元需求。经由对农户监测状态的预判和对农户监测状态变换成因的探究，同时结合孟婷（2023）研究中提到的返贫识别过程，本文设计了农户返贫致贫风险监测的研究框架，如图1所示。

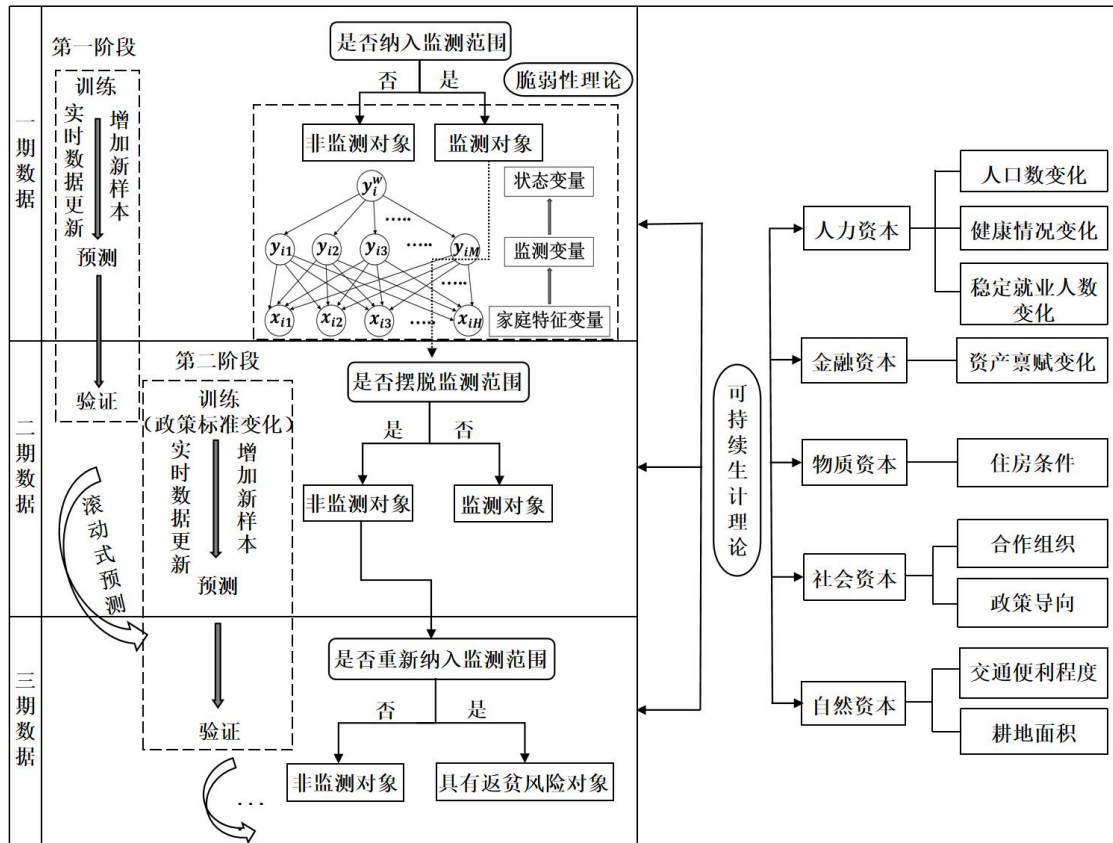


图1 监测过程和研究框架

基于概率推理的贝叶斯网络模型可以将不同脆弱性诱发因素转化为网络中的节点，连接节点的边表现为变量之间的依赖关系，由此形成的贝叶斯网络拓扑图能清晰表现这些因素如何通过不同路径影响监测状态。由 Bielza et al.（2011）首次提出的多维贝叶斯网络分类器能够在给定监测状态的前提下，根据农户特征的条件概率分布情况，结合分类状态的先验概率信息，估算农户处于监测状态的后验概率。而且，多维贝叶斯网络分类器可以通过监测状态的后验概率揭示脆弱性因素对返贫致贫风险的贡献率。返贫风险是脆弱性在时间维度的动态反映（周迪和王明哲，2019），能根据实时数据动态更新

^①此处所指监测对象与实际防止返贫监测工作中的监测对象不同，在调研过程中调查人员会根据农户实际情况进行风险标记，本文通过“是否被识别为监测户”和“风险是否已消除”两个问题判断农户的监测状态。

监测状态的后验概率是多维贝叶斯网络分类器的优势。当新观测或已观测农户的监测变量（收入、健康等）变动时，训练好的多维贝叶斯网络分类器可通过已学到的依赖关系重新计算农户的返贫致贫风险。因此，在脆弱性理论框架下将多维贝叶斯网络分类器作为防止返贫监测的方法，不仅能减轻基层工作负担、提升监测效率，还能灵活调整农户的监测状态。此外，本文根据三期监测状态将农户分为安全稳定户、轻度易返贫致贫户、中度易返贫致贫户和重度易返贫致贫户四种，利用有序 Logit 模型探究农户监测状态变换的驱动因素，进而根据农户的多元特征提出更具针对性的干预措施和帮扶政策。

三、研究设计

（一）研究方法

部分防止返贫监测实践存在重视事后评估、监测对象识别认定低效等问题，因此，对返贫致贫风险的预测和及时干预尤为重要。本文通过贝叶斯网络的参数学习量化脆弱性诱发因素与监测状态的关系，引入多维贝叶斯网络分类器预测农户潜在返贫致贫风险。多维贝叶斯网络分类器可以利用实时数据动态更新贝叶斯网络结构和概率分布，及时反映农户当前的风险状况，提高风险评估的时效性。经过标准下行半偏差调整后的返贫致贫风险分析可以将风险在判别阈值附近而被忽视的边缘群体纳入监测范围，在提高资源利用效率、减少冗余投入的同时实施前置性干预策略。

1. 通过多维贝叶斯网络分类器测量返贫致贫风险。贝叶斯网络是将概率论与图论结合起来的数学模型（Pearl, 1985）。该网络由 $B(G, X)$ 定义， G 是由节点集合和边集合组成的有向无环图，节点代表随机变量^①，箭头首尾两端分别连接子节点和父节点（见图2），边表示子节点与父节点的条件概率依赖关系，依赖关系的强度通过附加于父节点和子节点上的条件概率表示。假设网络中的随机变量为 x_1, \dots, x_J ， $P(x_j | \pi(x_j))$ 表示每个随机变量的条件概率， X 的联合概率分布可以形式化如下：

$$P(x_1, \dots, x_J) = \prod_{j=1}^J P(x_j | \pi(x_j)) \quad (1)$$

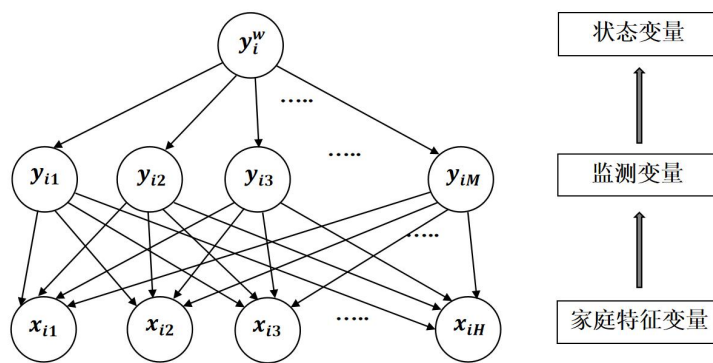


图2 基于多维贝叶斯网络分类器测量返贫致贫风险的基本思路

贝叶斯网络分类器是贝叶斯网络的一种应用形式，旨在根据一组特征变量预测父变量的分类结果。

^①本小节对变量和节点不做专门区分，两者可以混用。

多维贝叶斯网络分类器能够处理包含多层类变量的复杂分类问题，并允许变量之间存在条件依赖关系。这放宽了贝叶斯网络分类器变量之间必须完全独立的假设，拓宽了贝叶斯网络分类器的应用范围，提升了处理复杂数据结构和依赖关系的能力。

本文参照 Gallardo (2024) 的研究，将专家在多维贫困领域的经验与当前的防返贫动态监测和帮扶机制的监测内容结合起来，构建能够反映农户状态变量、监测变量、家庭特征变量之间关系的多维贝叶斯网络分类器拓扑图（见图 2）。图中三种变量的依赖关系之所以是反向的，是为了从每个子变量的条件概率获取信息来探究父变量的分类概率，再通过贝叶斯定理的反向因果关系完成推理。图 2 中，状态变量 y_i^w 表示第 i 户家庭是否为监测对象。如果是监测对象，则将状态变量赋值为 1，反之则赋值为 0。状态变量的取值取决于监测变量（福利维度），如果第 i 户家庭在第 m 个维度被剥夺，则将第 m 个监测变量赋值为 1，反之则赋值为 0。状态变量和监测变量都是二元变量，均服从伯努利分布。监测变量的取值取决于家庭特征变量，家庭特征变量可以有两个以上的分类结果，可以将家庭特征信息通过监测变量传递给状态变量进而影响状态变量的分类结果。

2. 基于多维贫困脆弱性预判监测对象。通过多维贝叶斯网络分类器可以得到状态变量的条件期望值： $\mu_i = E(y_i^w | y_i)$ 。这仅是对预期结果的估计，未充分考虑尾部风险，可能会低估极端不利情况的影响，无法精准识别返贫致贫风险群体。在平均风险行为框架下，要综合考虑观测目标的预期结果和风险两个因素（Gallardo, 2013）。该框架假定行为主体遵循“局部不敏感”和“风险规避”原则^①。在对贫困脆弱性的测量中，向下的风险至关重要（Povel, 2015）。因此，本文以标准下行半偏差^② σ_i^- （Gallardo, 2018）量化农户返贫致贫风险高于其条件期望值的程度，具体表达式如下：

$$\sigma_i^- = \sqrt{E\{\min[(y_i^w - \mu_i), 0]^2\}} \quad (2)$$

状态变量服从伯努利分布，其标准下行半偏差是状态变量后验概率的函数，转化后形式如下：

$$\sigma_i^- = \sqrt{p_i^2(1 - p_i)} \quad (3)$$

(3) 式中， p_i 是多维贝叶斯网络分类器估计的第 i 户家庭不是监测对象的后验概率。为了将预期结果和风险参数整合到一个表达式中，本文将 Ogryczak and Ruszczyński (2001) 用到的平均风险权衡参数^③ γ 作为调节因子，以调节某些风险因素对整体下行风险的影响。例如，如果投资者对某一特定类型的风险特别敏感，他们可能会增加与该风险相关的权衡参数，以便更准确地反映这种风险在整

^① “局部不敏感”和“风险规避”原则分别指行为主体对增加消费和减少风险的偏好，这里指在福利结果中对较高的平均值和较低的向下风险的偏好。

^② 标准下行半偏差是标准差的一种变形，通常用于计算资产或投资组合收益低于期望值的部分的标准差，只关注损失的部分，即只测量向下的偏差波动。

^③ $\gamma < 0$ 表示风险爱好， $\gamma = 0$ 表示风险中性， $\gamma > 0$ 表示风险厌恶。此外， $\gamma \in (0, 1]$ ，反映了风险相对于均值的权重。数值越大，对规避风险的估值越高。因此， $\gamma = 1$ 意味着规避风险与提高均值同等重要。

体下行风险中的重要性。状态变量 y_i^w 的条件期望值经过标准下行半偏差调整后变为 μ_i^{ra} ：

$$\mu_i^{ra} = \mu_i - \gamma\sigma_i^- = p_i - \gamma\sqrt{p_i^2(1-p_i)} \quad (4)$$

这一调整不仅考虑了农户在一般情况下的经济稳定程度，而且也将下行风险考虑在内，有助于更全面地了解农户面对不利环境时的脆弱性，为制定有效的前置性干预策略提供科学依据。

在平均风险行为框架下，(4) 式允许建立一个完整的偏好预排序，从而用标准下行半偏差调整后的期望与任意一对农户的状态变量 y_i^w 、 y_j^w 进行比较：

$$y_i^w > y_j^w \Leftrightarrow \mu_i^{ra} > \mu_j^{ra} \quad (5)$$

(5) 式的偏好关系符合二阶随机优势准则 (Ogryczak and Ruszczyński, 2001)。可根据以下标准识别监测对象：如果 $\mu_i^{ra} < z^v$ (z^v 为判别阈值)，则说明第 i 户家庭处于多维贫困脆弱性的风险状态。在多数分类算法的应用场景中，判别阈值默认为 0.5。这一设定基于正负类别先验概率均匀分布的假设。然而，现实中的类别分布往往是非平衡的，继续以 0.5 为判别阈值可能会使分类预测效果不佳 (Zou et al., 2016)。本文通过机器学习的五折交叉验证法并结合网格搜索法调整贝叶斯拟合参数和判别阈值，既能使判别阈值的选择更具说服力，也能确保监测对象的比例位于合理区间内。

3. 多维贫困脆弱群体的汇总和分解。在对多维贫困脆弱群体进行判别后，需要构建汇总框架度量并呈现多维贫困脆弱群体的总体情况。本文在判断脆弱个体时使用了判别阈值，参考 A-F 多维贫困指数 (Alkire and Foster, 2011)，用双界限法定义剥夺矩阵的方式量化脆弱群体。为此，将农户返贫致贫风险与判别阈值的差距 g_i^α 定义如下：

$$g_i^\alpha = \begin{cases} \left(\frac{z^v - \mu_i^{ra}}{z^v} \right)^\alpha, \forall \mu_i^{ra} < z^v \\ 0, \forall \mu_i^{ra} \geq z^v \end{cases} \quad (6)$$

为综合评估农户脆弱性的差异程度，本文以 FGT 指数 (Foster et al., 1984) 为度量工具，以此表明农户返贫致贫风险与判别阈值的关系，从而衡量多维贫困脆弱性的广度、深度和强度，具体如下：

$$V_\alpha = \frac{1}{n} \sum_{i=0}^n g_i^\alpha, \quad \alpha \geq 0 \quad (7)$$

(7) 式中：当 $\alpha=0$ 时， V_0 表示多维贫困脆弱群体占总样本的比例，即多维贫困脆弱性的广度；当 $\alpha=1$ 时， V_1 表示多维贫困脆弱性与判别阈值的差距，即多维贫困脆弱性的深度；当 $\alpha=2$ 时， V_2 表示多维贫困脆弱性与判别阈值的加权差距，即多维贫困脆弱性的强度。此外，可将 V_α 分解为高风险脆弱性 V_α^P 和边缘风险脆弱性 V_α^R 两个指标，具体如下：

$$V_\alpha = V_\alpha^P + V_\alpha^R \quad (8)$$

$$V_\alpha^P = \frac{1}{n} \sum_{i=0}^n g_i^\alpha I_{\mu_i \leq z^v}, \quad \alpha \geq 0 \quad (9)$$

$$V_{\alpha}^R = \frac{1}{n} \sum_{i=0}^n g_i^{\alpha} I_{z^v < \mu_i \leq (z^v + \gamma \sigma^-)}, \alpha \geq 0 \quad (10)$$

(9) ~ (10) 式中, $I_{\mu_i \leq z^v}$ 和 $I_{z^v < \mu_i \leq (z^v + \gamma \sigma^-)}$ 分别为判断高风险脆弱群体和边缘风险脆弱群体的示性函数。经过标准下行半偏差调整后得到的监测范围通常要大于实际的监测范围。多维贝叶斯网络分类器模型对 y_i^w 的预测准确率越高, 说明监测对象比例越接近高风险脆弱性的广度。此时, 高风险脆弱群体应被优先纳入重点帮扶计划, 持续对其进行密切关注。当然, 也需关注边缘风险群体并为其提供适当帮扶。此外, 还可测算各监测变量对多维贫困脆弱性的贡献程度。假设监测变量的权重为 w_m , 被识别为监测对象的农户数量记为 Q , 得到第 m 个监测变量对多维贫困脆弱性的贡献率 β_m :

$$\beta_m = \frac{\sum_{i=1}^Q w_m y_{im}}{\sum_{m=1}^M \sum_{i=1}^Q w_m y_{im}} (m = 1, \dots, M) \quad (11)$$

(二) 数据来源和样本特征

本文所用数据来自江西省 2021—2023 年防止返贫监测调研数据, 该省份曾是全国脱贫攻坚主战场之一, 其脱贫经验为全国提供了宝贵的示范作用。为验证研究样本的可行性, 本文用全国性的收入指标与该省份的收入指标进行对比, 发现该省农村人均可支配收入的中位数略高于全国农村人均可支配收入的中位数, 但两者并无显著差异。为探究农户监测状态动态变换的诱因, 本文以连续 3 年均参与调查的农户为实证研究对象。在删除缺失值后, 共得到 3328 户农户的 9984 个观测值, 2021 年、2022 年和 2023 年的风险人群比例分别为 8.28%、5.78%、4.54%。风险群体以非建档立卡户中的边缘易致贫户、突发困难户和已脱贫建档立卡户中的脱贫不稳定户为主, 非风险群体主要是脱贫稳定户。

(三) 变量选取

1. 监测变量。本文参考 Dercon (2001) 的“风险—脆弱性”理论框架, 根据《中央农村工作领导小组关于健全防止返贫动态监测和帮扶机制的指导意见》(以下简称《意见》) 提到的监测范围, 从“收入—福利—能力”角度将风险因素转化为具体的监测指标, 重点监测目标对象的收入与支出、“两不愁三保障”落实情况、饮水安全等状况, 同时关注其在就医、就学、就业等方面的实际困难和潜在风险, 在数据可得的前提下归纳了收入、教育、健康、生活条件、工作与社会保障五个维度(详见表 2^①)。在收入方面, 参照李莹等(2021)的研究, 用全国农村居民人均可支配收入中位数的 50% 界定收入门槛值。该省农村居民人均可支配收入的中位数接近全国平均水平, 后文将进一步验证收入门槛值设置的合理性。教育和健康方面的指标均以家庭为单位, 分别判断所有家庭成员受教育程度是否均在初中以下和是否均处于非健康状态。在生活条件方面, 以《意见》强调的饮水、住房安全为监测指标, 并将家庭有无卫生厕所纳入分析(周强, 2021), 以此关注农村居民的卫生状况和生活环境问题。此外, 用入户路类型直观反映农户的生活条件。顾宁和刘洋(2021)的研究已证实入户路类型对贫困脆弱性的显著负向影响。在工作与社会保障方面, 劳动技能会影响家庭收入, 医疗保险能为家庭的高额医疗

^①篇幅所限, 表 2 只展示 2023 年数据的部分变量情况, 其余变量可见《中国农村经济》网站或中国知网本文附表 1。

费用提供经济保障从而降低因病返贫风险。此外，低收入老年群体从劳动力市场退出还会引发因老返贫问题（李锐等，2024），需要对老年人养老保险情况进行监测。在指标权重选取方面，为避免客观赋权法处理跨期数据时可能产生的权重波动问题，参照多位学者的研究（郭熙保和周强，2016；平卫英和王佳，2023），本文采用等权重法以确保权重分配的稳定性和多维贫困指数测度结果的可比性。

表2 防止返贫动态监测的监测指标及动态变换影响因素

状态变量						
变量名称	变量描述	2021年风险 群体比例	2022年风险 群体比例	2023年风险 群体比例		
监测状态	农户是否为风险群体：是=1，否=0	8.28%	5.78%	4.54%		
监测变量						
维度	变量名称	变量描述	权重	全部样 本均值	已脱贫建档 立卡户均值	非建档立 卡户均值
收入	收入水平	2021年、2022年、2023年全国农村居民人均可支配收入中位数的50%分别为8451元、8867元、9374元，低于当年收入中位数=1，否则=0	1/5	0.13	0.19	0.08
	教育	受教育程度	家庭成员受教育程度均在初中以下=1，否则=0	1/5	0.47	0.53
健康	健康状态	家庭中无身体健康成员=1，否则=0	1/5	0.28	0.35	0.22
生活条件	饮水安全	饮水是否安全：不安全=1，安全=0	1/20	0.00	0.00	0.00
	住房安全	房屋等级为B、C、D级=1，否则=0	1/20	0.13	0.14	0.12
	卫生状况	有无卫生厕所：无=1，有=0	1/20	0.02	0.03	0.01
	入户路类型	泥土路=1，砂石路或硬化路=0	1/20	0.06	0.08	0.04
工作与 社会保障	劳动技能	家庭成员中没有劳动力=1，否则=0	1/15	0.20	0.28	0.13
	医疗保险	家庭中至少有1人无基本医疗保险=1，否则=0	1/15	0.01	0.01	0.01
	养老保险	家中60岁以上的老人都无养老金=1，否则=0	1/15	0.16	0.19	0.14
家庭特征变量						
维度	变量名称	变量描述		全部样 本均值	已脱贫建档 立卡户均值	非建档立 卡户均值
户主特征	户主性别	女=2，男=1		1.23	1.22	1.23
	户主年龄	65岁及以上=6，55~64岁=5，45~54岁=4，35~44岁=3，25~34岁=2，25岁以下=1		4.89	4.96	4.81
	户主受教育程度	高级教育=4，中级教育=3，初级教育=2，文盲=1		2.11	2.07	3.95
家庭特征	人口结构	只有老年人=5，只有中青年=4，只有中青年和老年人=3，只有中青年和少年儿童=2，三代同堂=1		3.47	3.61	3.35
	搬迁情况	家庭是否为易地搬迁户：是=1，否=0		0.03	0.03	0.03
	所在村历史 贫困状态	家庭所在村是否为脱贫村：是=1，否=0		0.47	0.46	0.48

表2 (续)

动态变换影响因素					
维度	变量名称	变量描述	方向	均值	标准差
被解释变量	农户监测状态变换类型	重度易返贫致贫户=4, 中度易返贫致贫户=3, 轻度易返贫致贫户=2, 安全稳定户=1		1.14	0.45
人力资本	人口数变化	当期家庭人口数-上一期家庭人口数(人)	正向	0.02	0.17
	健康情况变化	家庭大病、慢性病或残疾人口变化情况: 增加=1, 未变化=0, 减少=-1	正向	0.01	0.19
	稳定就业人数变化	当期稳定就业人数-上一期稳定就业人数(人)	负向	0.78	1.21
金融资本	资产禀赋变化	当期家庭务工收入、财产性收入、生产经营性收入之和与上一期的差值(元)	负向	4741.12	14241.42
物质资本	住房条件	住房面积(平方米)	负向	116.86	74.67
社会资本	政策导向	收到政府各类补贴金额(元)	正向	7757.73	7226.30
	合作组织	是否参与农民专业合作社: 是=1, 否=0	负向	0.30	0.46
自然资本	耕地面积	家庭拥有耕地面积(亩)	负向	2.91	3.98
	交通便利程度	住所离村主干道距离(千米)	正向	0.28	0.58

注: ①关于住房安全变量, 根据《危险房屋鉴定标准》(JGJ125-2016), 房屋的危险程度按 A、B、C、D 等级划分, A 级为无危险构件, B 级为有个别结构构件是危险构件, C、D 级就是通常说的危房。②人口结构方面, 将家庭成员年龄划分为 3 个阶段, 0~14 岁为少年儿童, 15~59 岁为中青年人, 60 岁以上为老年人。

2. 家庭特征变量。为描述农户特征变量对各监测变量的影响, 在户主特征层面选取户主的性别、年龄、受教育程度等因素, 在家庭特征层面选取家庭人口结构、易地搬迁情况(宁静, 2018)揭示家庭的经济承受能力、劳动力状况及社会适应能力, 以家庭所在村是否为脱贫村反映外部扶贫政策的影响和家庭所处的社会经济环境(章文光等, 2020)。

3. 动态变换影响因素。在动态变换影响因素的选取中, 本文根据可持续生计理论从人力资本、金融资本、物质资本、社会资本、自然资本五个维度探究农户的生计稳定性、适应性和恢复力(详见表 2)。人力资本是农户生计的核心要素, 影响农户生计活动的质量与效率。本文用人口数变化直观描述家庭规模的扩大或缩小, 用家庭大病、慢性病或残疾人口的变化衡量家庭成员健康情况变化, 以直观反映因病返贫风险。将全年工作时间超过六个月的人视为稳定就业群体, 利用稳定就业人数与上期的差值观测稳定就业情况。在金融资本方面, 用当期家庭务工收入、财产性收入、生产经营性收入之和与上一期的差值反映家庭资产禀赋的变化, 以观测家庭是否因为收入骤减或支出骤增而影响监测状态。在物质资本方面, 以住房面积反映住房条件, 良好的住房条件对生计活动至关重要(闫啸等, 2022), 不仅能为农户提供稳定的生活基础, 还能为其从事高效的生计活动创造条件。在社会资本方面, 政策导向及是否参与合作组织是评估农户家庭生计适应能力的重要维度。帮扶政策与产业扶持政策不仅直接影响农户生计策略, 还通过影响资源分配和社会网络构建间接促进农户生计的多元化与稳定性。自然资本是农户生计的基础, 耕地面积不仅直接影响农业生产的规模与潜力, 还与生计活动可持续性紧

密相连。交通便利程度是衡量是否存在地理和信息隔阂的重要指标，用住所离村主干道的距离表示。

四、实证结果与分析

（一）多维贝叶斯网络分类器参数选择

为避免估计中的过拟合问题，提高模型参数的稳健性，本文采用 K-fold 交叉验证方法和网格搜索法对多维贝叶斯网络分类器模型进行参数调优。K-fold 交叉验证的主要思路是在样本中反复抽取 K 次一定比例的样本形成训练集，剩余的样本为测试集，每次抽样后在训练集上训练模型，再利用测试集评估模型精度。K-fold 交叉验证能够通过多次评估不同的样本，减少模型评估的波动性，在小样本情况下有助于避免过拟合问题（Messner et al., 2020）。网格搜索法则可以通过尝试每一种给定的参数组合来寻找最优的超参数组合。本文使用五折交叉验证评估模型的性能，假定每一期训练集和测试集数据量之比为 4 : 1，需要优化的模型参数为贝叶斯狄利克雷评分^①和判别阈值。图 3 展示了 2021—2023 年不同参数组合下五折交叉验证的平均精度。可以发现，不同年份最优模型的平均精度都在 90% 以上。

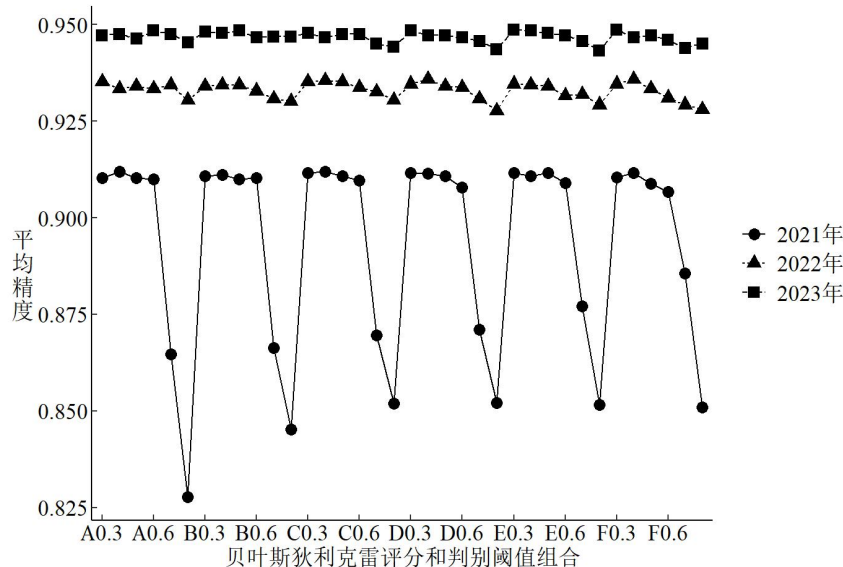


图 3 最优参数组合

注：横坐标对应多维贝叶斯网络分类器模型的参数组合，A、B、C、D、E、F 分别表示贝叶斯狄利克雷评分取值为 5、10、15、20、25、30，每个贝叶斯狄利克雷评分分别与判别阈值 0.3、0.4、0.5、0.6、0.7、0.8 结合形成不同的参数组合。

（二）多维贫困脆弱性的测度

本文基于调查数据中的风险标记和监测变量的剥夺情况，用多维贝叶斯网络分类器模型根据家庭的实际剥夺情况估计脱贫不稳定户或边缘易致贫户未来陷入多维贫困的可能性。该模型既关注那些在收入维度已遭受剥夺的家庭，也能进一步探讨这些家庭在其他非收入维度（如健康、教育、生活条件

^①贝叶斯狄利克雷评分是在已知结构时将样本数据的边际似然函数作为评分依据，计算时假定贝叶斯网络的参数服从狄利克雷形式的先验分布。

等)上的多维特征,而且,还能将在收入维度未遭受剥夺但在其他维度遭受多种剥夺的家庭纳入分析。表3汇总了农户总体、非建档立卡户和已脱贫建档立卡户多维脆弱性水平的测度结果。

表3 多维贫困脆弱性测度结果

调查时间	群体类别	多维贫困脆弱性广度			多维贫困脆弱性深度			多维贫困脆弱性强度		
		V_0	V_0^P	V_0^R	V_1	V_1^P	V_1^R	V_2	V_2^P	V_2^R
2021年	农户总体	0.099	0.007	0.093	0.024	0.006	0.018	0.010	0.006	0.004
	非建档立卡户	0.108	0.002	0.106	0.022	0.001	0.020	0.005	0.001	0.004
	已脱贫建档立卡户	0.090	0.012	0.077	0.027	0.012	0.015	0.015	0.012	0.003
2022年	农户总体	0.082	0.009	0.073	0.010	0.008	0.002	0.008	0.007	0.000
	非建档立卡户	0.097	0.004	0.092	0.005	0.003	0.002	0.002	0.002	0.000
	已脱贫建档立卡户	0.065	0.015	0.050	0.016	0.014	0.002	0.014	0.014	0.000
2023年	农户总体	0.056	0.011	0.045	0.011	0.010	0.001	0.009	0.009	0.000
	非建档立卡户	0.057	0.006	0.052	0.005	0.004	0.001	0.003	0.003	0.000
	已脱贫建档立卡户	0.055	0.018	0.037	0.018	0.017	0.001	0.016	0.016	0.000

从农户总体看,2021—2023年多维贫困脆弱性的广度分别为9.9%、8.2%和5.6%,与同年调研的风险群体相比,监测对象范围有所扩大。已脱贫建档立卡户中监测对象平均占建档立卡总人口的7%,符合《国务院扶贫开发领导小组关于建立防止返贫监测和帮扶机制的指导意见》(以下简称《建立意见》)中“监测对象规模一般为建档立卡人口的5%左右,深度贫困地区原则上不超过10%”^①的要求。从时间维度看,无论是农户总体,还是非建档立卡户和已脱贫建档立卡户,多维贫困脆弱性群体数量均逐年递减。这一测度结果意味着,防止返贫帮扶政策卓有成效。从脆弱性人群分类看,已脱贫建档立卡户多维贫困脆弱性水平高于非建档立卡户,原因在于非建档立卡户中包括边缘易致贫户和突发困难户两种监测对象,风险群体的总体比例较高。从多维贫困脆弱性的指数分解看,高风险脆弱性与边缘风险脆弱性的广度随时间变化的趋势相反,前者呈上升趋势,后者呈下降趋势。尽管这一现象看似与即时援助高风险群体以降低潜在风险的紧迫性不一致,实则表明高风险脆弱群体的绝对数量在减少,但减少的速度比边缘风险脆弱群体慢。背后的原因在于,边缘风险脆弱群体数量的减少趋势来自整体风险人口基数的缩减,当整体风险人口基数缩减时,处于边缘风险状态的个体比例自然也随之减少。从多维贫困脆弱性的广度、深度、强度看,其广度更具敏感性,而其深度和强度的变化幅度不明显。随着时间的推移,边缘风险脆弱人群的多维贫困脆弱性强度趋近于0。

各监测变量对多维贫困脆弱性的贡献率如表4所示。各监测变量的贡献率由高到低分为三个层次,第一类为收入水平、受教育程度、健康状态,第二类为劳动技能、养老保险,第三类为饮水安全、住房安全、卫生状况、入户路类型、医疗保险。收入水平、受教育程度和健康状态一直被公认为至关重要的监测指标,即便在等权重情况下,上述监测变量对多维贫困脆弱性的贡献率亦高于其他变量。边

^①参见《国务院扶贫开发领导小组关于建立防止返贫监测和帮扶机制的指导意见》, https://www.gov.cn/zhengce/zhengceku/2020-03/27/content_5496246.htm。

缘风险群体对收入水平有较高的敏感性，易受较大刚性支出或收入大幅缩减的影响。受教育程度低始终是农户无法在短期内有所改善的因素，健康问题依旧是农户返贫致贫的重要隐患，健康状态的改善能够显著提高家庭成员劳动参与率和劳动时间，减少因病致贫、因病返贫的风险（张鹏龙等，2024）。劳动能力是家庭经济收入的主要来源，缺乏劳动力是增加返贫致贫风险的重要原因。农村社会养老保险制度是促进农村老年群体实现“老有所养”的重要基本公共服务，养老金仍是多数农村老年群体的单一收入，为老年群体提供的物质保障比较有限，而且，养老保险在城乡间、区域间差距明显（文元全等，2024）。此外，实际调研发现，农户在饮水安全、住房安全、卫生状况、入户路类型、医疗保险等方面的条件已初步改善。以上指标虽对多维贫困脆弱性的贡献率不高，但在监测中仍不可忽视。

表4 多维贫困脆弱性各监测变量的贡献率 单位：%

维度	指标	各监测变量对多维贫困脆弱性的贡献率								
		2021年			2022年			2023年		
		总体	高风险脆弱群体	边缘风险脆弱群体	总体	高风险脆弱群体	边缘风险脆弱群体	总体	高风险脆弱群体	边缘风险脆弱群体
收入	收入水平	35.8	21.1	37.8	31.4	28.1	32.0	25.5	24.2	16.2
教育	受教育程度	30.9	30.8	30.9	26.9	29.3	26.6	24.4	26.6	16.2
健康	健康状态	20.7	34.0	19.0	23.8	29.2	23.0	25	28.1	32.4
生活条件	饮水安全	0.1	0.0	0.1	0.1	0.0	0.1	0.1	0.0	0.0
	住房安全	1.7	2.4	1.6	4.3	2.6	4.0	6.9	6.1	8.1
	卫生状况	1.1	0.8	1.1	0.8	0.0	0.7	1.4	0.8	4.1
	入户路类型	0.2	0.1	0.2	1.0	1.5	1.0	1.2	0.6	4.1
工作与社会保障	劳动技能	3.9	7.6	3.4	4.5	6.2	4.1	6.8	8.3	8.1
	医疗保险	0.2	0.5	0.2	0.2	0.4	0.1	0.1	0.3	0.0
	养老保险	5.4	2.7	5.7	7.0	2.7	8.4	8.6	5.0	10.8

（三）多维贝叶斯网络分类器精度检验

参照 Bielza et al. (2011) 和 Gil-Begue et al. (2021)，本文评估了多维贝叶斯网络分类器的类维度平均精度和整体精度。类维度平均精度是模型预测类维度变量（监测变量）准确度的均值，用 Acc_M 表示，计算公式如下：

$$\overline{Acc}_M = \frac{1}{M} \sum_{m=1}^M Acc_m \quad (12)$$

（12）式中， Acc_m 为类维度变量（监测变量） m 的预测精度，表示该变量正确分类的观测值个数与总观测值数的比值。由于超类变量（状态变量）由类维度变量构造而成，本文通过预测超类变量（状态变量）的精度来评估整体精度 Acc_{mp} ，定义如下：

$$Acc_{mp} = \frac{1}{N} \sum_{i=1}^N I_{mp} \quad (13)$$

(13) 式中, I_{mp} 是判别预测值是否正确的示性函数, 如果模型将观测家庭正确预测为监测对象或非监测对象, 则该函数等于 1, 否则等于 0。本文分别以 2021—2023 年调研数据的 80% 为训练集, 剩余样本为测试集进行测试, 计算得到类维度平均精度和整体精度的结果如表 5 所示。可以发现, 多维贝叶斯网络分类器整体精度均在 90% 以上, 一定程度上表明本文测算多维贫困脆弱性的方法可以很好地监测农户的返贫致贫风险。

	2021 年	2022 年	2023 年
类维度平均精度	83.7	83.9	83.1
整体精度	91.1	93.5	94.5

(四) 监测变量门槛设置合理性

在监测变量中, 收入水平的门槛受主观性影响较大。为了进一步验证门槛设置的合理性和监测方法的稳健性, 本文通过国家扶贫标准的 1.5 倍、OECD 国家收入比例标准、世界银行社会贫困线三种方式从低到高设置五个不同层次的收入门槛值, 对模型精度进行比较分析。

1. 国家扶贫标准的 1.5 倍。《建立意见》将人均可支配收入低于国家扶贫标准 1.5 倍左右的家庭纳入监测范围。2020 年的扶贫标准为人均年纯收入 4000 元^①, 该标准的 1.5 倍约为 6000 元。2021 年是巩固脱贫成果的第一年, 故本文以 6000 元作为 2021 年的监测标准。

2. OECD 国家常用的收入比例法。收入比例法具体有两种界定标准: 一种是不区分城乡的“全国一条线”方式, 统一界定相对贫困线; 另一种是区分城镇和农村的“城乡两条线”方式, 分别界定相对贫困线。在比例设置上, 大多数学者主张设置为基数的 40%~60% (李莹, 2021), 所以, 本文以江西省农村居民人均可支配收入中位数的 40%、50%、60% 为界定低、中、高三个人均相对贫困线的标准。

3. 世界银行的社会贫困线标准。世界银行将人均每天的消费低于 3.2 美元和 5.5 美元分别作为中等偏下收入国家和中等偏上收入国家的贫困线。中国属于中等偏上收入国家, 按照这个分类, 将人均每天消费 5.5 美元的标准换算为人民币后, 本文以每年人均可支配收入 14052 元为收入门槛值。

通过表 6 可以发现, 随着收入门槛值的提高, 模型精度有所提高。当收入门槛值达到人均可支配收入中位数的 50% 后, 进一步提高门槛值会使收入被剥夺概率大幅增长, 且对模型精度无明显影响。因此, 以人均可支配收入中位数的 50% 作为农户是否遭受收入剥夺的参考依据较为合理。

	国家扶贫标准的 1.5 倍	人均可支配收入中位数的 40%	人均可支配收入中位数的 50%	人均可支配收入中位数的 60%	世界银行中等偏上收入国家
收入门槛值 (元)	6000	6761	8451	10141	14052
收入被剥夺概率 (%)	7.8	9.1	21.1	34.6	67.3
模型精度 (%)	90.8	90.8	91.2	91.1	91.3

^①资料来源: 《2020 年底中国建档立卡贫困人口人均纯收入达 10740 元 远超国际极端贫困标准》, <https://baijiahao.baidu.com/s?id=1696268826517225237>。

（五）多种测度方法的结果对比

1. 多种分类模型评价结果对比。在探讨是否为监测对象时，根据真实类别与模型预测类别的组合情况，可以构建一个分类混淆矩阵，通过混淆矩阵计算准确率、精准率、召回率以及 F1 分数等关键评价指标。一般情况下，模型准确率越高，效果越好。本文将随机森林和神经网络两种机器学习方法的评价指标与多维贝叶斯网络分类器模型的评价指标进行对比，结果如表 7 所示。

表 7 不同机器学习方法模型效果（以 2021 年为例）

模型	准确率	精确率	召回率	F1 分数
随机森林	0.918	0.918	0.998	0.956
神经网络	0.902	0.931	0.965	0.947
多维贝叶斯网络分类器	0.914	0.917	0.997	0.955

注：①准确率是分类任务中最常用的性能指标，又称精度，与前文提到的精度计算方式一致。②精确率，又称查准率，表示在模型识别为正类的样本中，真正为正类的样本占比。③召回率，又称查全率，表示模型正确识别为正类的样本数量占总正类样本数量的比值。④F1 分数的计算公式为： $F1 \text{ 分数} = 2 \times \text{精确率} \times \text{召回率} / (\text{精确率} + \text{召回率})$ 。

可以发现，多维贝叶斯网络分类器在准确率、精准率、召回率和 F1 分数方面均有良好表现，能够根据变量的关联关系构建关系网络，在理解和解释变量之间关系方面具有明显优势。

2. 贫困领域多种监测方法结果对比。以 2021 年的数据为例，本文对比分析了 4 种主流贫困风险人群测度方法与多维贝叶斯网络分类器的关联和差异，用双色散点图可视化监测对象^①。可以发现，VEP 法和国家扶贫标准的 1.5 倍均以收入为单一监测维度来判断监测对象，所以易返贫致贫监测对象主要集中在低收入群体。A-F 多维贫困指数采用双界限法对多维贫困进行测度，该方法通过各福利维度临界值判断个体的多维贫困状况。多维特别关注指数借鉴了 A-F 多维贫困指数双界限法的原理，在一定程度上突出收入维度的重要性。在维度临界值 $K=4$ 的条件下，可以发现，用多维贝叶斯网络分类器测度的监测状态与用 A-F 多维贫困指数和多维特别关注指数两种方法测度的结果重合度较高。

五、农户监测状态变换的驱动因素与趋势研判

（一）农户监测状态变换情况

本文基于 2021—2023 年监测对象和非监测对象的变换情况，编制如表 8 所示相应的变换矩阵。

表 8 农户监测状态动态变换矩阵 单位：%

		2022 年		2023 年	
		监测对象	非监测对象	监测对象	非监测对象
2021 年	监测对象	3.21	6.86	0.60	9.47
	非监测对象	5.00	84.93	3.83	86.10
2022 年	监测对象			0.75	7.46
	非监测对象			3.69	88.10

^①具体结果见《中国农村经济》网站或中国知网本文附录图 1。

由表 8 可知，由于经济的稳步发展和农户收入水平的持续提升，纳入监测范围的农户规模总体呈下降趋势，成功退出监测范围的家庭数量有所上升。具体而言，2021 年，约 10%（3.21%+6.86%）的农户被识别为监测对象。2022 年，这一比例降为 8.21%（0.75%+7.46%）。2023 年，只有 4.43%（0.60%+3.83%）的农户被纳入监测范围。上述趋势表明，农户的返贫致贫风险有所降低。然而，从邻近调查年份看，监测对象和非监测对象的状态转移变化较为频繁。2021—2022 年，约 5%的家庭由非监测状态转变为监测状态，同时，6.86%的家庭则成功由监测状态转变为非监测状态，80%以上的农户始终保持在非监测状态。这些变动说明返贫风险具有动态性和阶段性，摆脱风险的同时也伴随新的返贫风险。本文将进一步揭示农户监测状态变换的频率、不稳定程度的演变规律及影响因素。

给定 $t-1$ 期、 t 期的监测状态， $t+1$ 期的状态取决于 t 期的监测状态。可将监测对象的状态分为以下类型：第一类为安全稳定户，三期均不是监测对象；第二类为轻度易返贫致贫户，在三期中只有一期为监测对象；第三类为中度易返贫致贫户，在三期中有两期为监测对象；第四类为重度易返贫致贫风险户，三期均为监测对象。本文将上述四种状态按顺序分别赋值为 1~4 的整数。由于监测状态的变换为有序变量，故可用有序 Logit 模型探讨农户风险监测状态变换的影响因素。此外，本文还引入了家庭结构特征，以分析家庭监测状态变换的关键因素，为制定相应政策提供理论依据。

（二）实证检验与分析

本文的实证分析以 2023 年的农户为研究样本，以监测模型预测的农户下一期的风险状态为研究对象，探究农户监测状态变换的影响因素。农户监测状态变换影响因素的估计结果和家庭人口结构的异质性分析结果如表 9 所示。人力资本对防止返贫及国家人口高质量发展具有重要意义，由表 9 可知，家庭人口数和非健康人数的增加都会加重农户的返贫致贫风险。稳定就业人数的变化对农户监测状态变换的影响不显著，可能与就业质量、家庭结构与负担、政策与制度因素、经济环境与市场变化等有关。农户的新增人口主要为非劳动力或低效率劳动力，这种增长对家庭经济状况的改善作用有限，如果中青年群体同时拥有未成年子女和老年亲属的双重抚养责任，可能因增加家庭消费而加剧返贫风险。家庭资产禀赋是衡量家庭抵御风险能力的重要标准，较高的家庭资产禀赋可以有效应对返贫致贫风险。

表 9 有序 Logit、稳健性检验和异质性分析结果

变量	被解释变量：农户监测状态变换类型					
	(1) 全体样本	(2) 变量替换	(3) 三代同堂	(4) 中青年和儿童	(5) 中青年和老年人	(6) 只有中青年
人口数变化	0.618*** (0.237)	0.626*** (0.235)	0.608** (0.298)	1.238 (0.821)	0.495 (0.568)	-10.271 (424.258)
健康情况变化	1.087*** (0.245)	0.487** (0.236)	0.215 (0.597)	1.209** (0.575)	1.407*** (0.494)	1.176*** (0.440)
稳定就业人数变化	-0.033 (0.076)	-0.030 (0.076)	0.090 (0.149)	0.133 (0.210)	-0.207 (0.156)	0.142 (0.160)
资产禀赋变化	-3.603** (1.818)	-3.933** (1.837)	-5.765 (4.911)	-5.555 (3.609)	2.374 (3.719)	-9.951* (5.435)

表9 (续)

住房条件	-0.232*** (0.069)	-0.223*** (0.069)	-0.242 (0.226)	-0.791* (0.459)	-0.072 (0.168)	-0.243*** (0.085)
政策导向	0.005 (0.022)	0.007 (0.022)	0.061 (0.060)	0.029 (0.053)	0.008 (0.049)	-0.045 (0.033)
合作组织	-0.672*** (0.221)	-0.672*** (0.221)	-0.353 (0.510)	-0.224 (0.611)	-1.718*** (0.619)	-0.309 (0.319)
耕地面积	-0.018 (0.021)	-0.018 (0.021)	-0.032 (0.042)	-0.146 (0.098)	0.007 (0.033)	0.003 (0.033)
交通便利程度	0.085 (0.085)	0.085 (0.086)	-0.214 (0.256)	0.116 (0.331)	-0.033 (0.247)	0.127 (0.110)
户主性别	-0.161 (0.152)	-0.174 (0.152)	-0.622 (0.384)	-0.342 (0.462)	-0.090 (0.297)	-0.096 (0.232)
户主年龄	-0.296*** (0.055)	-0.311*** (0.054)	-0.566*** (0.126)	0.053 (0.219)	-0.165 (0.124)	-0.317*** (0.092)
户主受教育程度	-0.783*** (0.041)	-0.797*** (0.041)	-1.050*** (0.122)	-0.910*** (0.136)	-0.716*** (0.080)	-0.696*** (0.065)
人口结构	-0.157*** (0.046)	-0.154*** (0.046)				
搬迁情况	0.448 (0.381)	0.458 (0.379)	-0.329 (1.188)	0.064 (1.135)	1.240** (0.521)	-0.367 (1.043)
所在村历史贫困状态	0.393*** (0.132)	0.383*** (0.131)	0.090 (0.325)	0.905** (0.384)	0.520* (0.275)	0.338* (0.203)
截距 1	-5.769*** (0.769)	-5.988*** (0.771)	-9.020*** (2.138)	-7.866*** (2.689)	-1.740 (1.636)	-6.637*** (1.686)
截距 2	-4.119*** (0.762)	-4.364*** (0.763)	-7.105*** (2.093)	-6.468** (2.665)	-0.191 (1.636)	-4.765*** (1.677)
截距 3	-0.975 (0.839)	-1.253 (0.840)	-3.854* (2.168)		2.111 (1.695)	-1.191 (1.934)
观测值数	3328	3328	585	297	922	1511
极大似然估计值	-997.941	-1021.238	-173.125	-120.452	-252.697	-430.141
卡方值	803.73	717.13	241.50	115.30	206.52	183.10
R ²	0.291	0.260	0.404	0.324	0.290	0.176

注：①***、**和*分别表示 1%、5%和 10%的显著性水平。②资产禀赋变化经标准化处理，住房条件和政策导向在回归中用的是对数值。③括号内为稳健标准误。

在物质资本层面，住房条件是衡量居住质量改善的重要指标，住房面积的扩大可减少因居住条件恶劣引发的潜在风险。在社会资本层面，发挥农民专业合作社、龙头企业等市场主体的作用，可增强村级集体经济对脱贫户的“输血”功能，辐射带动脱贫户增收，提高其自我发展能力、降低返

贫风险（董昕，2021）。政策导向对脆弱群体具有一定帮扶作用，但过度的政策补助会使农户形成政策的“悬崖效应”和“养懒汉”的依赖心理（黄薇和曹杨，2022），该变量正的估计系数意味着家庭获得更高政府资助时往往处于需要帮助的困境中，但该变量对返贫风险没有显著影响。在自然资本层面，耕地面积和距村主干道距离对返贫风险没有显著影响。原因在于，土地资源是有限的，难以根据生产需求随意增加耕地面积，而且，近年来土地抛荒现象越发严峻（余粮红等，2023），耕地面积与农业产出并不一定是正比关系。此外，随着交通工具的普及，地理距离对农户日常生活的实际影响也逐渐减弱。

从家庭层面看，户主的年龄和受教育程度是评估农户返贫致贫风险不可忽视的变量。年轻户主往往面临较大的需求压力，受较低的金融素养和资产配置能力、高风险偏好、社会保障体系不完善等因素的限制，家庭通常会面临较大的经济压力从而增加返贫致贫风险。户主的受教育程度在一定程度上决定了家庭教育资源的初始积累，会进一步影响劳动力质量和获取工资性收入的能力。是否为易地搬迁户对监测状态动态变换的影响在统计上不显著，可能的原因是，通过易地扶贫搬迁，搬迁户在安置住房、基础设施和公共服务等方面有了明显改善，已能够维持相对稳定的生活状态。原脱贫村农户更易受返贫风险的影响，他们的基础条件相对薄弱，经济稳定性不强。

（三）稳健性检验

1. 变量替换。将家庭健康水平变化替换为家庭健康人数变化，重新对主要结果进行检验，结果由表9（2）列所示，结论并未发生明显改变。

2. 检验方法替换。假设扰动项服从逻辑分布，采用有序 Probit 模型对监测状态动态变换诱因重新进行估计，结果由附录表2（2）列所示，结论并未发生明显改变^①。

3. 特殊样本剔除。为避免极端值的影响，本文对资产禀赋变化进行上下缩尾1%的处理后再次回归，结果如附录表2（3）列所示，除资产禀赋变化的显著性减弱外，其他结果未改变基本结论。

4. 内生性问题。针对遗漏变量、反向因果等导致的内生性，本文加入县级和村级层面可能的遗漏变量（如所在县自然灾害程度、所在村是否为脱贫村），结果如附录表2（4）列所示，同样不改变本文的基本结论。

（四）异质性分析

表9展示了不同群体监测状态变换的异质性。由表9（3）列可知，对三代同堂家庭而言，家庭人口数变化是影响其经济稳定性与贫困脆弱性的关键变量，表明了代际间资源分配与责任分担对家庭摆脱风险的重要性。由表9（4）列与表9（5）列可知，家庭所在村的原始贫困状态对有老年人和儿童的家庭具有深远影响。这两类群体的生理状况和社会基础往往较为脆弱，而且，原脱贫村在基础设施、社会服务等方面存在不足，双重脆弱性增加了此类家庭防止再次陷入贫困的挑战。由表9（6）列可知，对纯粹由中青年组成的家庭而言，金融资本和物质资本是抵御返贫致贫风险的重要因素。

^①检验方法替换、特殊样本剔除和内生性问题的具体结果见《中国农村经济》网站或中国知网本文附录表2（2）～（4）列。

六、结论与建议

在巩固拓展脱贫攻坚成果同乡村振兴有效衔接时期，防止规模性返贫至关重要。本文将各脆弱性诱发因素的依赖关系通过贝叶斯网络形成关联网络，利用调查数据进行实证分析，得到如下结论。

基于脆弱性理论的多维贝叶斯网络分类器方法可以较好地评估农户的返贫致贫风险。在江西省2021—2023年防止返贫监测的追踪调研数据中，农户的多维贫困脆弱性水平在10%以下。当整体风险人群规模缩小时，处于边缘风险状态个体的占比也随之减少，高风险群体的风险特征相对突出。该测量方法涵盖的弱势群体范围比实际陷入困境的人群更广，为“早预防”和“早干预”提供了充足的实施空间。进一步研究发现：农户监测状态的变换常伴随返贫致贫风险，人力资本和金融资本层面的家庭人口增加、家庭成员健康水平下降、家庭资产禀赋减少等更易使农户陷入高风险状态，参与社会组织、改善住房条件等则能有效削弱农户返贫风险。家庭人口结构对监测状态变换的影响具有异质性。

根据以上研究结论，本文对完善防止返贫致贫监测和帮扶机制给出如下建议。

第一，构建基于多源融合数据与大数据技术的防止返贫动态监测机制是提升帮扶工作效率与精准度的重要路径。在完善和优化返贫风险测度方法的同时，应利用行业、部门数据等多维度数据，结合先进数据挖掘和人工智能技术构建智能化防止返贫动态监测机制，变“人找人”为“数据找人”，找出潜在风险对象，做到“应纳尽纳”，找出稳定的风险消除群体，做到“应退尽退”。要解决因人工操作导致的效率低下、资源消耗大、易出错等难题，减轻基层工作负担，使帮扶工作更加科学、高效。

第二，依据监测对象的风险特征和发展需求，实施精准化帮扶。对风险单一的个体或群体，采取专项援助措施，防止其过度依赖福利资源，规避“福利陷阱”。对存在多重风险的复杂情况，则需采取因人而异、量身定制的综合性帮扶方案，确保帮扶措施的有效性和针对性，使风险在萌芽状态即得到有效遏制。在帮扶策略上，应坚持需求导向、精准施策原则，脱贫不稳定户、边缘易致贫户的脆弱性较强，要持续优化对此类群体的兜底保障政策，特别是要提高70岁以上老年群体农村养老保险补助标准，确保其不因老返贫。对不同劳动技能的群体也要针对性施策：对有劳动能力和劳动意愿的监测户，应通过提供劳动机会增强其内生动力，帮助其提高劳动能力，鼓励勤劳致富；对弱劳力、半劳力，要探索落实开发式帮扶措施，完善补助机制，减轻家庭负担；对整户无劳动能力的兜底保障户或因灾而出现突发返贫致贫风险的农户，可通过应急救助绿色通道实施紧急救助。

第三，关注易返贫群体内生发展能力，推进构建防止返贫长效机制。产业帮扶增收效果明显，政府可通过支持龙头企业、农民专业合作社等，带动易返贫群体入股、分红，共享发展成果。此外，通过加快建设高效灵活的劳动力市场打破就业壁垒和信息壁垒，完善灵活就业形态下劳动者的权益保障机制，畅通企业和易返贫群体的交流渠道。在打破就业壁垒和信息壁垒的同时，要做好后期就业情况追踪，切实保障易返贫群体的合法权益，提升其依靠自身力量抵御内外部风险的能力，减少

对政府补助的过分依赖。在后脱贫时代，依然要将健全风险保障体系与巩固拓展脱贫攻坚成果衔接，防止返贫，同时要关注非建档立卡户中的高风险脆弱性家庭，防止其陷入贫困，做到“早发现、早干预、早帮扶”。

虽然基于多维贝叶斯网络分类器的防止返贫监测方法的整体精度已达到90%以上，但仍有必要进一步完善和优化现有测度方法。未来的研究应着眼于对贝叶斯网络内部各变量间依赖关系的详尽剖析，注意与计量经济模型中协变量相互作用机理的结合，进一步思考规避设置单一阈值的突变效应，以提高防止返贫监测工作的精准度和时效性。

参考文献

- 1.董昕, 2021:《三峡库区(重庆段)农户稳定脱贫评价与返贫阻断机制研究》,重庆工商大学博士学位论文。
- 2.顾宁、刘洋, 2021:《产业扶贫降低了贫困农户的脆弱性吗》,《农业技术经济》第7期,第92-102页。
- 3.郭熙保、周强, 2016:《长期多维贫困、不平等与致贫因素》,《经济研究》第6期,第143-156页。
- 4.韩崢, 2004:《脆弱性与农村贫困》,《农业经济问题》第10期,第8-12页。
- 5.胡世文、曹亚雄, 2021:《脱贫人口返贫风险监测:机制设置、维度聚焦与实现路径》,《西北农林科技大学学报(社会科学版)》第1期,第29-38页。
- 6.黄薇、曹杨, 2022:《常态化精准扶贫政策的完善:反福利依赖的视角》,《经济研究》第4期,第172-190页。
- 7.贾男、王赫, 2022:《脱贫农户返贫风险防范政策研究》,《经济研究》第10期,第121-137页。
- 8.蒋南平、郑万军, 2017:《中国农民工多维返贫测度问题》,《中国农村经济》第6期,第58-69页。
- 9.兰宇、张鹏, 2024:《“脱贫不脱政策”何以防止脱贫农户返贫》,《中国农村经济》第5期,第167-184页。
- 10.李博、张全红、周强、Mark Yu, 2018:《中国收入贫困和多维贫困的静态与动态比较分析》,《数量经济技术经济研究》第8期,第39-55页。
- 11.李芳华、张阳阳、郑新业, 2020:《精准扶贫政策效果评估——基于贫困人口微观追踪数据》,《经济研究》第8期,第171-187页。
- 12.李锐、李佩伦、贾敏雪, 2024:《基本养老保险低收入参保者养老待遇适度吗?——基于替代率和缴费负担的研究》,《保险研究》第6期,第102-116页。
- 13.李莹、于学霆、李帆, 2021:《中国相对贫困标准界定与规模测算》,《中国农村经济》第1期,第31-48页。
- 14.李玉山、卢敏、朱冰洁, 2021:《多元精准扶贫政策实施与脱贫农户生计脆弱性——基于湘鄂渝黔毗邻民族地区的经验分析》,《中国农村经济》第5期,第60-82页。
- 15.刘胜强、赵雪雁, 2024:《生态脆弱区农户的返贫脆弱性评估及其影响因素——以陇南山区为例》,《生态学报》第3期,第1009-1023页。
- 16.刘子宁、郑伟、贾若、景鹏, 2019:《医疗保险、健康异质性与精准脱贫——基于贫困脆弱性的分析》,《金融研究》第5期,第56-75页。
- 17.吕光明、崔新新、孙伯驰, 2021:《防止返贫动态监测和精准帮扶的着力点——基于CFPS数据的实证分析》,《财政研究》第8期,第16-30页。

- 18.孟婷, 2023: 《返贫识别与成因探究——基于 Triple-Hurdle 模型的分析》, 《南方经济》第 12 期, 第 19-36 页。
- 19.宁静、殷浩栋、汪三贵、王琼, 2018: 《易地扶贫搬迁减少了贫困脆弱性吗? ——基于 8 省 16 县易地扶贫搬迁准实验研究的 PSM-DID 分析》, 《中国人口·资源与环境》第 11 期, 第 20-28 页。
- 20.平卫英、王佳, 2023: 《多维特别关注指数: 防止返贫致贫监测对象的识别与致贫影响因素研究》, 《统计研究》第 6 期, 第 118-133 页。
- 21.汪三贵、周园翔, 2022: 《构建有效的防规模性返贫的机制和政策》, 《农业经济问题》第 6 期, 第 12-22 页。
- 22.王兆林、王洁仪, 2024: 《易地扶贫搬迁户返贫风险因素识别及防范策略——基于武陵山区典型项目区的调查研究》, 《西南大学学报(社会科学版)》第 3 期, 第 166-179 页。
- 23.文元全、马鑫旺、高红, 2024: 《农村社会养老保险制度变迁、现存短板与实践进路——基于共同富裕目标前景的审思》, 《地方财政研究》第 4 期, 第 86-96 页。
- 24.谢治菊, 2023: 《防止规模性返贫的核心议题与研究转向》, 《中州学刊》第 12 期, 第 14-23 页。
- 25.闫啸、李录堂、李晗, 2022: 《宅基地退出降低了农户的贫困脆弱性吗? ——来自安徽金寨的证据》, 《中国土地科学》第 4 期, 第 38-48 页。
- 26.严小燕, 2022: 《贫困退出背景下的返贫脆弱性评价、预警及精准防贫机制》, 福建师范大学博士学位论文。
- 27.余粮红、高堃、高强, 2023: 《休戚与共: 土地托管企业与农户利益联结机制重塑》, 《农业经济问题》第 7 期, 第 49-63 页。
- 28.张鹏龙、钟建乐、胡羽珊, 2024: 《防止返贫帮扶政策效果评估——基于劳动收入的视角》, 《管理世界》第 3 期, 第 127-152 页。
- 29.张全红、李博、周强, 2019: 《中国农村的贫困特征与动态转换: 收入贫困和多维贫困的对比分析》, 《农业经济问题》第 12 期, 第 31-42 页。
- 30.章文光、吴义榕、宫钰, 2020: 《建档立卡贫困户的返贫风险预测及返贫原因分析——基于 2019 年 25 省(区、市)建档立卡实地监测调研数据》, 《改革》第 12 期, 第 110-120 页。
- 31.赵普、龙泽美、王超, 2022: 《规模性返贫风险因素、类型及其政策启示——基于西南民族地区的调查》, 《管理世界》第 11 期, 第 146-158 页。
- 32.周迪、王明哲, 2019: 《返贫现象的内在逻辑: 脆弱性脱贫理论及验证》, 《财经研究》第 11 期, 第 126-139 页。
- 33.周强, 2021: 《我国农村贫困的动态转化、持续时间与状态依赖研究——基于收入贫困与多维贫困的双重视角》, 《统计研究》第 10 期, 第 90-104 页。
- 34.左停、李泽峰, 2022: 《风险与可持续生计为中心的防返贫监测预警框架》, 《甘肃社会科学》第 5 期, 第 35-46 页。
- 35.Alkire, S., and J. Foster, 2011, "Counting and Multidimensional Poverty Measurement", *Journal of Public Economics*, 95(7-8): 476-487.
- 36.Bielza, C., G. Li, and P. Larrañaga, 2011, "Multi-dimensional Classification with Bayesian Networks", *International Journal of Approximate Reasoning*, 52(6): 705-727.
- 37.Chambers, R., 1989, "Editorial Introduction: Vulnerability, Coping and Policy", *ID Bulletin*, 20(2): 1-7.

38. Dercon, S., 2001, "Assessing Vulnerability", https://www.researchgate.net/publication/228378379_Assessing_Vulnerability.
39. DFID, 2000, "Sustainable Livelihoods Guidance Sheets", <https://www.livelihoodscentre.org/-/sustainable-livelihoods-guidance-sheets>.
40. Foster, J., J. Greer, and E. Thorbecke, 1984, "A Class of Decomposable Poverty Measures", *Econometrica*, 52(3): 761-776.
41. Gallardo, M., 2013, "Using the Downside Mean-semideviation for Measuring Vulnerability to Poverty", *Economics Letters*, 120(3): 416-418.
42. Gallardo, M., 2018, "Identifying Vulnerability to Poverty: A Critical Survey", *Journal of Economic Surveys*, 32(4): 1074-1105.
43. Gallardo, M., M. E. Santos, P. Villatoro, and V. Pizarro, 2024, "Measuring Vulnerability to Multidimensional Poverty in Latin America", *Review of Income and Wealth*, 70(3): 661-691.
44. Gil-Begue, S., C. Bielza, and P. Larrañaga, 2021, "Multi-dimensional Bayesian Network Classifiers: A survey", *Artificial Intelligence Review*, Vol.54: 519-559.
45. Hu, S., Y. Ge, M. Liu, Z. Ren, and X. Zhang, 2022, "Village-level Poverty Identification Using Machine Learning, High-resolution Images, and Geospatial Data.", *International Journal of Applied Earth Observation and Geoinformation*, Vol.107, 102694.
46. Ligon, E., and L. Schechter, 2003, "Measuring Vulnerability", *The Economic Journal*, 113(486): 95-102.
47. Mahanta, R., and D. Das, 2021, "Vulnerability to Poverty: A Survey", *Journal of Business Strategies*, 32(2): 151-172.
48. Messner, J. W., P. Pinson, J. Browell, M. B. Bjerregard, and I. Schicker, 2020, "Evaluation of Wind Power Forecasts-An Up-to-date View", *Wind Energy*, 23(6): 1461-1481.
49. Moser, C., 1998, "The Asset Vulnerability Framework: Reassessing Urban Poverty Reduction Strategies", *World Development*, 26(1): 1-19.
50. Ogryczak, W., A. Ruszczyński, 2001, "On Consistency of Stochastic Dominance and Mean-semideviation Models", *Mathematical Programming*, Vol.89: 217-232.
51. Pearl, J., 1985, "Bayesian Networks: A Model of Self-Activated Memory for Evidential Reasoning", http://ftp.cs.ucla.edu/tech-report/198_-reports/850017.pdf.
52. Polsky, C., R. Neff, and B. Yarnal, 2007, "Building Comparable Global Change Vulnerability Assessments: The Vulnerability Scoping Diagram", *Global Environmental Change*, 17(3-4): 472-485.
53. Povel, F., 2015, "Measuring Exposure to Downside Risk with an Application to Thailand and Vietnam", *World Development*, Vol.71: 4-24.
54. Zhang, R., Y. He, W. Cui, Z. Yang, J. Ma, H. Xu, and D. Feng, 2022, "Poverty-Returning Risk Monitoring and Analysis of the Registered Poor Households Based on BP Neural Network and Natural Breaks: A Case Study of Yunyang District, Hubei Province", *Sustainability*, 14(9), 5228.
55. Zou, Q., S. Xie, Z. Lin, M. Wu, and Y. Ju, 2016, "Finding the Best Classification Threshold in Imbalanced Classification", *Big Data Research*, Vol.5: 2-8.

Dynamic Monitoring to Prevent the Return to Poverty and Exploring the Causes of the Change in Monitoring Status: Analysis Based on the Multi-dimensional Bayesian Network Classifier Model

PING Weiyang^{1,2} GUO Yutang¹ HUANG Fei¹

(1. School of Statistics and Data Science, Jiangxi University of Finance and Economics;

2. Applied Statistics Research Center, Jiangxi University of Finance and Economics)

Summary: Based on the urgent need for dynamic monitoring and assistance to prevent poverty return in China at the current stage, it is of great significance to build an efficient and reliable monitoring and early warning mechanism for the risk of returning to poverty to prevent a large-scale return to poverty and consolidate and expand the achievements of poverty alleviation. Based on the data of the longitudinal survey in Jiangxi Province from 2021 to 2023, this paper uses a multi-dimensional Bayesian network classifier model to predict the risk of rural households returning to poverty. On this basis, this paper measures the multi-dimensional poverty vulnerability level of rural households through the standard downward half-deviation adjusted probability, thus more effectively identifying the monitoring objects of governance of returning to poverty after 2020. To explore the frequency of monitoring state transformation, the evolution law of instability degree, and influencing factors, the risk triggers of the dynamic transformation of the monitoring state of the farmers to prevent the return of poverty are analyzed using the ordered Logit model.

This paper finds that the overall prediction accuracy of the multidimensional Bayesian network classifier model based on the vulnerability theory for the monitoring status of farmers is more than 90%. In addition, the overall multidimensional poverty vulnerability level of rural households is below 10%, covering a wider range of vulnerable groups than those who are actually in distress. Further research shows that the reduction in human capital and financial capital is an important factor for rural households to get rid of the risk of returning to poverty while participating in social organizations and improving housing conditions can effectively reduce the risk of returning to poverty. Based on the above conclusions, this paper proposes the following suggestions. First, it is necessary to improve the monitoring mechanism for preventing the return to poverty based on multi-source fusion data and big data technology. Second, the implementation of precise assistance should be based on the risk characteristics and development needs of the monitored groups. Third, attention should be paid to the endogenous development capacity of groups prone to returning to poverty, and the construction of a long-term mechanism for preventing the return to poverty should be promoted.

The possible marginal contributions of this paper are in two aspects. First, using a multi-dimensional Bayesian network classifier, the dependency relationship of each monitoring variable is intuitively depicted and the conditional probability value of the Bayesian network is dynamically updated using real-time data to achieve rolling prediction of the monitoring objects. Second, this paper explores the influencing factors of livelihood capital on the multi-period monitoring status changes of farmers and combines the household head and family structure characteristics to capture the diverse needs of farmers during the monitoring status changes.

Keywords: Dynamic Monitoring to Prevent the Return to Poverty; Multi-dimensional Bayesian Network Classifier; Downside Risk; Dynamic Transformation

JEL Classification: I32; I38; C53

(责任编辑：马太超)